5-2016

Investigation of Novel Functions for DNA Damage Response and Repair Proteins in Escherichia coli and Humans

Benjamin A. Hilton
East Tennessee State Universtiy

Investigation of Novel Functions for DNA Damage Response and Repair Proteins in *Escherichia coli* and Humans

---

A dissertation

presented to

the faculty of the Department of Biomedical Sciences

East Tennessee State University

In partial fulfillment

of the requirements for the degree

Doctor of Philosophy in Biomedical Sciences,

Biochemistry & Molecular Biology Concentration

---

by

Benjamin A. Hilton

May 2016

Yue Zou, Ph.D. Chair

Phillip R. Musich, Ph.D.

Robert V. Schoborg, Ph.D.

Antonio E. Rusiñol, Ph.D.

Krishna Singh, Ph.D.

Keywords: Nucleotide Excision Repair, UvrABC, XPC, XPA, ATR, Apoptosis

ABSTRACT

Investigation of novel functions for DNA Damage Response and Repair proteins in *Escherichia coli* and Humans

by

Benjamin A. Hilton

Endogenous and exogenous agents that can damage DNA are a constant threat to genome stability in all living cells. In response, cells have evolved an array of mechanisms to repair DNA damage or to eliminate the cells damaged beyond repair. One of these mechanisms is nucleotide excision repair (NER), which is the major repair pathway responsible for removing a wide variety of bulky DNA lesions. Deficiency, or mutation, in one or several of the NER repair proteins is responsible for many diseases, including cancer. Prokaryotic NER involves only three proteins to recognize and incise a damaged site, while eukaryotic NER requires more than 25 proteins to efficiently recognize and incise a damaged site. XPC-RAD23B (XPC) is the damage recognition factor in eukaryotic global genome NER. The association rate of XPC to damaged DNA has been extensively studied; however, our data suggests that the dissociation of the XPC-DNA complex is the rate-limiting step in NER. The factor that verifies DNA-damage downstream of XPC is XPA. XPA also has been implicated in binding of ds-ssDNA junctions and has been found to bind at or near double-strand break sites in the premature aging syndrome Hutchinson-Gilford progeria (HGPS). This role for XPA is outside of its known function in NER and suggests that XPA may bind at

collapsed replication forks in HGPS that are unprotected due to a lack of binding by replication proteins. Along with XPC and XPA, a*taxia telangiectasia* and Rad3-related (ATR) is activated in response to DNA damage and initiates the cell cycle checkpoint pathway to rescue cells from genomic instability. We found that ATR functions outside of its known role in the checkpoint-signaling cascade. Our data demonstrate that ATR can rescue cells from apoptosis by inhibiting cytochrome c release at the mitochondria though direct interaction with the outer mitochondrial membrane and the proapoptotic protein tBid. The role of ATR in apoptosis is regulated by Pin1, which can change the structure of ATR at the backbone level. All of the results presented here suggest novel roles for DNA repair proteins in the maintenance of genome stability.

DEDICATION

I dedicate this manuscript to my family and friends whose support has made this achievement possible. Without my parents', Fred and Lisa, and grandmother's, Jacqueline Frye, continuous support and encouragement I would not have been able to complete my Ph.D.

ACKNOWLDEGMENTS

Thank you to Drs. Yue Zou and Phillip Musich for all their guidance, instruction, and patience through my graduate career at East Tennessee State University. Thank you Dr. Musich for giving me a solid foundation to build my career as a scientist on and to Dr. Zou for giving me direction through my time as a graduate student. I also appreciate the members of my committee for their support of my education. A special thank you to Antonio Rusiñol who was always next-door and willing to listen to and guide me through the minutia. In addition, I want to thank Hui Wang, Moises Serrano, and especially Zhengke Li for their friendship and critical input during my graduate career. I would also like to thank the following for their help in my research: Dr. Bongsup Cho, University of Rhode Island, for synthesizing adducted DNA substrates and advice in completion of the NER projects, Dr. Mamuka Kvaratskhelia, Ohio State University, for bringing me to his laboratory to learn mass spectrometry, and Dr. Kun Pin Lu, Harvard University, for antibodies and reagents that were pertinent to completion of the ATR project.

TABLE OF CONTENTS

Page

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS

aa, amino acid

AAF, *N*-acetyl-2-acetylaminofluorene

AF, aminofluorene

AP, 1-aminopyrene

AT, ataxia telangiectasia

ATM, ataxia telangiectasia mutated

ATP, adenosine triphosphate

ATR, ATM- and Rad3-related

ATRIP, ATR interacting protein

Bad, Bcl-2-associated death promoter

Bak, Bcl-2 homologous antagonist/killer

Bax, Bcl-2–associated X protein

BcL-2, B-cell lymphoma 2

Bid, BH3 interacting-domain death agonist

Bim, Bcl-2-like protein 11

BER, base excision repair

bp, base pair

BPDE, benzo[*a*]pyrene diol epoxide

Cdk, cyclin-dependent kinase

Cdk1, cyclin-dependent kinase 1

ChIP, chromatin immunoprecipitation

Chk1, checkpoint kinase-1

Chk2, checkpoint kinase-2

CPD, cyclobutane pyrimidine dimer

CPT, camptothecin

Cyp40, cycolphilin 40

CyPA, cycolophilin A

Cyto, cytoplasm

CytoC, cytochrome C

DAPI, 4',6-diamidino-2-phenylindole

DBD, DNA binding domain

DDB, damaged DNA binding protein

DDR, DNA damage response

DMEM, Dulbecco's modified Eagle medium

DMSO, dimethyl sulfoxide

DNA, deoxyribonucleic acid

DSB, double strand break

dsDNA, double-stranded DNA

ds/ssDNA, double-strand/single-strand DNA

DTT, dithiothreitol

EDTA, ethylenediamine tetraacetic acid

EMEM, Eagles minimum essential medium

EMSA, electromobility shift assay

ERCC1, excision-repair cross-complementing group 1

FAF, fluoro-(2)-aminofluorene

FBS, fetal bovine serum

FTI, farnesyltransferase inhibitor

GAPDH, glyceraldehyde-3-phosphate dehydrogenase

GFP, Green fluorescent protein

GGR, global genomic repair

GST, glutathione s-transferase

HEAT, huntingtin, elongation factor 3, protein phosphatase 2A and yeast TOR1

HEPES, N-2-hydroxyethylpiperazine-N'-2-ethanesulfonic acid

HGPS, Hutchinson-Gilford progeria syndrome

HR, homologous recombination

HU, hydroxyurea

IF, immunofluorescence

IP, immunoprecipitation

k_a, association rate

k_d, dissociation rate

K_D, equilibrium dissociation constant

kDa, kilodalton

LAΔ50, progerin or lamin A with 50 aa deletion

MALDI-TOF, matrix-assisted laser desorption time of flight

MCM7, minichromosome maintenance complex component 7

MEF, mouse embryonic fibroblasts

Mito, mitochondria

MMR, mismatch repair

MS, mass spectrometry

MHSP70, mitochondrial heat shock protein 70

MS/MS, tandem mass spectrometry

m/z, mass/charge ration

NER, nucleotide excision repair

NHEJ, nonhomologous end-joining

NHS-biotin, *N*-hydroxysuccinimidobiotin

NMR, nuclear magnetic resonance

nt, nucleotide

PARP-1, poly[ADP-ribose]polymerase 1

PBS, phosphate buffered saline

PCNA, proliferating cell nuclear antigen

PDB, Protein Data Bank

PI, propidium iodide

PIKK, phosphoinositide 3-kinase-like kinase

Pin1, peptidylprolyl *cis*/*trans* isomerase NIMA-interacting 1

PVDF, polyvinylidene difluoride

Q-TOF, quadrupole-time of flight

RAD23B, UV excision repair protein RAD23 homolog B

RD, restrictive dermopathy

RFC, replication factor C

RNAi, RNA interference

RPA, replication protein A

SDS-PAGE, sodium dodecyl sulfate-polyacrylamide gel electrophoresis

siRNA, small interfering RNA

SSB, single-stranded DNA binding

ssDNA, single-stranded DNA

T4-PNK, T4 polynucleotide kinase

TBE, Tris-Borate-EDTA

TCR, transcription coupled repair

TFIIH, transcription factor II H

TTD, trichothiodystrophy

UV, ultraviolet

UV-C, ultraviolet C

WT, wild-type

XP, xeroderma pigmentosum

XPA, XPB etc., xeroderma pigmentosum group A, B etc.

XPC-RAD23B, xeroderma pigmentosum protein C-RAD23B

CHAPTER 1

INTRODUCTION

DNA Damage and Cellular Responses

Endogenous and exogenous agents that can damage DNA are a constant threat to genome stability in all living cells. These agents include reactive oxygen radicals that are byproducts of many metabolic processes and environmental factors such as ultraviolet (UV) radiation and chemotherapeutics. When DNA damage-induced chemical and structural changes to the DNA occur the ultimate outcome is genomic instability and cancer (Kolodner et al. 2002). To avoid this fate, cells have evolved a multitude of damage sensors that initiate a DNA damage response that can effectively determine whether the damage is capable of being repaired or not (Zhou and Elledge 2000; Sancar et al. 2004). The response that is elicited is crucial in the prevention of disease. It has been estimated that 80-90% of human cancers are the result of DNA damage (Doll and Peto 1981; Sancar and Hearst 1993).

There are four major DNA repair pathways initiated in response to DNA damage: DNA damage checkpoint-mediated cell cycle arrest, activation of transcriptional programs, DNA repair, or apoptosis (Figure 1-1) (Sancar et al. 2004). If the repair machinery determines that the damage is repairable the DNA damage checkpoint proteins will activate the cell cycle checkpoint, which in turn arrests cell cycle progression. This stall in the cell cycle alters the transcription program allowing time for the DNA repair machinery to remove the damage, restoring genomic integrity and preventing mutation. Following repair, the transcription machinery is reactivated to fill-in the correct bases. When the

damage is deemed unrepairable the apoptosis pathway is activated, which eliminates the effected cells. When the DNA damage is not correctly addressed through checkpoint

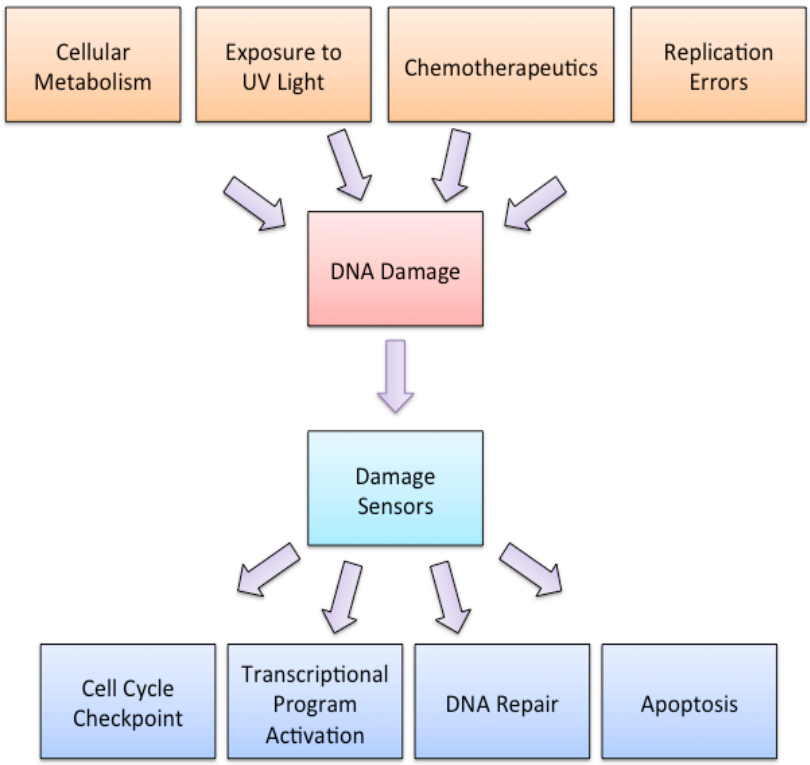


Figure 1-1. Molecular mechanisms of mammalian DNA damage responses (Modified from Sancar, 2004).

activation followed by repair or clearance via apoptosis, accumulation of genomic mutations and instability that promotes tumorigenesis may occur which could lead to the development of disease.

There are many ways that damage can occur to the DNA. These include base damage (oxidized or chemically adducted), DNA backbone damage (abasic sites and single- or double-strand DNA breaks), or cross-links (DNA-protein cross-links and interstrand/intrastrand cross-links) (Postel-Vinay et al. 2012). To address the variety of DNA damage five different DNA repair pathways can be activated including direct repair, base excision repair (BER), mismatch repair (MMR), double-strand break (DSB) repair, and

nucleotide excision repair (NER). Each of these pathways is further divided into subpathways that specialize in clearance of specific damage to the DNA. My dissertation research has focused on NER, prokaryotic and eukaryotic, and the DNA damage checkpoint protein ataxia telangiectasia and Rad3-related (ATR); therefore, NER and the DNA damage checkpoints will be introduced as follows.

Nucleotide Excision Repair

Nucleotide excision repair (NER) is the major repair pathway responsible for removing a wide variety of structurally and chemically unrelated bulky DNA lesions such as UV-induced pyrimidine dimers and chemically-adducted nucleotides. In humans, defects in proteins involved in NER result in diseases such as *Xeroderma pigmentosum* (XP) and trichothiodystrophy (TTD) (Hoeijmakers 2001; Thoma and Vasquez 2003). For instance, the phenotype for XPA deficiency is extreme UV-sensitively and often result in a shortened lifespan due to the development of skin cancer (Bootsma and Hoeijmakers 1991). The NER pathway is common between prokaryotes and eukaryotes with a similar general mechanism (Petit and Sancar 1999). NER occurs through a five-step process: 1. DNA damage recognition, 2. dual incisions flanking the damage site, 3. release of the damaged oligonucleotide, 4. DNA synthesis to fill the gap, and 5. ligation of the newly synthesized DNA into the helix. The ability of the NER machinery to recognize and eliminate a wide range of different types of damage lies in the fact that the initial damage-recognition factors sense the distortion to the Watson-Crick DNA structure induced by the adduct (Friedberg et al. 2005). Even though the prokaryotic and eukaryotic systems have similar mechanisms

there is no evidence that they are evolutionarily linked; therefore, there are no protein homologs between the two systems.

Prokaryotic NER

In prokaryotes the UvrABC protein complex carries out the major steps in NER. These three distinct proteins serve to recognize DNA damage, unwind the DNA flanking the adduct, and incise the damaged oligonucleotide(s) via dual incisions, at 3' then 5' to the damage. This mechanism is summarized in Figure 1-2 (Van Houten 1990; Y. Zou and Van Houten 1999). There are two major ways in which NER is activated in prokaryotes. First, the damage can be directly recognized and processed by UvrA in complex with UvrB. Otherwise, RNA polymerase (RNAP) stalls at a damage site and activates the transcriptional-repair coupling factor, which can dislodge RNAP and recruit the $UvrA_2B$ complex to the damage for repair (Kisker et al. 2013). In both cases, repair of the damage site is initiated by homodimerization of UvrA into $UvrA_2$ (Orren and Sancar 1990). $UvrA_2$ then forms a complex with UvrB ($UvrA_2B$) that recognizes and unwinds the DNA flanking the damage. The unwinding occurs in a two-step process that requires hydrolysis of ATP (Pakotiprapha et al. 2008). The current available data suggests that the $UvrA_2$ component of the $UvrA_2B$ complex initially recognizes distortion of the Watson-Crick structure of DNA (Van Houten and Kad 2014 Jan 25). Once helical distortion is identified, the $UvrA_2B$ complex opens the strands at which point UvrB verifies the presence of the adduct through strand invasion (Y. Zou and Van Houten 1999; Y. Zou et al. 2001). Strand opening is believed to facilitate the formation of a stable UvrB-DNA complex as well as initiate the

dissociation of $UvrA_2$ from the damage site (Oh and Grossman 1986; Shi et al. 1992). Two molecules of UvrC then are recruited through direct interaction with the DNA-scaffold formed by the UvrB-DNA complex resulting in the formation of the $UvrBC_2$ pre-incision complex. Formation of the $UvrBC_2$-DNA complex is believed to trigger the 3' and 5' endonuclease activity of $UvrBC_2$ at the phosphodiester bonds 8 nucleotides 5' and 4-5 nucleotides 3' to the damaged site (Y. Zou et al. 1995). The post-incision complex, $UvrBC_2$ bound to the damaged oligonucleotide, is displaced by UvrD/DNA polymerase I (Pol I) following the cleavage event (Van Houten 1990; Y. Zou et al. 1995). Pol I then fills-in the resulting gap and the newly created repair patch is ligated into the helix (Caron et al. 1985; Yeung et al. 1986). The UvrABC nuclease system represents an efficient mechanism for the removal of a variety of damage to the DNA and has been well studied as a model for nucleotide excision repair and its interaction with damaged DNA (Van Houten 1990).

Eukaryotic NER

Although similar to prokaryotic NER described above, the mechanism behind eukaryotic NER incorporates several major differences. Foremost, while prokaryotic NER involves only three proteins to recognize and incise a damaged site, eukaryotic NER involves more than 25 proteins to efficiently recognize and incise a damaged site (Riedl et al. 2003; Thoma and Vasquez 2003; C.-J. Park and Choi 2006; Scharer 2013). Additionally, while the prokaryotic NER proteins are specific to repair of damaged DNA, the proteins involved in some factors of eukaryotic NER also are involved in replication, recombination,

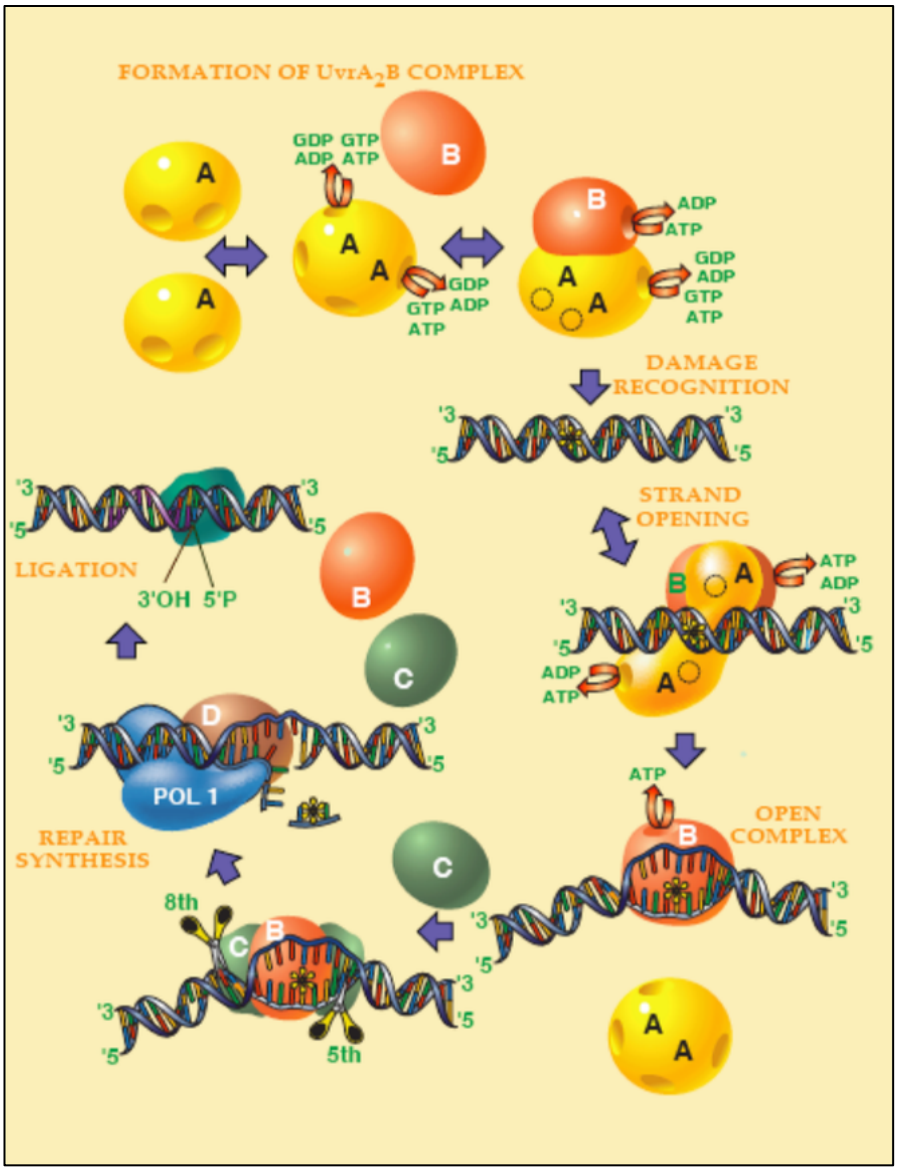


Figure 1-2: Prokaryotic nucleotide excision repair (Modified from Van Houten, 1991). Prokaryotic NER is performed by the UvrABC nuclease system and is ATP-dependent. In two-steps the $UvrA_2B$ complex senses DNA damage, identifying both distortion in the DNA-duplex and the adduct. Finally, $UvrA_2B$ unwinds the DNA flanking the damage site. Strand opening facilitates dissociation of $UvrA_2$ and formation of a stable UvrB-DNA complex. UvrC then is recruited via interaction with UvrB to form the $UvrBC_2$ pre-incision complex. $UvrBC_2$ then makes the dual incisions to remove the damaged oligonucleotide.

gene transcription and many other pathways outside of NER (Sancar et al. 2004; Marteijn et al. 2014).

There are two subpathways in eukaryotic NER: global genomic repair (GGR) and transcription-coupled repair (TCR). These subpathways are defined by how the damage is encountered. In GGR-NER the repair machinery surveys the entire genome for sites of DNA damage. In TCR-NER, if replication has stalled during active transcription on the chromatin DNA damage repair factors are recruited to repair the damage so that replication can continue. The general mechanism of NER is summarized in Figure 1-3 (Scharer 2013). Eukaryotic NER is highly ordered and involves orchestration of all 25+ factors to successfully recognize, incise, excise, repair, and ligate a damage site.

In GGR-NER the XPC-RAD23B protein complex (henceforth, XPC) scans the genome for destabilized duplex DNA. XPC then binds to the damage site to initiate DNA-damage repair. XPC binding is essential in that XPC is required to recruit the downstream factors to the damage site. Although XPC is sufficient to recognize DNA damage, damaged DNA binding complex (DDB) and centrin2 protein stimulate UV-induced damage recognition when bound to XPC (Nishi et al. 2005; Takedachi et al. 2010). Following damage-site recognition, transcription factor IIH (TFIIH), replication protein A (RPA), XPD, XPB and XPA are recruited to verify the damage site and unwind the DNA in an ATP-dependent manner in preparation for dual-incision. After the DNA containing the damage site is unwound, the endonucleases XPG and XPF-ERCC1 (excision-repair cross-complementing group 1) are sequentially recruited to the damage site. Binding of XPG/XPF-ERCC1 releases the XPC and initiates the dual-incisions at downstream then upstream to the damage by XPG and XPF-ERCC1, respectively. Upon excision of the damaged strand XPA, TFIIH, XPG and XPF-ERCC1 dissociate from the site while RPA remains bound to protect the single-stranded DNA (ssDNA) through synthesis of the new strand. In TCR, damage repair is initiated when RNA

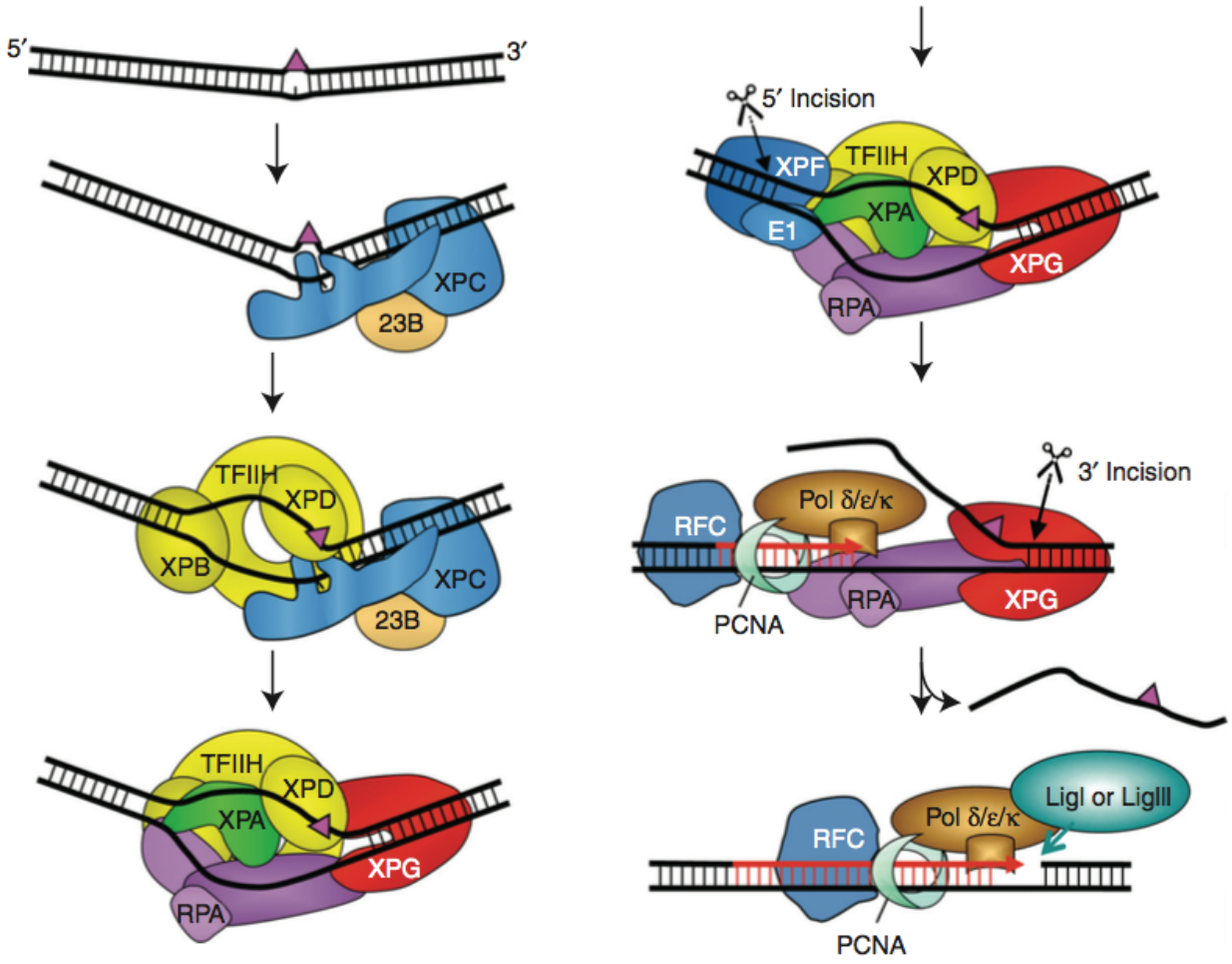


Figure 1-3: General Mechanism of Eukaryotic NER (Modified from Scharer, 2013). GG-NER occurs through a mechanism similar to that described for the UvrABC nuclease system. Damage recognition, dual incision, excision, and re-synthesis of the damaged patch are performed by a series of protein complexes, pictured above, in which factors dissociate or are recruited depending on the stage of repair.

polymerase II becomes stalled at a damaged nucleotide. The subsequent TCR-NER process requires the same proteins as GGR-NER with the exception of the XPC in the initiation step (Riedl et al. 2003; C.-J. Park and Choi 2006; Scharer 2013).

The Role of XPA in Laminopathy-Based Premature Aging

DNA damage accumulation and DNA repair defects have been linked to the phenotypes associated with premature aging and also have been suggested to be a potential cause of normal aging (Lombard et al. 2005). Often, DNA repair defects are a result of mutations within one of the genes encoding a DNA repair protein (Lieber and Karanjawala 2004). Hutchinson-Gilford progeria syndrome (HGPS) is a premature aging disease caused by defective maturation of lamin A and is characterized by DNA repair defects (B. Liu et al. 2005; Varela et al. 2005; Y. Liu et al. 2006). Lamin A along with lamin C and lamin B form a scaffold underlying the inner nuclear membrane collectively called the nuclear lamina (Goldman et al. 2002; Dechat et al. 2010). The meshwork of lamina in the nucleus serves as a mechanical support for the nucleus; however, it also serves in other roles such as chromatin organization, DNA replication, and protein signaling (Moir et al. 1995; Moir et al. 2000). HGPS is a rare (1 in 4 million cases) and ultimately fatal disease. HGPS patients exhibit retarded growth, partial lipodystrophy, osteoporosis, osteolytic lesions, thin skin, micrognathia, and premature atherosclerosis (Eriksson et al. 2003). Most of the patients die in their teens from complications from myocardial infarction or stroke (Gonzalo and Kreienkamp 2015).

The severe phenotype observed in HGPS is due to a *de novo* point mutation (1824C → T) in *LMNA* gene (De Sandre-Giovannoli et al. 2003; Eriksson et al. 2003). The mutation in *LMNA* leads to the sporadic activation of a cryptic splice donor site in the pre-mRNA of prelamin A, producing a farnesylated-carboxymethylated truncated lamin A protein named progerin or LAΔ50. The splice donor site activation is known to cause deletion of 50 amino acids near the C-terminus that includes an endoprotease (Zmpste24) cleavage site required

for maturation of normal lamin A (Musich and Y. Zou 2009). The consequence of this mutation is the accumulation of progerin in patient cells. Interestingly, a low level of progerin also is produced in cells of healthy individuals and the progerin level increases with age and has been correlated with aging-related human diseases (Aliper et al. 2015).

Past studies have shown that the genomic integrity is compromised and that DNA repair is defective in human HGPS fibroblasts and Zmpste24-deficient mouse embryonic fibroblasts (Zmpste24$^{-/-}$ MEFs) (B. Liu et al. 2005; Varela et al. 2005; Y. Liu et al. 2006; Musich and Y. Zou 2009; Musich and Y. Zou 2011). However, recent work has established that an active p53 signaling pathway exists in human HGPS fibroblasts and Zmpste24-deficient MEF (Varela et al. 2005; Y. Liu et al. 2008). Our past data suggested that the accumulation of DNA damage in these patient cells is not due to mutations within repair machinery genes; rather, the strand breaks that are generated by progerin are not repaired by the double-strand break (DSB) repair machinery. We know that XPA accumulates at sites of DSBs; however, the mechanism is not fully understood (Y. Liu et al. 2008; Musich and Y. Zou 2011).

Xeroderma pigmentosum Group A

XPA is a member of the *Xeroderma pigmentosum* (XP) family of proteins, each of which is lacking in patients with XP. XPA patients lacking functioning XPA protein have the most severe phenotype of all of the XP syndromes (Cleaver and States 1997). XPA is a 32kDa zinc metalloprotein that is required for NER, specifically, the damage verification step following damage-site binding by XPC (C.H. Park and Sancar 1994; Lehmann 2003; Y. Liu et al. 2005; Nasir et al. 2013). XPA has been shown to bind to damaged duplex DNA

with low affinity and specificity; however, XPA is required for both GGR- and TCR-NER (Cleaver and States 1997; Cortez et al. 2001).

Comparing the role XPA plays in damage verification to its indispensable involvement in both subpathways of NER suggests a larger role for XPA outside of its known functions (S.H. Lee et al. 1995; Cleaver and States 1997; Batty and Wood 2000). Past work has demonstrated that XPA binds with high affinity to non-damaged DNA containing double-strand/single-strand DNA junctions (ds-ssDNA) (Yang et al. 2006). XPA also is thought to be involved in the recruitment of other NER proteins to the damage site. This is illustrated by the fact that XPA interacts with many of the proteins downstream of the verification step (Shell and Y. Zou 2008). Observing binding of XPA to a bulky lesion elucidated the DNA-binding domain; however, the low binding affinity to this type of substrate and the high affinity to non-damaged ds-ssDNA junctions (intermediates formed during DNA replication and repair) suggest that there may be additional contacts outside of

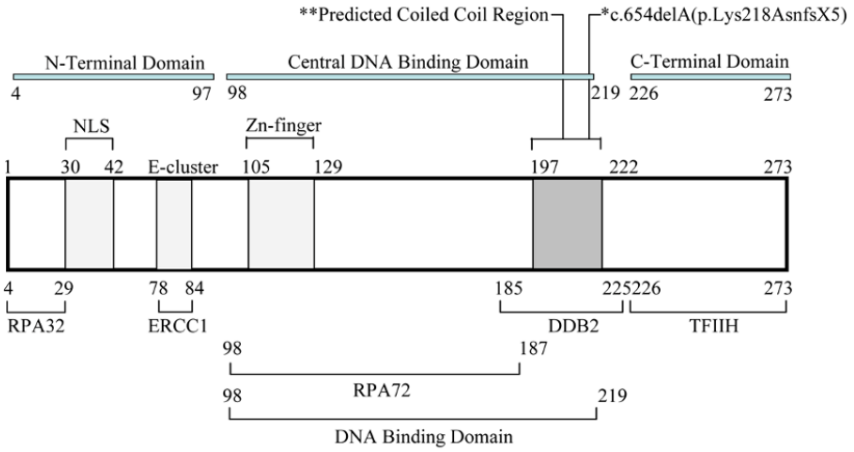


Figure 1-4: Schematic of XPA domains (Modified from Nasir, 2013). XPA contains a DNA binding domain in the central part of the protein and protein interaction domains at the termini and within the DNA binding domain. XPA interacts with several of the proteins involved in DNA replication as indicated above.

the known domain involved in junction binding. The experimentally verified domains within the XPA protein, including the DNA-binding domain are shown in Figure 1-4.

The high-affinity binding to junction DNA also may shed light on the mislocalization of XPA to DSB sites in HGPS patient fibroblasts. XPA may bind at these sites, much like binding to repair intermediates, to stabilize the replication fork and prevent fork collapse, while at the same time preventing rapidly accumulating DSBs in HGPS that leads to early replicative senescence in these cells.

Replication Factor C1

The replication factor C complex has been shown to consist of five subunits, one large subunit of approximately 140 kDa and four small (36-40 kDa) subunits (Ellison and Stillman 1998). Replication factor C1 (RFC1) is the catalytic subunit of the RFC complex. RFC1 contains domains that are indicative of proteins involved in cell cycle regulation and DNA metabolism. Furthermore, the same domains are found in poly (ADP-ribose) polymerase (PARP), which is known to be involved in binding and sensing ssDNA breaks, suggesting that RFC1 also may play a role in sensing nicks (Hashiguchi et al. 2007). It is not entirely clear what role the extended N-terminal domain plays; however, it has been demonstrated that RFC1 interacts directly with proliferating cell nuclear antigen (PCNA) and with DNA (Rhéaume et al. 1997; Dionne et al. 2008; Chen et al. 2009). PCNA promotes the recruitment and binding of many of the proteins known to be involved in DNA replication and repair.

In eukaryotes, PCNA must be loaded onto dsDNA by an alternative protein due to its stability as a homotrimeric ring structure. Of the RFC complex subunits, RFC1 exhibits the

most contact with PCNA and positions the DNA so that PCNA can be loaded (Bermudez et al. 2003). Complete DSB repair requires synthesis of DNA to rejoin the newly synthesized DNA. To restart replication at the arrested fork RFC1 is involved through two pathways, Rad51-mediated recombination or Polδ-dependent synthesis, both resulting in branched intermediates (Moriel-Carretero and Aguilera 2010). RFC also has been shown to recruit and stimulate polymerases supporting the evidence that RFC serves as more than just a clamp loader (Ogi et al. 2010). In HGPS cells XPA is bound at the ds-ssDNA junction left by stalled replication. In addition to XPA being bound at the junctions, RFC1 is no longer bound and proteolyticly cleaved. Cleavage of RFC1 is one reason XPA accumulates at the stalled replication forks in HGPS. The DNA-damage response machinery can repair chemically-induced damage in HGPS but not the progerin-induced damage. The lack of repair at progerin-induced damage is likely due to XPA binding at ds-ssDNA junctions, which prevents binding of not only DSB repair proteins but also the replication machinery in HGPS.

Ataxia Telangiectasia and Rad3-related

An additional pathway designed to defend against mutations caused by insults to the genome in eukaryotic cells is the DNA damage checkpoint pathway. The DNA damage checkpoints are a series of pathways that coordinate DNA damage repair and regulate the cell cycle to maintain genomic integrity (Zhou and Elledge 2000; Sancar et al. 2004). The DNA damage checkpoint consists of four components: sensors, mediators, signal transducers, and effectors (Figure 1-5) (Sancar et al. 2004). Ataxia telangiectasia mutated (ATM) and ataxia telangiectasia and Rad3-related (ATR) are protein kinases in the

phosphoinositide 3-kinase-like kinase (PIKK) family, that play central roles in initiating the checkpoints in response to DNA damage (Abraham 2001; Harper and Elledge 2007). ATM is activated primarily in response to DSBs (Shiloh 2003); in contrast, ATR is activated by a wide range of DNA damage and stalled replication (Abraham 2001; L. Li and L. Zou 2005).

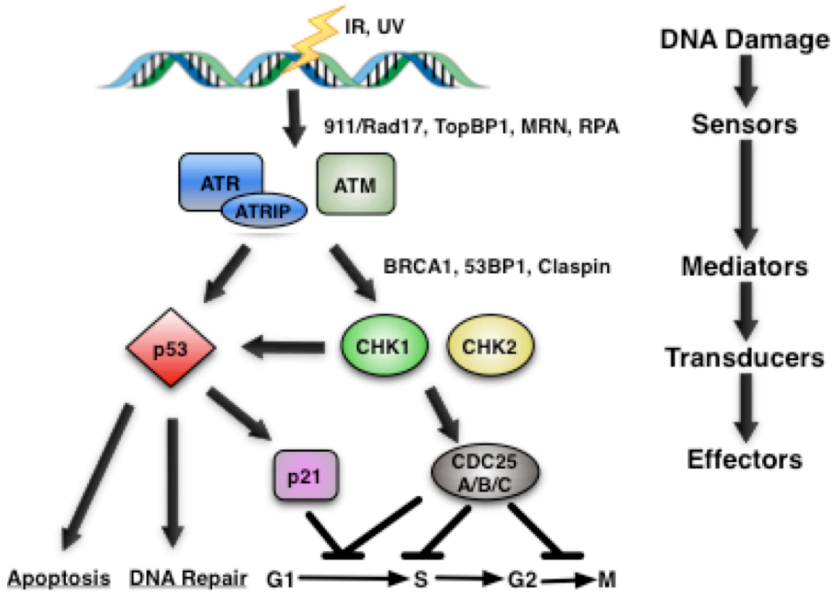


Figure 1-5: Schematic of DNA damage checkpoint activation following IR/UV irradiation (Modified from Li, 2013).

Following sensing of the DNA damage by ATM and ATR, phosphorylate the signal-transducer proteins, also kinases, Chk1 and Chk2, initiating a signaling cascade in which downstream targets, such as p53, alter cell cycle progression (Sancar et al. 2004; L. Li and L. Zou 2005).

ATR in the DNA Damage Response and Apoptosis

ATR is an essential gene as demonstrated by the fact that knockout of ATR in mice is embryonically lethal (Brown and Baltimore 2000; de Klein et al. 2000). Interestingly, cells

collected from ATR-knockout mice die of apoptosis (Brown and Baltimore 2000; Cortez et al. 2001). There is evidence that links ATR to apoptosis through the DNA damage checkpoints (Heffernan et al. 2009). However, it is of note that knockout of ATM is not embryonically lethal and promotes cell survival and carcinogenesis (Bartkova et al. 2005). Additionally, inhibition of ATR was shown to enhance apoptosis through a p53-independent mechanism (Heffernan et al. 2009; Myers et al. 2009; Kawasumi et al. 2011) regardless of ATR kinase activation (Huntoon et al. 2013). To explain how ATR can regulate such events outside of the traditional role in cell cycle checkpoints several mechanisms have been proposed; however, all of the current studies have been on ATR/chromatin interaction or replication regulation (Shechter et al. 2004; Petermann et al. 2006; Petermann et al. 2010).

Furthermore, a review of human epidemiologic and mouse model studies demonstrate that caffeine consumption significantly decreases the risk of cancer (Sarkaria et al. 1999; Brown and Baltimore 2003; Abel et al. 2007; Rees et al. 2007). Caffeine is known to inhibit the cell cycle checkpoint (Sarkaria et al. 1999); however, caffeine can block the checkpoints independently of ATM/ATR inhibition (Cortez 2003). Given that suppression of ATM has been shown to increase carcinogenesis the decrease in cancer risk could be due to functions of ATR outside of the kinase activity. Taken together the evidence points to an anti-apoptotic activity for ATR outside of checkpoint activity such that when ATR is inhibited carcinogenesis is suppressed. ATR is an attractive target for chemotherapy given the potential antitumor activity of ATR inhibition. By better understanding the molecular mechanism through which ATR inhibition suppresses carcinogenesis novel therapies can be explored.

Peptidylprolyl Cis/Trans Isomerase, NIMA-Interacting 1

Peptidylprolyl *cis/trans* isomerase, NIMA-interacting 1 (Pin1) is a critical regulator of numerous key biological processes (K.P. Lu et al. 1996; Driver and K.P. Lu 2010; Liou et al. 2011; Z. Lu and Hunter 2014). Dysfunction of Pin1 has been related to human diseases such as cancer, neurodegeneration, aging, heart disease, *etc* (P.J. Lu et al. 1999; Zheng et al. 2002; Driver and K.P. Lu 2010; Nakamura et al. 2012). Knockdown of Pin1 has been shown to increase cell survival which is interesting in that Pin1 regulates many proteins in a variety of pathways (Shen et al. 2009). Previous work implicates Pin1 in the regulation of the apoptotic response through modulating p53-mediated checkpoint control (Zheng et al. 2002). Further work demonstrated that p53 directly interacts with pro-apoptotic BAX to promote apoptosis (Follis et al. 2015). Specifically, Pin1 recognizes phosphorylated Ser/Thr-Pro bonds in certain proteins leading to isomerization. In most cases the phosphorylation is a prerequisite to Pin1 binding and isomerization. Pin1 catalytically regulates protein conformation, *cis* or *trans*, to control protein function. This change in confirmation also can be accompanied by a translocation event, such is the case with p53 isomerization which is isomerized and translocated from the nucleus to the mitochondria (Zheng et al. 2002). Previous studies have suggested that conformational changes, an effect of Pin1 isomerization, regulate a wide range of protein functions such as protein interactions, subcellular localization, and protein stability (Zacchi et al. 2002; Steger et al. 2013). Therefore, Pin1 may play a role in changing the backbone architecture of protein targets to allow for new protein-protein interactions. Pin1 is regulated by phosphorylation, which suppresses Pin1 isomerization activity. One inactivating kinase of interest is death associated protein kinase 1 (DAPK1) which was observed previously to inactivate Pin1

activity following UV-induced DNA damage (T.H. Lee et al. 2011; Kim et al. 2014). The phosphorylation of S71 in Pin1 is in the catalytic active site and prevents isomerization, hence, change of protein target function (Z. Lu and Hunter 2014).

Questions to be Answered in these Studies

The UvrABC nuclease system has been studied for decades and has given insight into how both prokaryotic and eukaryotic cells cope with constant insults to the genome. The mechanism by which prokaryotic cells incise damage sites is well understood; however, more information is needed to elucidate the mechanisms involved in the eukaryotic system. By utilizing the excellent model provided by the simple and established UvrABC system investigations into the specific mechanism(s) of NER can be illuminated. The studies presented in this dissertation have confirmed that DNA damage-recognition in NER is not sensing of the damaged base. Rather, DNA damage-recognition is the sensing of thermodynamic instability on the duplex that is generated by the damage. My studies implementing the UvrABC system to investigate the effects of DNA sequence, structure and lesion clustering on damage recognition in NER were published over the course of this dissertation. The first study demonstrates a novel 3'-next-flanking sequence effect on NER processing of bulky arylamine lesions and the second indicates the importance of base stacking and related thermal and thermodynamic destabilization in the repair of bulky cluster arylamine lesions. This work has been published in Nucleic Acids Research (2012, 41:869-880) and Chemical Research in Toxicology (2013, 26:1251-1262) and is highlighted in Chapters 2 and 3, respectively.

In eukaryotes, the XPC-RAD23B protein complex (XPC) is the unequivocal initial damage recognition factor in GGR-NER. XPC is required for assembly of the NER core factors and progression through the GGR-NER pathway both *in vitro* and *in vivo* (Evans 1997; Riedl et al. 2003; Tapias et al. 2004). The NER pathway can recognize a wide range of substrates with one set of enzymes, orchestrated by XPC, including UV-induced photoproducts, adducts formed by exogenous mutagens such as benzo[*a*]pyrene, various aromatic amines such as bulky arylamines, endogenous oxidative lesions such as cyclopurines and adducts formed by cancer chemotherapeutic drugs such as cisplatin (Friedberg 2005; Gillet and Scharer 2006). These lesions do not share chemical structure similarity; therefore, much like in the case with the UvrABC system, thermodynamic destabilization of DNA-duplex that can disrupt base-pairing, bending and flexibility contribute to damage-recognition by XPC which is related to repair outcomes (Mocquet et al. 2007; Mu et al. 2012; Yeo et al. 2012). Recent work has attempted to further understand the molecular mechanisms behind XPC recognition of damage sites. The data suggest that XPC residence-time on the strand opposite the damage is extremely important to producing an efficient-incision and that this correlates with the thermodynamic destabilization (Yeo et al. 2012; Mu et al. 2015 Aug 13). To further define the interaction of XPC with damaged DNA duplex surface plasmon resonance (SPR) was implemented to determine the XPC-DNA association and dissociation rates which were then correlated with a human NER repair assay. These data demonstrates that the XPC-DNA dissociation rate is an important factor in XPC residence-time and directly effects repair efficiency. This work is currently under review and is summarized in Chapter 4.

Following recognition of DNA damage XPC recruits several proteins including XPA, which is thought to be the major factor in DNA-damage verification (C.-J. Park and Choi 2006; C.-L. Li et al. 2015). An interesting observation that was made previously in our lab was that XPA has a higher affinity for binding to ds-ssDNA junctions than for DNA-damage sites (Yang et al. 2006). In addition to this finding there is evidence that XPA colocalizes with DSB sites in HGPS patient fibroblasts (Y. Liu et al. 2008). Taken together, these data suggest that XPA may play an important role outside of its known function as a DNA-damage verifier and serve to protect intermediate DNA fork structures that are present during DNA repair and DNA synthesis. By using a modified gel mobility shift assay, mass spectrometry footprinting and a limited proteolysis assay I found that XPA binds to ssDNA-overhang and Y-shape structures using additional surface contacts outside of the defined DNA-binding domain. Further characterization of XPA-junction binding is presented in Chapter 5 and was published in Bioscience Reports (2014, 34:e00162).

The finding that XPA can bind ds-ssDNA junctions of repair intermediates shed light on work in our lab that suggests that XPA is mislocalized to DSB sites in HGPS patient fibroblasts (Y. Liu et al. 2008). Using a modified chromatin immunoprecipitation assay (ChIP) XPA was found in association with the long-lasting DSBs generated by progerin but not at breaks generated by DNA-damaging agents. Previously, it was unclear whether XPA was found at these sties due to protein-protein interaction or if XPA was binding the DNA-junction. Coupled with our structural analysis we proposed that XPA plays a role in binding to ds-ssDNA junctions to prevent cell death induced by collapsed replication forks in HGPS. To test this hypothesis, immunofluorescence, chromatin association and a modified ChIP

assay were implemented. This work is presented here in Chapter 7. The manuscript is currently in progress and will be submitted soon.

RFC1 has been shown previously to localize to sites of DNA damage along with XPA and also has been identified as a substrate for ATR in response to DNA damage (Rouse and Jackson 2002; Matsuoka et al. 2007; Gourdin et al. 2014). RFC1, the largest subunit of the RFC complex, exhibits catalytic activity and is a component of the post-repair complex. In aged HGPS patient fibroblasts RFC1 is no longer associated with the chromatin which is in agreement with previous studies that found the replication capacity in aged HGPS is greatly diminished with cells in a senescent-like state (Y. Liu et al. 2006). My work demonstrates that RFC1 is proteolyticly cleaved as HGPS cells age in culture. The results are summarized in Aging Cell (2012, 11:363-365) and presented in Chapter 6.

ATR previously has been thought to be involved in apoptosis in response to replicative stress (Heffernan et al. 2009; Myers et al. 2009). However, the mechanism in which ATR acts as an antiapoptotic factor is not yet defined. The involvement of checkpoint activity in apoptosis of ATR cannot be ignored; nevertheless, ATR recently has been shown to be involved in kinase-independent antiapoptotic functions in cancer cells exposed to DNA-damaging agents (Huntoon et al. 2013). ATR is thought also to promote apoptosis when initiation of DNA replication fails due to low doses of UV-damage (Al-Khalaf et al. 2012). The mechanism behind the antiapoptotic function of ATR was still unclear. Therefore, we hypothesized that ATR may play a more direct role in apoptosis suppression. Our results suggest that ATR contains a BH3-like domain, commonly found in apoptotic regulator proteins, which promotes binding to Bid in response to DNA damage and prevents Bid-mediated release of cytochrome c. This interaction and important function for

ATR is further regulated by Pin1 at the S428-P429 site in a structurally disordered region of ATR. The data presented in this dissertation in Chapter 8 proposes that following DNA damage ATR can interact quickly and directly with the apoptotic machinery to suppress apoptosis. This data is published in Molecular Cell (2015, 60:35-46) and is presented in Chapter 8.

CHAPTER 2

UNUSUAL SEQUENCE EFFECTS ON NUCLEOTIDE EXCISION REPAIR OF ARYLAMINE LESIONS: DNA BENDING/DISTORTION AS A PRIMARY RECOGNITION FACTOR

Abstract

The environmental arylamine mutagens are implicated in the etiology of various sporadic human cancers. Arylamine-modified dG lesions were studied in two fully paired 11-mer duplexes with a -G*CN- sequence context, in which G* is a C8-substituted dG adduct derived from fluorinated analogs of 4-aminobiphenyl (FABP), 2-amino-fluorene (FAF) or 2-acetylaminofluorene (FAAF), and *N* is either dA or dT. The FABP and FAF lesions exist in a simple mixture of 'stacked' (S) and 'B-type' (B) conformers, whereas the *N*-acetylated FAAF also samples a 'wedge' (W) conformer. FAAF is repaired three to four times more efficiently than FABP and FAF. A simple A- to -T polarity swap in the G**CA*/G**CT* transition produced a dramatic increase in *syn*-conformation and resulted in 2- to 3-fold lower nucleotide excision repair (NER) efficiencies in *Escherichia coli*. These results indicate that lesion-induced DNA bending/thermodynamic destabilization is an important DNA damage recognition factor, more so than the local S/B-conformational heterogeneity that was observed previously for FAF and FAAF in certain sequence contexts. This work represents a novel 3'-next flanking sequence effect as a unique NER factor for bulky arylamine lesions in *E. coli*.

Introduction

Structural and conformational damage in specific areas of the genome can trigger tumorigenesis. For example, disruption of a gene that encodes the tumor suppressor p53 protein has been found in the majority of sporadic human cancer (1). Although human cells are equipped with repair pathways to safeguard the genome from various DNA damage, some lesions may go unrepaired, thereby serving as a faulty template to produce a complex array of mutations and genomic instability, ultimately leading to cancer initiation (2).

Arylamines and heterocyclic amines are well-known environmental mutagens/carcinogens, which have been implicated in the etiology of breast, liver and bladder cancers in humans (3). Metabolic activation of these amines *in vivo* produces C8-substituted dG as the major DNA adducts (4). For example, the human bladder carcinogen 4-aminobiphenyl produces ABP (Figure 1a). Similarly, AF and AAF are the major DNA adducts derived from 2-aminofluorene, 2-nitrofluorene and 2-acetylaminofluorene (Figure 1a). The ABP and AF adducts in fully paired duplex DNA have been shown to adopt an equilibrium of two prototype conformers: 'B-type', in which the carcinogen resides in the major groove of a relatively unperturbed double helical DNA, and 'stacked (S)', in which the carcinogen is base displaced and the glycosidic linkage to the modified guanine is *syn* (Figure 1c) (5, 6). The aromatic moieties of ABP are not coplanar as in AF, which results in a much lower S-state population than AF. AF-induced S/B-heterogeneity is dependent on the flanking sequence, which modulates mutational and repair outcomes (6, 7). AAF is chemically identical to AF except for a single acetyl group on the central nitrogen (Figure 1a), leading to sampling of an additional W-conformation, in which the fluorene moiety is

in the minor groove along with a *syn* glycosidic linkage (Figure 1c) (7, 8). The B and S conformers exhibited by AAF are similar to those obtained for ABP and AF.

Nucleotide excision repair (NER) is the major cellular pathway for removing bulky DNA lesions in cells. Accumulated evidence suggests that efficiency of NER is governed by various structural, cellular and biological factors (9-11). Sequence context, in particular, plays an important role in NER of bulky DNA lesions (7, 12). The most notable sequence effects were observed in the *Nar*I sequence (5'-. . .CG1G2CG3CC. . .-3'), which is well known for inducing higher frequency of −2 deletion mutations when adducted by AAF at G_3 position despite the similar chemical reactivities of three guanines (13, 14). Fuchs and coworkers have shown that AAF in duplex is an excellent substrate for *Escherichia coli* UvrABC and human exonuclease repair systems (15-18). They reported that in *E. coli*, the relative repair efficiencies of AAF at G_1, G_2 and G_3 were in a ratio of 100:18:66, respectively, whereas the human exonuclease exhibited 38:100:68 ratio (17, 18). Mu et al. (19) have recently carried out a human NER study of these lesions in HeLa cell extracts and found similar sequence-dependent NER efficiencies. Their molecular dynamics (MD) simulation data indicated that the greater NER efficiencies are correlated with base sequence–dependent local untwisting and minor groove opening together with weaker stacking interactions (19). Recently, we conducted *E. coli* UvrABC NER studies on the *Nar*I sequence duplexes (5'–$G_1G_2CG_3CC$–3'), in which guanines are modified by either AF or AAF (7). Results showed that the bulky AAF adducts repair in a conformation-specific manner, with the highly S/W-conformeric G_3 and G_1 duplexes incised considerably more efficiently than the highly B-conformeric G_2 duplex ($G_3 \sim G_1 > G_2$). Conversely, the repair rate of *N*-deacetylated AF was 2- to 3-fold lower than AAF, and the order of incision efficiencies was

opposite of that observed for the AAF case. We have coined the term '*N*-acetyl factor' to describe the complexity of NER recognition of AF versus AAF (7).

Here, we describe an unusual 3'-next flanking base effect on the conformational properties and *E. coli* NER efficiencies of three prototype arylamine adducts in the G**CN* sequence context (Figure 1b: G*=ABP, AF, or AAF; *N*=A or T). Results from spectroscopy (^{19}F NMR and induced circular dichroism [ICD]), thermodynamic quantification (differential scanning calorimetry [DSC]/ultraviolet (UV)-melting experiment) and gel electrophoresis, as well as MD/potential of mean force (PMF) calculations, show that sequence-dependent lesion-induced DNA bending coupled with thermodynamic destabilization is responsible for the altered repair recognition of bulky arylamine-DNA adducts in *E. coli*. This work represents a novel 3'-next flanking sequence effect as a unique NER factor for bulky arylamine lesions in *E. coli*.

Materials and Methods

Adduct synthesis

Modified duplexes were prepared following the published procedures (7, 8, 20-23). The modified oligos were characterized by electrospray time-of-flight mass spectrometry analysis as reported previously (24). An identical set of unmodified duplexes was also prepared as controls.

Differential scanning calorimetry

DSC measurements were performed using a VP-DSC Micro-calorimeter from Microcal Inc. (Northampton, MA) according to the procedures published previously (22). All sample solutions were 0.12 mM concentration. T_m was the temperature at half the peak area. ΔG and ΔS values were determined by the procedures of Chakrabarti et al. (25). The uncertainties in the values of Tm, ΔH, ΔG and ΔS represent the random errors inherent in the DSC measurements.

UV-Melting (Cary100 Bio, Beckman) and Circular Dichroism (CD) (J-810, Jasco) experiments were performed using the previously reported procedures (7, 20, 23, 26).

Dynamic ^{19}F NMR

Duplex samples (about 20-30 ODS) were dissolved in 300 ml of pH 7.0 buffer (100 mM NaCl, 10 mM Na_3PO_4 and 100 μM EDTA in 10% D_2O/90% H_2O) and filtered into through a Shigemi tube using a 0.2-mm membrane filter. All ^1H and ^{19}F NMR experiments were conducted using a dedicated 5-mm $^{19}F/^1H$ dual probe on a Bruker DPX400 Avance spectrometer operating at 400.0 and 376.5 MHz, respectively. Imino proton spectra were obtained using phase-sensitive jump-return sequences at 5°C. ^{19}F NMR spectra were acquired in the ^1H-decoupled mode and referenced to $CFCl_3$ by assigning external C_6F_6 in C_6D_6 at −164.90 ppm. Temperature dependence spectra were processed as reported previously (20, 27).

EMSA assay

The *N*-(20 -deoxyguanosin-8-yl)-4-fluoro-4-aminobiphenyl (FABP), *N*-(20 -deoxyguanosin-8-yl)-7-fluoro-2-amino-fluorene (FAF) and *N*-(20-deoxyguanosin-8-yl)-7-fluoro-2-acetylaminofluorene (FAAF)-modified 19-mer GCT and GCA sequences were each (100 nM) annealed with an equimolar complementary sequence, in which the 5'-end was γ-^{32}P-labeled using T4 polynucleotide kinase and [γ-^{32}P] ATP (Perkin-Elmer radiochemical, Boston, MA) in a buffer containing NaCl (25 mM) and Tris–HCl (25 mM). The mixture was heated at 95°C for 5 min and then cooled to room temperature overnight. The duplexes were subjected to 15% non-denaturing poly- acrylamide (acrylamide:bisacrylamide::29:1, w/w) gel electrophoresis at 1800 V, and the temperature was maintained at 4–8°C by regularly replacing the running buffer with the ice-cold Tris/Borate/EDTA (TBE) buffer. Gel was exposed to Kodak phosphor imaging screen overnight and scanned on Typhoon 9410.

Nucleotide excision assay

DNA substrates of 58 bp containing a single FABP, FAF or FAAF, each adducted at either G*CT or G*CA sequences, were constructed as described previously (28, 29). UvrA, UvrB and UvrC proteins were overexpressed in *E. coli* and then purified as described previously (30). The 5'-terminally labeled DNA substrates were incubated and incised by UvrABC as described previously (28, 29). Briefly, the DNA substrates (2 nM) were incubated in the UvrABC reaction buffer (50 mM Tris– HCl, pH 7.5, 50 mM KCl, 10 mM $MgCl_2$, 5 mM DTT) at 37°C in the presence of UvrABC (10 nM UvrA, 250 nM UvrB and 100 nM UvrC). The Uvr proteins were diluted and premixed in Uvr storage buffer before

addition to the reaction. Aliquots were collected at 0, 5, 10, 15 and 20 min into the reaction. The reaction was terminated by heating at 95°C for 5 min. The products were denatured by addition of formamide loading buffer and heating to 95°C for 5 min, followed by quick chilling on ice. The incision products were then analyzed by electrophoresis on a 12% polyacrylamide sequencing gel under denaturing conditions with Tris/Borate/EDTA (TBE) buffer.

To quantify the incision products, radioactivity was measured using a Fuji FLA-5000 Image Scanner with MultiGauge V3.0 software. The DNA incised (in femtomoles) by UvrABC was calculated based on the total molar amount of DNA used in each reaction and the ratio of the radioactivity of incision products to total radioactivity of DNA. At least three independent experiments were performed for determination of the rates of incision.

MD and PMF calculations

PMF calculations were performed on the GCA and GCT 11-mers initiated from the canonical B form of DNA for 'anti' simulations where the glycosidic bond is in the *anti* form. 'Syn' simulations were initiated from models based on an NMR structure (PDB:1C0Y) in which the glycosidic bond is in the *syn* form. MD simulations were performed with the programs CHARMM and NAMD, using the CHARMM27 additive nucleic acid force field. Modified G* lesions were created based on the CHARMM General Force Field followed by additional optimization of the dihedral parameters linking the G base to the adduct. Determination of the PMFs followed the protocol of Banavali and MacKerell (31) with details of the simulations included in the supporting information.

Results

Model systems

As model systems, 11-mer DNA duplexes [d(5'-CCATCG*CNACC-3').d(5'-GGTNGCGATGG-3')] were prepared, in which G* is FABP, FAF or FAAF and *N* is either dA or dT (designated as G**CA* and G**CT* duplexes, respectively) (Figure 1b). The two sequences are chemically isomeric, differing only on the polarity of the 3'-next flanking A:T versus T:A. The utility of fluorine-tagged lesions as effective structure probes has been documented (32). Both the G**CA* and G**CT* sequences have been used previously for the studies of bulky adducts (22, 33).

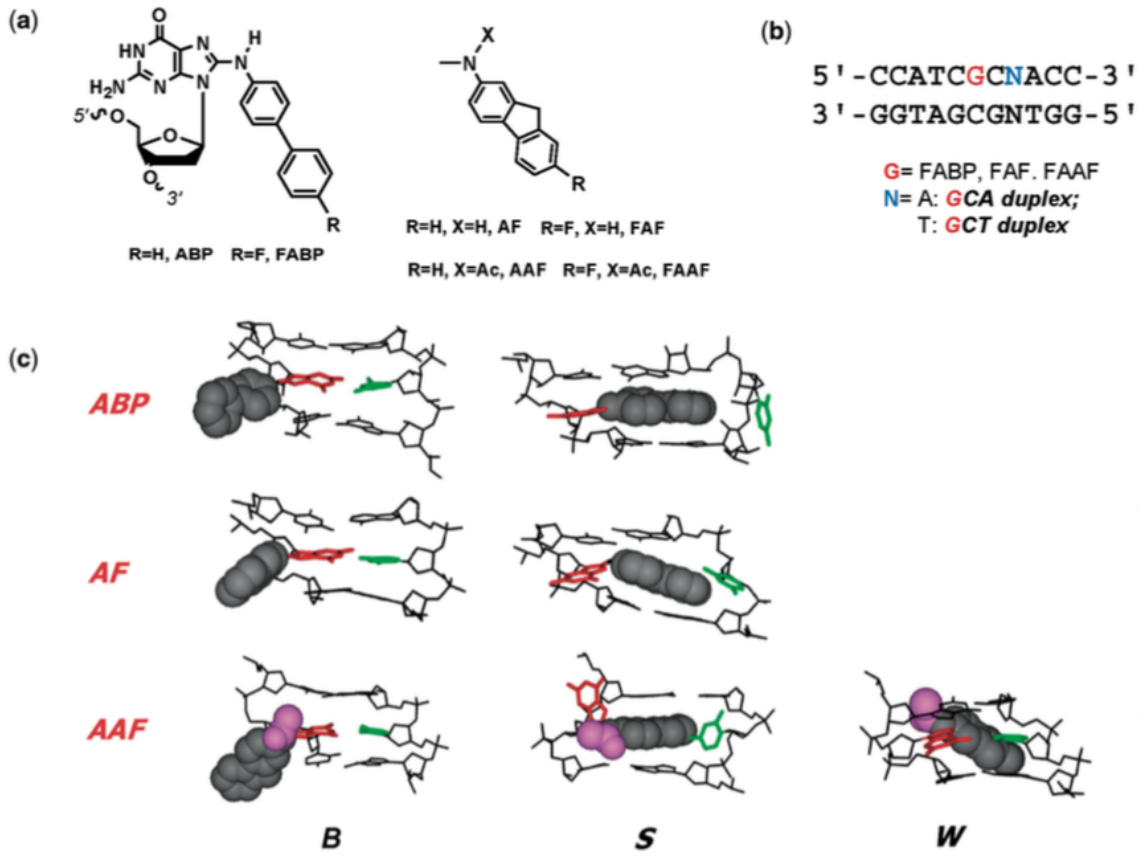


Figure 1. (**a**) Structures of ABP [*N-(2'-deoxyguanosin-8-yl)-4-aminobiphenyl*], AF [*N-(2'-deoxyguanosin-8-yl)-2-aminofluorene*] and AAF [*N-(2'-deoxyguanosin-8-yl)-2-acetylaminofluorene*] *and their fluoro models,* FABP [*N-(2'-deoxyguanosin-8-yl)-4-fluoro-4-aminobiphenyl*], *FAF* [*N-(2'-deoxyguanosin-8-yl)-7-fluoro-2-aminofluorene*] and FAAF [*N-(2'-deoxyguanosin-8-yl)-7-fluoro-2-acetylaminofluorene*]; (**b**) 11-mer GCA and GCT duplexes used in this study; (**c**) Major groove views of the B, S and W-conformers of ABP, AF and AAF. Modified-dG (red), dC (green) opposite the lesion site (orphan C), fluorene (grey CPK), acetyl (AAF only, magenta).

^{19}F NMR spectroscopy

Figure 2a–c compares the ^{19}F NMR spectra (−114 to −121 ppm) of modified DNA duplexes at 20°C for the G*_CT_ and G*_CA_ sequence contexts (see Supplementary Figure. S1 for full temperature ranges). ^{19}F signal assignments were made based on the H/D solvent effect, exchange spectroscopy, adduct-induced CD ($ICD_{290–350\ nm}$) and chemical shifts as previously described (6, 26, 32, 34).

FABP-duplexes

A clear conformational difference exists between the two isomeric FABP-modified G*_CA_ and G*_CT_ duplexes (Figure 2a). The single signal at −116.9 ppm for FABP-G*_CA_ has been previously assigned to the B-conformer (22). In contrast, FABP-G*_CT_ exhibited two signals at −116.9 (B) and −118.0 (S) ppm in a 40:60% ratio and adopted a two-site exchange (EXSY spectra at 5 and 17°C, inset, Supplementary Figure S1a). A large chemical shift gap (~1 ppm) between the two signals suggests differences of their electronic environments. In addition, the −116.9 ppm signal revealed a large H/D effect (+0.24 ppm) compared to the −118.0 ppm signal (+0.14 ppm) (data not shown) on increasing the D_2O content from 10 to 100%. The results indicate the exposed fluorine atom in the B-conformer, as observed in the MD/PMF simulations (Supplementary Figure S2).

FAF-duplexes

Although not as dramatic, a similar sequence effect was observed for FAF. The FAF-G*_CA_ duplex showed signals at −117.4 and −118.6 ppm in a 34:66% ratio (Figure 2b),

which have been assigned to B- and S-conformers, respectively (20, 22, 23, 26, 35). The S conformer population was increased to 90% in the FAF-G*CT duplex. Consistent with this assignment, the downfield −117.4 ppm signal revealed a larger H/D effect (+0.19 ppm) compared with the −118.8 ppm signal (+0.03 ppm) (data not shown), again consistent with the MD/PMF simulations (Supplementary Figure S2).

FAAF-duplexes

At least three major signals are present (Figure 2c) for FAAF in the G*CA and G*CT sequences. These signals are shifted considerably (∼2 ppm) to the downfield compared with FAF, a phenomenon associated with the *N*-acetyl factor. We reported previously S/B/W-conformer assignment of FAAF-modified 12- and 16-mer duplexes in several sequence contexts (i.e. TG*A, CG*C, CG*G and GG*C) (7, 8). The results revealed that ^{19}F signals of B-, S- and W-conformers appear consistently going from downfield to upfield in order of −115.0 to −115.5, −115.5 to −117.0 and −116.5 to −118.0 ppm, respectively. The major ^{19}F signals in Figure 2c fit that pattern. In particular, the signal patterns of the FAAF-modified G*CA/G*CT duplexes (5' -CCATCG*CNAC-3') match well with those observed for 16-mer (5'-CTCTCG$_1$G$_2$CG$_3$*CCATCAC-3') (7) and 12-mer (5'-CTTCTCG*CCCTC-3') duplexes (8), both of which have the CG*C sequence context (underlined). In line with this observation, their proton spectra displayed a mixture of broad imino signals arising not only from those involved in Watson–Crick hydrogen bonds (12–14 ppm) but also from the lesion site and its vicinity (11–12 ppm) (Supplementary Figure S3). Although the two sequences are similar in the total *syn* conformation (S+W) (i.e. 83 vs. 78%, respectively, for

G*CT and G*CA), the W-conformer population appears to be significantly greater for G*CT (30%) compared with G*CA (14%) (Supplementary Figure S4).

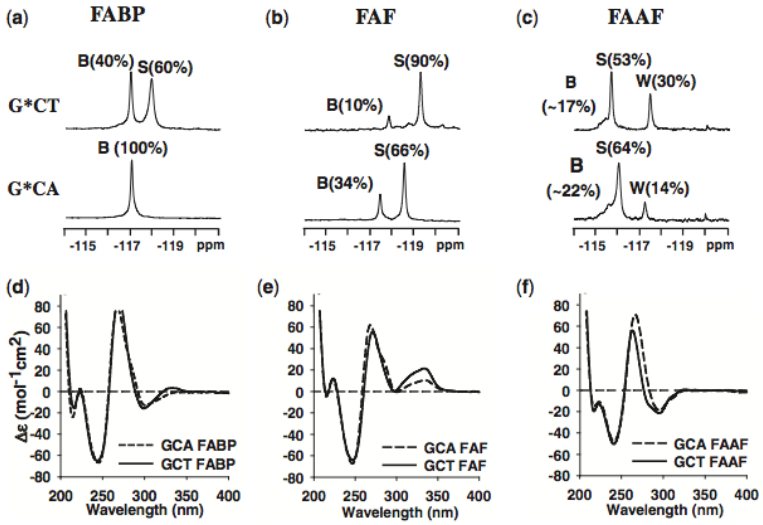


Figure 2. ^{19}F NMR (**a–c**) and CD (**d–f**) spectra of FABP-, FAF- and FAAF-modified 11-mer duplexes, respectively, in the G*CA and G*CT sequence contexts at 20°C. The B, S and W notation used in the ^{19}F NMR (a–c) signal assignments represent major groove 'B', base displaced stacked 'S' and minor groove 'W' conformers, respectively (see Figure 1c legend). The CD spectra (d–f) of modified duplexes show the typical B-DNA characteristic (positive and negative ellipticity at 275 and 245 nm, respectively) and different $ICD_{290-350\ nm}$ patterns represent the lesion conformation in the duplex (see CD in Results).

Induced circular dichroism

Figure 2d–f show the CD of the modified G*CA and G*CT duplexes. We reported that B- and S-conformers are characterized by positive and negative $ICD_{290-350\ nm}$, respectively (26). Accordingly, the B-conformeric FABP-G*CA displayed a strongly negative $ICD_{290-350\ nm}$, whereas an S-shape curve was observed for the S/B-mixture G*CT duplex (Figure 2d). These results are in good agreement with the ^{19}F NMR results (Figure 2a). Unlike FABP, FAF on both sequences exhibited a strong positive $ICD_{290-350\ nm}$ with the effect much greater for G*CT (Figure 2e), consistent with the greater S-conformer population determined by ^{19}F NMR (Figure 2b). The ICD of FAAF (Figure 2f), which is confined in the narrow 290–320 nm range, has not been defined as clearly as FAF (8).

In addition, the modified duplexes displayed a significant blue shift relative to their respective control duplexes (Supplementary Figure S5a and b and Table 1). All except for FAF-G*_CT_ exhibited significant blue shifts up to 8 nm. The bulky *N*-acetylated FAAF exhibited greater shifts (Δ_{G^*-G} = 4–8 nm) than FAF and FABP (Δ_{G^*-G} = 0–4n m). GCA sequences, which are prone to the B-conformer, displayed greater blue shift ($\Delta_{G^*CA-\ G^*CT}$ = 2–4 nm, Table 1) compared with GCT. It is well known that protein-induced DNA bending exhibits significant CD shift at 275 nm of regular B-type DNA (36-38). For instance, the HMG box protein SOX-5 bends DNA by ~74° on binding, which resulted in a significant blue CD shift (37). These reports suggest that the blue shifts observed in this study result from the distortion of the DNA backbone, particularly bending.

Gel mobility assay

Two 19-mer sequences (5' -CTTACCATCGC*N*ACCATT C-3' , *N* = T or A) were used to investigate the impact of the A/T polarity swap at the *N* position on the gel mobility of the modified duplexes. Initially, the abovementioned 11-mer sequences were used but they denatured in the 15% native polyacrylamide gel at 1800 V (data not shown). Figure 3 compares the electrophoretic mobility of the 19-mer GCA and GCT sequences with and without modifications. Differential mobility between the single strand and double strand (ds) oligonucleotides confirmed the integrity of the duplexes (Figure 3). The modified duplexes exhibited retardation in the mobility in a lesion-dependent manner. In both sequences, major retardation effect was observed for FAAF followed by FABP, whereas no retardation was observed for FAF, results consistent with the CD blue shift data above (Table 1). It should be noted that the magnitude of retardation in mobility observed in this

Table 1. Lesion-induced CD blue shifts

Lesion	GCA (nm)	GCT (nm)	Blue Shift $\Delta^{a}_{(G^*-G)}$ (nm)		Blue shift $\Delta^{b}_{(G^*CA-G^*CT)}$ (nm)
Control	271	271	GCA	GCT	
FABP	267	269	4	2	2
FAF	268	271	3	0	3
FAAF	263	267	8	4	4

[a]Difference in the wavelength of positive band between the modified and control duplexes.
[b]Difference in the wavelength of positive band between the G*CA and G*CT duplexes.

study is significantly lower than what was previously observed in benzo[*a*]pyrene modified duplexes (39-41). We cannot rule out a possibility of C8-subsitituted dG (this study) versus *N*2- substituted dG (benzo[*a*]pyrene) binding. However, the rationale behind this small difference in mobility could be well due to the significantly longer oligonucleotides (19-mer) used in this study as opposed to the aforementioned benzo[*a*]pyrene cases (11 and 15-mer). The utilization of longer sequences will reduce the number of adducts per helical turn, which might result into diminishing of the lesion-induced bending effect. Tsao *et al.* reported similar effects of oligonucleotide length on the electrophoretic mobility of benzo[*a*]pyrene-modified duplexes (41).

Thermodynamics

Thermodynamics results from UV-optical melting (Supplementary Table S1) and DSC, which is not dependent on melting patterns and stoichiometry (22) (Table 2), are comparable. Supplementary Figure S6 shows the DSC thermograms of modified duplexes in the G**CA* and G**CT* sequences relative to the unmodified controls. These curves were transformed into the corresponding thermodynamic charts (Figure 4a and b), and the

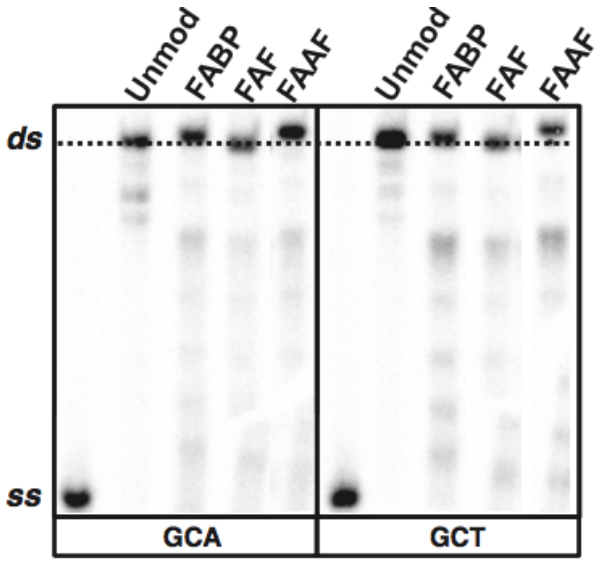


Figure 3. Autoradiograph of 15% (w/v) native polyacrylamide gel (acrylamide/bis-acrylamide 29:1, w/w) showing the relative mobility of ss and GCA/GCT 19-mer ds; both unmodified and single site specifically modified by FABP, FAF and FAAF.

results are tabulated in Table 2. According to the NMR results (Figure 2), FABP and FAF display a S/B-equilibrium, whereas FAAF produces a complex S/B/ W-equilibrium; thus, they will be compared separately.

Table 2. Thermal and thermodynamic parameters derived from DSC curves

	CCATC*G*CA*ACC GGTAGCGTTGG				CCATC*G*CT*ACC GGTAGCGATGG			
	$-\Delta H$ (kcal/mol)	$-\Delta S$ (eu)	$-\Delta G_{37}$ (kcal/mol)	T_m (°C)	$-\Delta H$ (kcal/mol)	$-\Delta S$ (eu)	$-\Delta G_{37}$ (kcal/mol)	T_m (°C)
Control[a]	79.1	214.9	12.4	64.5	75.0	203.9	11.8	63.0
FABP[a]	76.4	214.1	10.0	54.3	64.2	178.3	8.9	51.8
FAF[a]	65.7	180.7	9.7	55.4	58.6	160.1	8.9	53.4
FAAF[a]	39.7	106.1	6.8	42.7	33.9	88.8	6.4	39.7
	$\Delta\Delta H$ (kcal/mol)[b]	$\Delta\Delta S$ (eu)[c]	$\Delta\Delta G_{37}$ (kcal/mol)[d]	ΔT_m (°C)[e]	$\Delta\Delta H$ (kcal/mol)[b]	$\Delta\Delta S$ (eu)[c]	$\Delta\Delta G_{37}$ (kcal/mol)[d]	ΔT_m (°C)[e]
FABP[a]	2.7	0.8	2.4	−10.2	10.8	25.6	2.9	−11.2
FAF[a]	13.4	34.2	2.7	−9.1	16.4	43.8	2.9	−9.6
FAAF[a]	39.4	108.8	5.6	−21.8	41.1	115.1	5.4	−23.3

The average standard deviations for ΔG, ΔH, and T_m are ±0.2, ±2.0, and ±0.4, respectively.
[a]The results were calculated from integration of the DSC curve directly. ΔG and ΔH represent the heat absorbed during duplex melting at 0.12 mM.
[b]$\Delta\Delta H$ = ΔH (modified duplex) − ΔH (control duplex).
[c]$\Delta\Delta S$ = ΔS (modified duplex) − ΔS (control duplex).
[d]$\Delta\Delta G$ = ΔG (modified duplex) − ΔG (control duplex).
[e]ΔT_m = T_m (modified duplex) − T_m (control duplex).

FABP/FAF-DNA duplexes

Both FABP and FAF resulted in destabilization (Figure 4 and Table 2). FABP reduced T_m for the G*CA and G*CT duplexes by −10.2 and −11.2°C and $\Delta\Delta G_{37°C}$ by 2.4 and 2.9 kcal/mol, respectively. The G*CA/G*CT transition produced major effect on ΔΔH (2.7 vs. 10.8 kcal/mol) and ΔΔS (0.8 vs. 25.6 eu), consistent with a significant increase in the S-conformer population from 0 to ~60% (Figure 2a). The B-conformer FABP-G*CA is expected to lead to small entropy compensation, and consequently, the enthalpy reduction dominated the free-energy destabilization (Figure 4a). As expected, the structural disturbance caused by the S/B mixture FABP-G*CT duplex leads to a considerable reduction of melting enthalpy; however, most of it is compensated by entropy (Figure 4b).

FAF modification resulted in a similar destabilization effect: ΔT_m by −9.1 and −9.6°C and $\Delta\Delta G_{37°C}$ by 2.7 and 2.9 kcal/mol, respectively, for G*CA and G*CT. However, compared with FABP, FAF in both sequences yielded significantly larger reduction in enthalpy (ΔΔH = 13.4 and 16.4 kcal/mol) and entropy (ΔΔS = 34.2 and 43.8 eu) compensation (Figure 4 and Table 2). FAF exhibits more S-conformer than FABP in both sequences because of stronger stacking effect. As a result, sequence dependence on the thermodynamics was not as dramatic as in FABP. It is clear from Figure 4 that FAF (over FABP) and G*CT (over G*CA) combinations produce consistently greater enthalpy/entropy compensation, that is, FAF/G*CT > FAF/G*CA > FABP/G*CT > FABP/G*CA from the highest to the lowest. As expected the *N*-deacetylated FABP-G*CT and FAF-G*CA exhibited two site exchange (Supplementary Table S2).

FAAF-duplexes

FAAF modification resulted in the most significant reduction of ΔT_m by −21.8 and −23.3°C and ΔΔH by 39.4 and 41.1 kcal/mol, respectively, for G*CA and G*CT sequence contexts (Figure 4a and b and Table 2). This is due to the bulky acetyl group in FAAF (Supplementary Figure S2). Like FABP and FAF, however, entropy compensation contributed a stabilization, that is, ΔΔG = 5.6 and 5.4 kcal/mol, respectively, for the G*CA and G*CT.

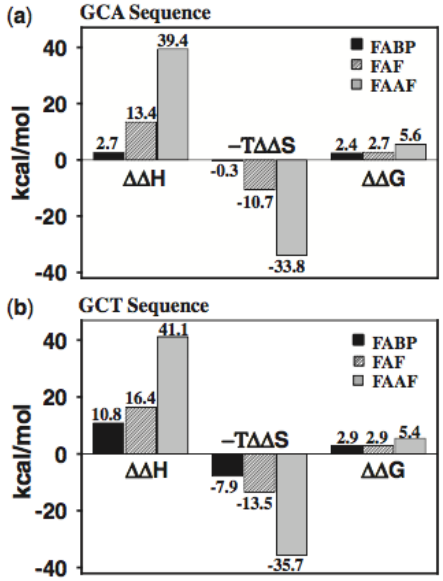


Figure 4. Comparative thermodynamic parameters histogram of FABP (black), FAF (hatched) and FAAF (gray) in (**a**) G*CA duplexes and (**b**) G*CT duplexes. The ΔΔ values represent: $\Delta\Delta H = \Delta H$ (modified duplex) $- \Delta H$ (control duplex), $\Delta\Delta S = \Delta S$ (modified duplex) $- \Delta S$ (control duplex) and $\Delta\Delta G = \Delta G$ (modified duplex) $- \Delta G$ (control duplex).

Molecular dynamics/potential mean force calculations

To further understand the impact of lesion modification on the G*CA and G*CT duplexes, MD simulations were performed in combination with potential mean force (PMF) calculations. PMF calculations yield the free energy as a function of the extent of flipping of

the modified G* base (Supplementary Figure S7). Individual PMFs were determined with G* in the either *anti* or *syn* orientations, whereas only the *anti* orientation was studied for the unmodified duplexes.

Figure 5 shows the free energies from the PMFs. In the unmodified and *anti*-G* PMFs, there is a deep minimum in 330–360° corresponding to the Watson-Crick (WC)-base paired state (31), which in case of lesions corresponds to the B-state. Deep minima are also present in 330–360° in the *syn*-G* PMFs, which corresponds to the S-state. The conformer assignments were made based on the published experimental NOE data (Supplementary Figures S8–S11 and Supplementary Tables S3–S6) (42-45). The relative energies of the flipped states are highest in the unmodified duplexes in all cases indicating that the lesions lower the relative energies of the flipped state and/or destabilize the low- energy B- or S-states. Notably, the free-energy surfaces show the relative energies of the flipped states to be lower in the *syn* PMFs, consistent with the conformational thermodynamic data discussed above. As such, the lowest energies of the flipped state occur with FAAF (Table 2) indicating, that the *syn* FAAF may sample a wider range of conformations as compared with FABP and FAF, consistent with the ^{19}F NMR data (Figure 2). Further validation of the PMFs is the energies of the minima being about 15 kcal/mol, which is in good agreement with the experimental $\Delta G^{\neq}$ of 14.1 kcal/mol required for a B/S conversion. Representative B/S/W structures from the NOE-based PMFs are shown in Supplementary Figures S12–S14. For all three lesions, the presence of WC base pairing in the B-state and the stacking of the adduct into the duplex in the S-state are evident. In contrast, the structure of the W-state, which is

only populated by FAAF, is highly distorted.

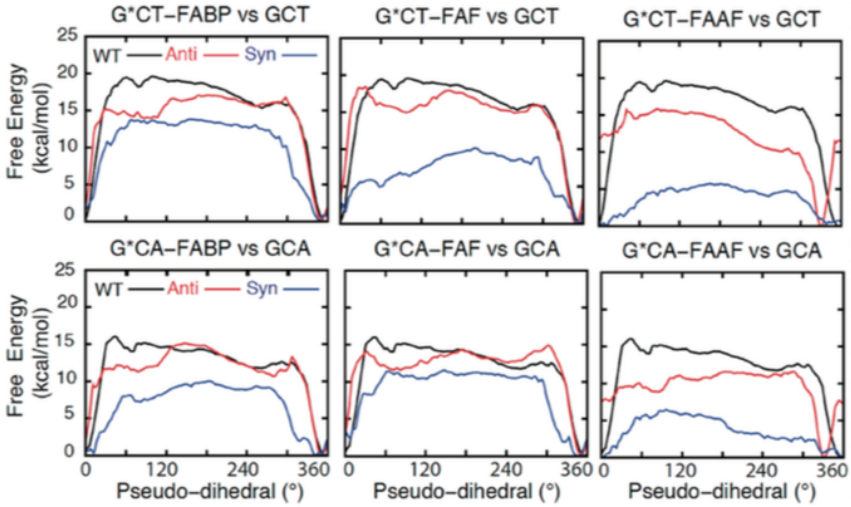


Figure 5. Free-energy profiles as a function of the pseudo-dihedral angle phi (Supplementary Figure S7) from PMF calculations over the sampling range of 0.5–3 ns for G*CT-FABP, G*CA-FABP, G*CT-FAF, G*CA-FAF, G*CT-FAAF and G*CA-FAAF modified sequences (red and blue) and the unmodified GCT and GCA (black).

In the *anti*-G* PMFs, the average solvent accessible surface areas (SASAs) of both the lesion and fluorine atom are high in the vicinity of the B state, consistent with the location adduct in the major groove (Supplementary Figure S2). For the *syn* PMFs, the SASA values are low in the regions corresponding to the S-state, consistent with adducts being stacked inside the helices. However, the SASA values are higher with FAAF compared with FABP and FAF, suggesting an altered environment for FAAF. In addition, the *syn* PMFs of FAAF exhibits SASA minima in the 60–120° region, which encompasses the W-state. These suggest a fundamental difference in conformational properties of FAF/FABP versus FAAF, consistent with the significant difference in the ^{19}F chemical shifts for the FAAF species and thermodynamic data.

Shown in Supplementary Figure S15 are bending probability distributions for the B-, S- and W-states for the three lesions in both the G*_CA_* and the G*_CT_* contexts. In the B- and S-states, the extent of bending is significantly larger with FAAF (cyan) versus FABP (red)

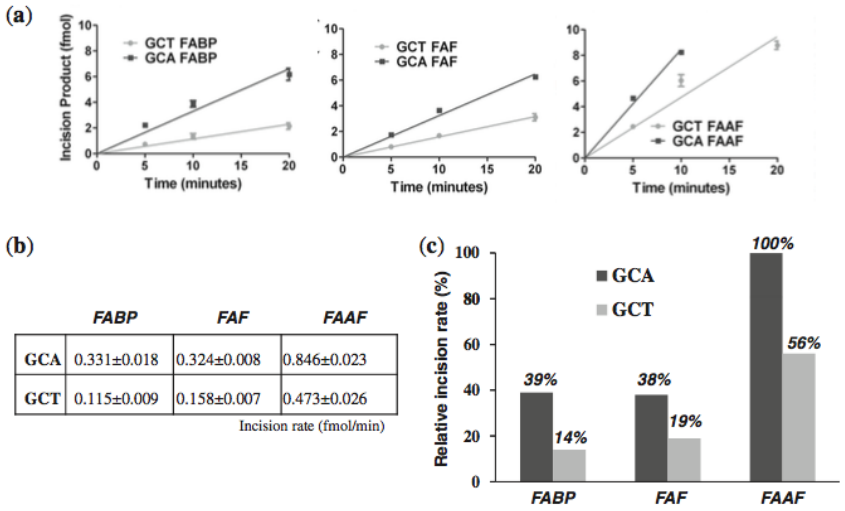
and FAF (blue). These results are consistent with the experimental data obtained from the greater blue shift in CD (Supplementary Figure S5 and Table 1), although the changes in G*CA occur only in the S-state (Supplementary Figure S15). In addition, the simulations indicate the extent of bending for FABP to be similar to that of FAF. The significant increase in bending in FAAF is consistent with the greater destabilization of the duplexes caused by the FAAF lesion (Table 2). Concerning the bending, calculation of local helicoidal parameters revealed significant differences in twist and tilt for base pair 8, where the A/T switch occurs (Supplementary Table S7). For example, twist values are systematically larger and tilt values are less negative in GCT versus GCA sequences. These differences suggest that the local structural alteration associated with the A/T switch is being propagated to the overall helix.

Escherichia coli UvrABC incision

DNA substrates containing lesions in the defined sequences were subjected to incisions by the *E. coli* UvrABC system in a kinetic assay. These substrates were radioactively labeled at the 5'-end of the adducted strand and the major incision products separated on a urea-PAGE gel running under denaturing conditions (Supplementary Figure S16). The incision occurred at the eighth phosphate bond 5' to G*, which is consistent with the currently accepted mechanisms of UvrABC-based NER (28, 29). The substrates were incised at differing efficiencies depending on not only the type of DNA adduct but also the sequence context (Figure 6). Specifically, the G*CA sequences were incised at higher rates by ~2-fold than G*CT, while the order of incision rate of adducts is FAAF > FAF ≈ FABP for

both sequences, with FAAF being incised with 2- to 3-fold greater efficiency than the other lesions.

It should be noted that the 5'-incision products appeared as doublet bands (Supplementary Figure S16). Similar incision products of this type of lesion have been observed previously (35, 46, 47). This is likely either due to the type of arylamine lesions or due to the structural heterogeneity exhibited by this type of lesions as demonstrated in this study and previous studies, suggesting that UvrABC may make the 5'-incision at the site different by one nucleotide for the different conformers of arylamine lesion.



	FABP	FAF	FAAF
GCA	0.331±0.018	0.324±0.008	0.846±0.023
GCT	0.115±0.009	0.158±0.007	0.473±0.026

Incision rate (fmol/min)

Figure 6. Absolute incision rates (**a** and **b**) of FABP, FAF and FAAF-modified 55-mer substrates in the G*CA and G*CT sequence contexts. (**c**) Relative percent incision rate histograms with respect to G*CA-FAAF modified duplex.

Discussion

Conformational and thermodynamic effects on the G**CA*/G**CT* transition

The NMR/ICD results indicate that lesion stacking is affected considerably by a polarity swap at the 3'-next flanking base (*GCA* → *GCT*). The effect was most significant for FABP, which produced a dramatic increase in S-conformer (0–60%) (Figure 2a). *This is an extraordinary DNA sequence effect*. A similar trend was observed for FAF, although the S-conformer was only 24% greater in G**CT* than in G**CA* (Figure 2b). Unlike FABP and FAF, the impact of the A/T swap on FAAF was minimal; specifically, the *syn*-glycosidic S- and W-conformers remained relatively unchanged (~78 to ~83%) (Figure 2c). Interestingly, the increase of W-conformer (14–30%) appeared to be compensated by a concomitant decrease of S-conformer (64–53%). These data indicate that the *N*-acetyl group in FAAF can push the low-energy *syn*-S/W-equilibrium toward W (see '*N*-acetyl factor'). Overall, these results indicate that the A/T swap has the largest impact on the most stable system, whereas the least stable FAAF lesion is the least impacted.

As expected, all modified duplexes were consistently destabilized compared with the controls (Figure 4 and Table 2): FAAF > FAF ≈ FABP. The G**CA*/G**CT* transition led to further destabilization, which was associated with increases in lesion stacking (greater S/W) for all three lesions. Obviously, a higher population of the *syn*-S-/W-conformer states is expected to disrupt the double helical structure, which would significantly reduce the enthalpy, accompanied by a compensatory increase in entropy (22).

Lesion-induced DNA bending as a major NER recognition factor

For each lesion, a greater proportion of B-conformer was observed in G**CA** (FABP: 100%, FAF: 34% and FAAF: 22%) than G**CT** (FABP: 40%, FAF: 10% and FAAF: 17%). Moreover, good correlation between the magnitudes of change in conformer populations and incision efficiencies was found among the lesions. In FABP, the A/T polarity swap caused a 60% change in S-conformer proportion and a 3-fold reduction in repair efficiency. The changes were significantly lower in FAF and FAAF (24 and 5–15%, respectively), as were the repair efficiencies (2.0- and 1.8-fold, respectively). At a glance, the results seem to suggest B-conformers have greater reparability than S-conformers. This feature is in clear contrast to the trend that has been observed previously for AF and AAF in certain sequence contexts, that is, the S-conformer is more reparable than the B-conformer (7, 35). This type of conformation-specific repair is not only restricted to arylamines but also applied to other bulky lesions. For instance, Geacintov et al. have reported that the base-displaced *cis*-N^2-dG adducts of benzo[*a*]pyrene are incised more efficiently than the minor groove–orientated *trans*-N^2-dG adducts (33).

However, the *E. coli* repair results in this study seem to match well with events of adduct-induced DNA bending/distortion, as evidenced by blue shifts in CD (Table 1) and retardation of mobility in electrophoretic mobility shift assay (Figure 3). The slowed mobility indicates flexibility at the lesion site as observed by Tsao et al. for (+)-*trans*-*anti*-[BP]-N^2-dG lesion in the TG*T sequence context with concomitant thermal destabilization (41, 48). Similarly, the bulky *N*-acetyl FAAF exhibited significantly slower electrophoretic mobility compared with FAF and FABP within the same sequence context. In case of sequence, the G**CA** duplex exhibited consistently greater bending than its G**CT*

counterpart, with the effect being significantly greater for FAAF than FAF and FABP. A similar CD pattern has been reported for AAF-modified *Nar*I duplexes related to the formation of a B-Z junction (49). Clearly, the A/T swap alters the conformational equilibrium *anti* (B-) to *syn* (S- or W-). It should be noted that the G**CA* (–TCG*CAA–) sequence contains a stretch of alternating pyrimidine:purine bases, which are predisposed to DNA bending (50-52). In contrast, such a stretch is interrupted in the highly S-conformeric G**CT* (–TCG*CTA–) sequence. It is possible that the B-conformer may facilitate DNA bending, due to the exposure of the carcinogen moiety to the major groove's hydrophilic environment. In both sequences, a major effect was observed with FAAF, followed by FABP and FAF (Figure 3). However, MD/PMF simulations indicate that the major changes in G**CA* occur in the S-state. Also, unlike the CD data, there were no significant differences in electrophoretic mobility between the two sequences (Figure 3). The reason for the inconsistency in the mobility, CD and MD data demonstrating the sequence effect is not apparent, but the greater bending and flexibility of FAAF over FABP or FAF is in good agreement with the observed repair efficiencies (FAAF >> FAF ≈ FABP; Figure 6).

The repair results in Figure 6 along with previously reported work on polycyclic aromatic hydrocarbons (41, 53) and arylamines (7) indicate that lesion-induced destabilization of DNA is a major determining factor for repair. However, these lesions were consistently repaired two to three times more efficiently in G**CA* than in G**CT*, which was not consistent with relative thermodynamic stabilities observed for each. The inconsistence is likely due to the second step of damage recognition (54) that becomes much more significant for FAAF versus FABP and FAF within a given sequence. Unlike the

initial step of damage recognition by $UvrA_2$, which depends on DNA conformation and sequence, the second step of recognition is well known to be characterized by the direct interaction of UvrB with adduct itself on DNA strand opening (47, 54-56). In other words, the structure and chemistry of the lesions matter more with UvrB than $UvrA_2$. Recently, Liu et al. reported the NER incision efficiencies of the bulky benzo[*a*]pyrene and equine estrogen substrates using human HeLa cell extracts and bacterial UvrABC proteins (53). They demonstrated that despite having differences in the prokaryotic and eukaryotic NER proteins, XPC-RAD23B and UvrB, respectively, they exhibit common feature of β-hairpin intrusion for damage recognition. In addition, it was found that local thermodynamic destabilization near the lesion site assists the insertion of β-hairpin, thus recognition.

Clearly, this study shows that the thermodynamic destabilization of the DNA duplex along with lesion flexibility promotes strand opening and thus the second step of damage recognition. The presence of the *N*-acetyl group (see below) may make FAAF more efficiently recognized than FAF and FABP at the second step due to its flexible nature and greater destabilization of the DNA double helix. As for the G**CA*/G**CT* transition, the initial recognition step conducted by $UvrA_2$ should be a major determinant factor as the same efficiency of recognition at the second step is expected for the same type of lesion. Thus, bending appears to be an important factor for the DNA damage recognition. Indeed, a recent crystal study by Jaciuk et al. (10) found that in the active site of UvrA, the fluorescein-modified duplexes were bent by ~15° and the structure was related to the kinked structure of psoralen and PAH adducts according to NMR (57). They concluded that the $UvrA_2$ protein does not have direct chemical contacts with a lesion *per se*, but indirectly senses the overall helical distortion (unwinding and bending) (10). Because energy is

required for the bending, formation of the pre-bent DNA induced by bulky lesions would likely enhance the $UvrA_2$ binding and thus damage recognition.

Cai *et al.* have reported a similar repair trend for the 5'-CACACCG*CACAC-3' sequence versus 5'-CCATCCG*CTACC-3', in which G* is the major mutagenic lesion derived from the environmental carcinogen benzo[*a*]pyrene, 10S (+)-*trans*-*anti*-B[*a*]P-N^2-dG (39). A greater repair (1.6-fold) of the *CG*CA* duplex over the *CG*CT* counterpart was attributed to its higher bending of the distant 5'-end sequences, as evidenced by gel experiments and MD simulations; these findings are consistent with the bending argument made in this study. Although the sequence contexts (underlined above) near the lesion site, including the 3'-next flanking base, are identical to those used in this study, they did not consider the structural and repair consequences of the *GC*A/G*CT* swap.

N-acetyl factor

Although a relatively small modification, the *N*-acetyl group has an important structural consequence. As shown in Figure 7, the lack of the acetyl moiety in G**CT*-FAF allows the G* moiety (red licorice representation) to point away from the sugar and stay in the plane of the GC base pair, where the *N*-H bond is directed toward the sugar. However, in G**CT*-FAAF, the acetyl group will have a steric clash with the sugar moiety of G* (identified with a black arrow), thereby leading the fluorene moiety (cyan) to be perpendicular to the G* ring system. This persistent 'perpendicular' lesion orientation is predicted to lead to more disruption of the DNA duplex. A similar observation regarding the differences in the orientation of AF and AAF was reported by Mu *et al.* (19) who have conducted a NER study of these lesions in human HeLa cells. MD simulations in that work indicate that the greater

repair susceptibility of AAF stems from steric hindrance effects of the acetyl group, which significantly diminish the adduct base stabilizing van der Waals stacking interactions relative to AF. The persistent 'perpendicular' FAAF mentioned earlier could raise barriers between conformations of FAAF modified DNA, resulting in the overall lower free energy of the *syn*-G* PMFs for FAAF, compared with FABP and FAF. In other words, the *N*-acetyl group in FAAF could act as a 'conformational locker' (7) that orients the adduct in a position that will lead to greater destabilization of the DNA duplex (Figure 4 and Table 2), as well as the increased bending observed in CD (Table 1) and mobility assays (Figure 3). As a result, FAAF lesions are repaired at significantly greater rate compared with the FABP and FAF lesions (7).

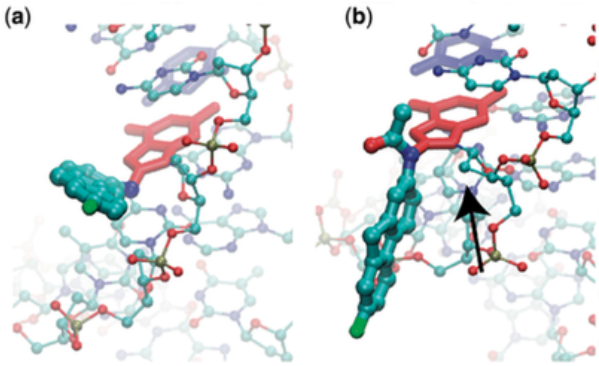


Figure 7. Image of the (**a**) G*CT-FAF adduct in the B state (350°) and (**b**) G*CT-FAAF adduct in the B state (330°). The flipping G6 base is red licorice, the orphan C6 is blue licorice, adducts are atom-colored thick CPK and the reminder of the DNA is atom-colored thin CPK. (b) The arrow indicates the sugar moiety with which the acetyl or aryl moiety on FAAF is suggested to form a steric clash.

Conclusion

The A to T polarity swap in the arylamine-modified G*CA/G*CT transition produced a dramatic increase in destabilized stacked conformation but resulted in unexpected 2- to 3-fold lower NER efficiencies. These results are consistent with lesion-induced DNA

bending/distortion. As for lesions, FAAF was repaired three to four times more efficiently than FABP and FAF lesions, which is consistent with the extent of bending and helix destabilization, as well as the steric constraint in the duplex ('*N*-acetyl factor') (7). A number of different damage recognition parameters have been implicated in the molecular mechanisms of NER (9, 55, 58). However, it is known that thermal/thermodynamic destabilization and DNA distortion/bending are important factors for damage recognition by repair proteins (9, 39, 53). The results of this study show that lesion-induced DNA bending/thermodynamic destabilization is a more important NER factor than the usual S/B conformational heterogeneity, as has been observed previously for AF and AAF in certain sequence contexts (7, 35). This work represents a novel 3'-next flanking sequence effect as a unique NER factor for bulky arylamine lesions in *E. coli*. Taken together, the results of this study demonstrate the complexity in DNA recognition factors for repair of bulky arylamine lesions in *E. coli*.

References

1. Greenblatt,M.S., Bennett,W.P., Hollstein,M. and Harris,C.C. (1994) Mutations in the p53 tumor suppressor gene: clues to cancer etiology and molecular pathogenesis. *Cancer Res*, **54**, 4855–4878.

2. Friedberg,E.C., Walker,G.C., Siede,W. and Wood,R.D. (2005) DNA Repair and Mutagenesis American Society for Microbiology Press.

3. Luch,A. (2005) Nature and nurture - lessons from chemical carcinogenesis. *Nat Rev Cancer*, **5**, 113–125.

4. Beland,F.A. and Kadlubar,F.F. (1990) Metabolic Activation and DNA Adducts of Aromatic Amines and Nitroaromatic Hydrocarbons. In Cooper,C.S., Grover,P.L. (eds), *Chemical Carcinogenesis and Mutagenesis I*, Handbook of Experimental Pharmacology. Springer Berlin Heidelberg, Berlin, Heidelberg, Vol. 94 / 1, pp. 267–325.

5. Cho,B. (2010) Structure–Function Characteristics of Aromatic Amine-DNA Adducts. *The Chemical biology of DNA damage*, 10.1002/9783527630110.ch10.

6. Meneni,S.R., Shell,S.M., Gao,L., Jurecka,P., Lee,W., Sponer,J., Zou,Y., Chiarelli,M.P. and Cho,B.P. (2007) Spectroscopic and theoretical insights into sequence effects of aminofluorene-induced conformational heterogeneity and nucleotide excision repair. *Biochemistry*, **46**, 11263–11278.

7. Jain,V., Hilton,B., Patnaik,S., Zou,Y., Chiarelli,M.P. and Cho,B.P. (2012) Conformational and thermodynamic properties modulate the nucleotide excision repair of 2-aminofluorene and 2-acetylaminofluorene dG adducts in the NarI sequence. *Nucleic Acids Res*, **40**, 3939–3951.

8. Patnaik,S. and Cho,B.P. (2010) Structures of 2-acetylaminofluorene modified DNA revisited: insight into conformational heterogeneity. *Chem. Res. Toxicol.*, **23**, 1650–1652.

9. Geacintov,N.E., Broyde,S., Buterin,T., Naegeli,H., Wu,M., Yan,S. and Patel,D.J. (2002) Thermodynamic and structural factors in the removal of bulky DNA adducts by the nucleotide excision repair machinery. *Biopolymers*, **65**, 202–210.

10. Jaciuk,M., Nowak,E., Skowronek,K., Tańska,A. and Nowotny,M. (2011) Structure of UvrA nucleotide excision repair protein in complex with modified DNA. *Nature Publishing Group*, **18**, 191–197.

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

12. Cai,Y., Patel,D.J., Broyde,S. and Geacintov,N.E. (2010) Base sequence context effects on nucleotide excision repair. *J Nucleic Acids*, **2010**, 1–9.

13. Broschard,T.H., Koffel-Schwartz,N. and Fuchs,R.P.P. (1999) Sequence-dependent modulation of frameshift mutagenesis at NarI-derived mutation hot spots. *J Mol Biol*, **288**, 191–199.

14. Burnouf,D., Koehl,P. and Fuchs,R.P. (1989) Single adduct mutagenesis: strong effect of the position of a single acetylaminofluorene adduct within a mutation hot spot. *Proc Natl Acad Sci USA*, **86**, 4147–4151.

15. Delagoutte,E., Bertrand-Burggraf,E., Lambert,I.B. and Fuchs,R.P.P. (1996) Binding and Incision Activities of UvrABC Excinuclease on Slipped DNA Intermediates that Generate Frameshift Mutations. *J Mol Biol*, **257**, 970–976.

16. Fuchs,R.P. and Seeberg,E. (1984) pBR322 plasmid DNA modified with 2-acetylaminofluorene derivatives: transforming activity and in vitro strand cleavage by the Escherichia coli uvrABC endonuclease. *The EMBO Journal*, **3**, 757–760.

17. Mu,D., Bertrand-Burggraf,E., Huang,J.C., Fuchs,R.P., Sancar,A. and Fuchs,B.P. (1994) Human and E.coli excinucleases are affected differently by the sequence context of acetylaminofluorene-guanine adduct. *Nucleic Acids Res*, **22**, 4869–4871.

18. Seeberg,E. and Fuchs,R.P. (1990) Acetylaminofluorene bound to different guanines of the sequence -GGCGCC- is excised with different efficiencies by the UvrABC excision nuclease in a pattern not correlated to the potency of mutation induction. *Proc Natl Acad Sci USA*, **87**, 191–194.

19. Mu,H., Kropachev,K., Wang,L., Zhang,L., Kolbanovskiy,A., Kolbanovskiy,M., Geacintov,N.E. and Broyde,S. (2012) Nucleotide excision repair of 2-acetylaminofluorene- and 2-aminofluorene-(C8)-guanine adducts: molecular dynamics simulations elucidate how lesion structure and base sequence context impact repair efficiencies. *Nucleic Acids Res*, **40**, 9675–9690.

20. Jain,N., Li,Y., Zhang,L., Meneni,S.R. and Cho,B.P. (2007) Probing the sequence effects on NarI-induced -2 frameshift mutagenesis by dynamic 19F NMR, UV, and CD spectroscopy. *Biochemistry*, **46**, 13310–13321.

21. Jain,N., Meneni,S., Jain,V. and Cho,B.P. (2009) Influence of flanking sequence context on the conformational flexibility of aminofluorene-modified dG adduct in dA mismatch DNA duplexes. *Nucleic Acids Res*, **37**, 1628–1637.

22. Liang,F. and Cho,B.P. (2010) Enthalpy-entropy contribution to carcinogen-induced DNA conformational heterogeneity. *Biochemistry*, **49**, 259–266.

23. Meneni,S.R. (2006) Sequence effects of aminofluorene-modified DNA duplexes: thermodynamic and circular dichroism properties. *Nucleic Acids Res*, **34**, 755–763.

24. Gao,L., Zhang,L., Cho,B.P. and Chiarelli,M.P. (2008) Sequence verification of oligonucleotides containing multiple arylamine modifications by enzymatic digestion and liquid chromatography mass spectrometry (LC/MS). *Journal of the American Society for Mass Spectrometry*, **19**, 1147–1155.

25. Chakrabarti,M.C. and Schwarz,F.P. (1999) Thermal stability of PNA/DNA and DNA/DNA duplexes by differential scanning calorimetry. *Nucleic Acids Res*, **27**, 4801–4806.

26. Liang,F., Meneni,S. and Cho,B.P. (2006) Induced Circular Dichroism Characteristics as Conformational Probes for Carcinogenic Aminofluorene−DNA Adducts. *Chem. Res. Toxicol.*, **19**, 1040–1043.

27. Cho,B.P. and Zhou,L. (1999) Probing the conformational heterogeneity of the acetylaminofluorene-modified 2'-deoxyguanosine and DNA by 19F NMR spectroscopy. *Biochemistry*, **38**, 7572–7583.

28. Luo,C., Krishnasamy,R., Basu,A.K. and Zou,Y. (2000) Recognition and incision of site-specifically modified C8 guanine adducts formed by 2-aminofluorene, N-acetyl-2-

aminofluorene and 1-nitropyrene by UvrABC nuclease. *Nucleic Acids Res*, **28**, 3719–3724.

29. Zou,Y., Liu,T.M., Geacintov,N.E. and Van Houten,B. (1995) Interaction of the UvrABC nuclease system with a DNA duplex containing a single stereoisomer of dG-(+)- or dG-(-)-anti-BPDE. *Biochemistry*, **34**, 13582–13593.

30. Zou,Y. and Van Houten,B. (1999) Strand opening by the UvrA(2)B complex allows dynamic recognition of DNA damage. *The EMBO Journal*, **18**, 4889–4901.

31. Banavali,N.K. and Mackerell,A.D. (2002) Free energy and structural pathways of base flipping in a DNA GCGC containing sequence. *J Mol Biol*, **319**, 141–160.

32. Meneni,S., Liang,F. and Cho,B.P. (2007) Examination of the long-range effects of aminofluorene-induced conformational heterogeneity and its relevance to the mechanism of translesional DNA synthesis. *J Mol Biol*, **366**, 1387–1400.

33. Geacintov,N.E., Cosman,M., Hingerty,B.E., Amin,S., Broyde,S. and Patel,D.J. (1997) NMR solution structures of stereoisometric covalent polycyclic aromatic carcinogen-DNA adduct: principles, patterns, and diversity. *Chem. Res. Toxicol.*, **10**, 111–146.

34. Zhou,L., Rajabzadeh,M., Traficante,D.D. and Cho,B.P. (1997) Conformational Heterogeneity of Arylamine-Modified DNA: 19F NMR Evidence. *J. Am. Chem. Soc.*, **119**, 5384–5389.

35. Meneni,S., Shell,S.M., Zou,Y. and Cho,B.P. (2007) Conformation-specific recognition of carcinogen-DNA adduct in escherichia coli nucleotide excision repair. *Chem. Res. Toxicol.*, **20**, 6–10.

36. Blazy,B., Culard,F. and Maurizot,J.C. (1987) Interaction between the cyclic AMP receptor protein and DNA. *J Mol Biol*, **195**, 175–183.

37. Connor,F., Cary,P.D., Read,C.M., Preston,N.S., Driscoll,P.C., Denny,P., Crane-Robinson,C. and Ashworth,A. (1994) DNA binding and bending properties of the postmeiotically expressed Sry-related protein Sox-5. *Nucleic Acids Res*, **22**, 3339–3346.

38. Meierhans,D., Sieber,M. and Allemann,R.K. (1997) High affinity binding of MEF-2C correlates with DNA bending. *Nucleic Acids Res*, **25**, 4537–4544.

39. Cai,Y., Kropachev,K., Xu,R., Tang,Y., Kolbanovskii,M., Kolbanovskii,A., Amin,S., Patel,D.J., Broyde,S. and Geacintov,N.E. (2010) Distant neighbor base sequence context effects in human nucleotide excision repair of a benzo[a]pyrene-derived DNA lesion. *J Mol Biol*, **399**, 397–409.

40. Suh,M., Ariese,F., Small,G.J., Jankowiak,R., Liu,T.M. and Geacintov,N.E. (1995) Conformational studies of the (+)-trans, (-)-trans, (+)-cis, and (-)-cis adducts of anti-benzo[a]pyrene diolepoxide to N2-dG in duplex oligonucleotides using polyacrylamide

gel electrophoresis and low-temperature fluorescence spectroscopy. *Biophys. Chem.*, **56**, 281–296.

41. Tsao,H., Mao,B., Zhuang,P., Xu,R., Amin,S. and Geacintov,N.E. (1998) Sequence dependence and characteristics of bends induced by site-specific polynuclear aromatic carcinogen-deoxyguanosine lesions in oligonucleotides. *Biochemistry*, **37**, 4993–5000.

42. Abuaf,P., Hingerty,B.E., Broyde,S. and Grunberger,D. (1995) Solution Conformation of the N-(Deoxyguanosin-8-yl)aminofluorene Adduct opposite Deoxyinosine and Deoxyguanosine in DNA by NMR and Computational Characterization. *Chem. Res. Toxicol.*, **8**, 369–378.

43. Mao,B., Hingerty,B.E., Broyde,S. and Patel,D.J. (1998) Solution Structure of the Aminofluorene [AF]-Intercalated Conformer of the syn-[AF]-C 8-dG Adduct Opposite dC in a DNA Duplex †. *Biochemistry*, **37**, 81–94.

44. O'Handley,S.F., Sanford,D.G., Xu,R., Lester,C.C., Hingerty,B.E., Broyde,S. and Krugh,T.R. (1993) Structural characterization of an N-acetyl-2-aminofluorene (AAF) modified DNA oligomer by NMR, energy minimization, and molecular dynamics. **32**, 2481–2497.

45. Eckel,L.M. and Krugh,T.R. (1994) Structural Characterization of Two Interchangeable Conformations of an N-2-Aminofluorene-Modified DNA Oligomer by NMR and Energy Minimization. *Biochemistry*, **33**, 13611–13624.

46. Yang,Z., Colis,L.C., Basu,A.K. and Zou,Y. (2005) Recognition and incision of gamma-radiation-induced cross-linked guanine-thymine tandem lesion G[8,5-Me]T by UvrABC nuclease. *Chem. Res. Toxicol.*, **18**, 1339–1346.

47. Zou,Y., Shell,S.M., Utzat,C.D., Luo,C., Yang,Z., Geacintov,N.E. and Basu,A.K. (2003) Effects of DNA adduct structure and sequence context on strand opening of repair intermediates and incision by UvrABC nuclease. *Biochemistry*, **42**, 12654–12661.

48. Ruan,Q., Liu,T., Kolbanovskiy,A., Liu,Y., Ren,J., Skorvaga,M., Zou,Y., Lader,J., Malkani,B., Amin,S., *et al.* (2007) Sequence context- and temperature-dependent nucleotide excision repair of a benzo[a]pyrene diol epoxide-guanine DNA adduct catalyzed by thermophilic UvrABC proteins. *Biochemistry*, **46**, 7006–7015.

49. Koehl,P., Valladier,P., Lefèvre,J.-F. and Fuchs,R.P.P. (1989) Strong structural effect of the position of a single acetylaminofluorene adduct within a mutation hot spot. *Nucleic Acids Res*, **17**, 9531–9541.

50. Lankaš,F., Sponer,J., Langowski,J. and Cheatham,T.E.,III (2003) DNA Basepair Step Deformability Inferred from Molecular Dynamics Simulations. *Biophys. J.*, **85**, 2872–2883.

51. McConnell,K.J. and Beveridge,D.L. (2001) Molecular dynamics simulations of B′-DNA: sequence effects on A-tract-induced bending and flexibility. *J Mol Biol*, **314**, 23–40.

52. Lyubchenko,Y.L., Shlyakhtenko,L.S., Appella,E. and Harrington,R.E. (1993) CA runs increase DNA flexibility in the complex of .lambda. Cro protein with the OR3 site. *Biochemistry*, **32**, 4121–4127.

53. Liu,Y., Reeves,D., Kropachev,K., Cai,Y., Ding,S., Kolbanovskiy,M., Kolbanovskiy,A., Bolton,J.L., Broyde,S., Van Houten,B., *et al.* (2011) Probing for DNA damage with β-hairpins: Similarities in incision efficiencies of bulky DNA adducts by prokaryotic and human nucleotide excision repair systems in vitro. *DNA Repair (Amst)*, **10**, 684–696.

54. Zou,Y., Luo,C. and Geacintov,N.E. (2001) Hierarchy of DNA Damage Recognition in EscherichiacoliNucleotide Excision Repair †. *Biochemistry*, **40**, 2923–2931.

55. Truglio,J.J., Croteau,D.L., Van Houten,B. and Kisker,C. (2006) Prokaryotic nucleotide excision repair: the UvrABC system. *Chem. Rev.*, **106**, 233–252.

56. Van Houten,B., Croteau,D.L., DellaVecchia,M.J., Wang,H. and Kisker,C. (2005) 'Close-fitting sleeves': DNA damage recognition by the UvrABC nuclease system. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis*, **577**, 92–117.

57. Van Houten,B. and Snowden,A. (1993) Mechanism of action of theEscherichia coli UvrABC nuclease: Clues to the damage recognition problem. *BioEssays*, **15**, 51–59.

58. Malta,E., Verhagen,C.P., Moolenaar,G.F., Filippov,D.V., van der Marel,G.A. and Goosen,N. (2008) Functions of base flipping in E. coli nucleotide excision repair. *DNA Repair (Amst)*, **7**, 1647–1658.

Supplemental Data

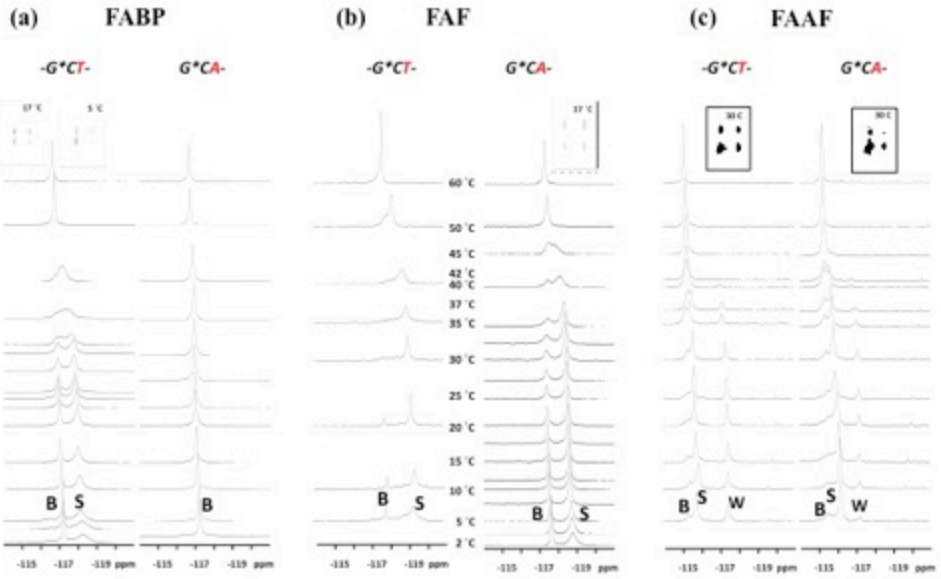


Figure S1: Temperature dependent dynamic 1D ^{19}F NMR spectra of (a) FABP-, (b) FAF-, and (c) FAAF-modified 11-mer duplexes on the G*CT and G*CA sequence contexts. The signals were assigned as B-, S- and W-conformers based on H/D solvent effect, adduct-induced circular dichroism ($ICD_{290\text{-}350nm}$) and chemical shift patterns. Insets are the NOESY spectra, the off-diagonal signals indicate that the conformers are in exchangeable equilibrium with each other

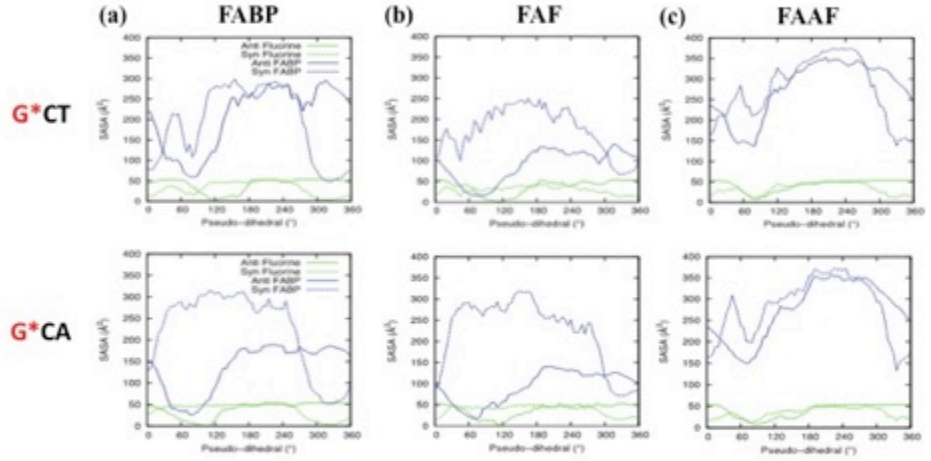


Figure S2: Solvent Accessible Surface Area (SASA, Å^2) calculated for the G*CT and G*CA duplexes modified by either a) FABP, b) FAF, or c) FAAF adduct. Results for the full adduct (blue line) and fluorine atom (green line) shown. In the *anti*-G* PMFs, the average solvent accessible surface areas (SASA) of both the lesion and fluorine atom are high in the vicinity of the B state (330 ~ 360°) whereas in the *syn* PMFs, the SASA values are low in the regions of the minima corresponding to the S-state (330 ~ 360°). Moreover, the *syn* PMFs of FAAF exhibit SASA minima in 60 ~ 120°, which encompasses the W state.

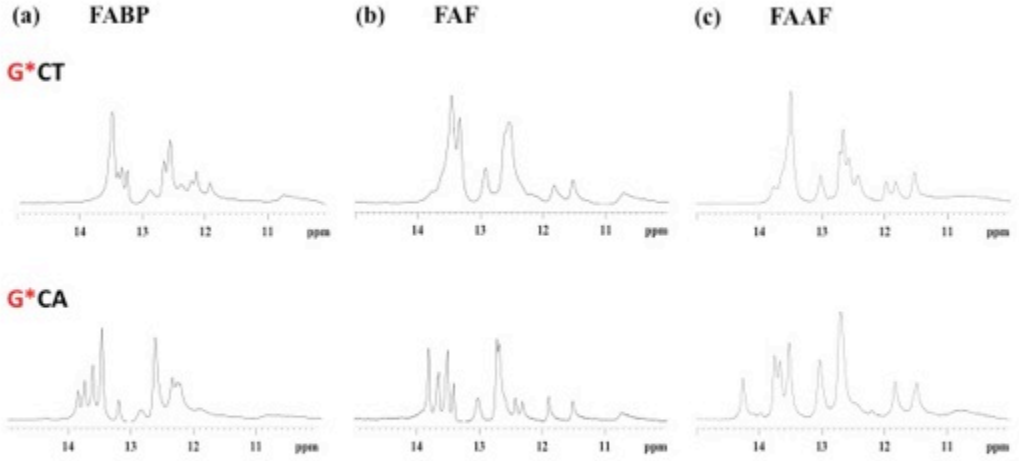


Figure S3: Imino proton region (10-15 ppm) of proton NMR of G*CT and G*CA duplexes modified by a) FABP, b) FAF and c) FAAF at 5 ˚C. The spectra displays a mixture of broad imino signals arising not only from those involved in Watson-Crick hydrogen bonds (12 ~ 14 ppm), but also from the lesion site and its vicinity (11-12 ppm).

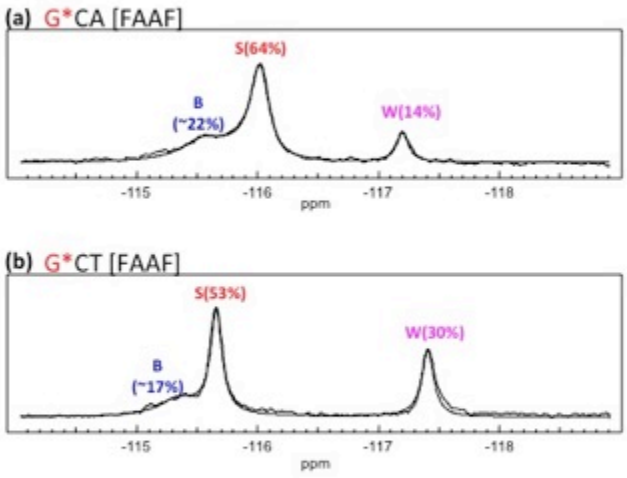


Figure S4: Line simulation of fully paired FAAF-modified (a) G*CA and (b) G*CT duplexes at 20 ˚C. The simulations were performed to calculate the population ratio of B-, S- and W-conformers in FAAF modified G*CA and G*CT duplexes using WINDNMR-Pro (version 7.1.6; *J. Chem. Educ.* Software Series; Reich, H. J., University of Wisconsin, Madison, WI).

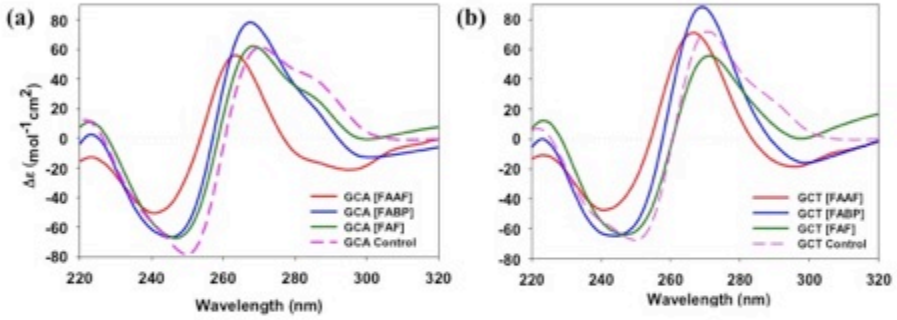


Figure S5: Overlay CD spectra of FABP-, FAF-, and FAAF-modified 11-mer duplexes in (a) the G*CA and (b) G*CT sequence contexts. Both unmodified and modified duplexes displayed a positive and negative ellipticity at around 270 and 250 nm, respectively, which is an S-curve characteristic of a B-type DNA double helix. In addition, the modified duplexes displayed significant blue shifts relative to the unmodified duplex and the results are tabulated in Table 1.

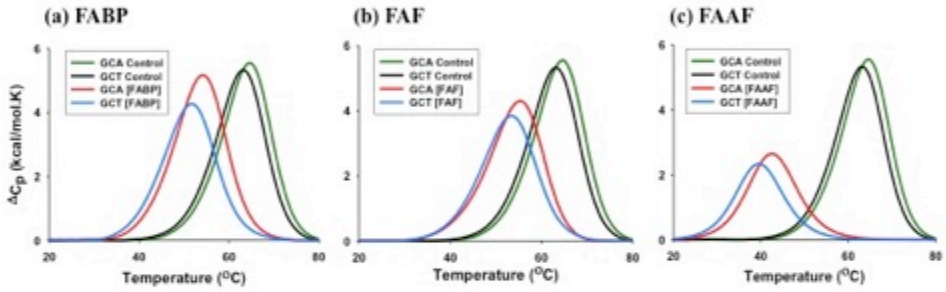


Figure S6: DSC thermograms of (a) FABP- (b) FAF- and (c) FAAF-modified relative to control duplexes. GCA control duplex (Green); GCT Control Duplex (Black); GCA-modified duplex (Red); GCT-modified duplex (Blue). The area of the curve is proportional to the transition heat, which was normalized for the number of moles of the sample to provide the transition enthalpy, ΔH. T_m was the temperature at half the peak area. The presence of lesion (FABP, FAF or FAAF) resulted into the thermal and thermodynamic destabilization of the duplexes as evident from their reduced Tm and area under the curve values(ΔH).

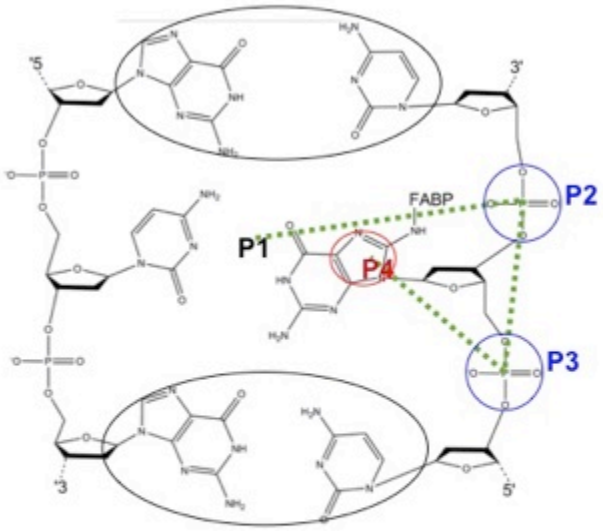


Figure S7: Definition of the reaction coordinate: The Pseudo-dihedral angle P1-P2-P3-P4 is an angle between the Center Of Mass (COM) of the following atom groups: P1 is defined as a COM of the base pairs above and below the flipping guanine base (3'-C5:G7, 5'-C7:G5), P2 and P3 are defined as the COM of the phosphate group in 3'- and 5'- direction from the flipping base and P4 is taken as a COM of the 5-membered ring of the flipping guanine base G6. Based on the above definition 330~360° in *anti* simulations corresponds to the Watson-Crick base paired B state, 330~360° in *syn* simulations corresponds to the S state, and 60~90° in *syn* simulations corresponds to W state. Major groove path ranges from 180° to 360° and the minor groove path ranges from 0° to 180°.

Note: Figures S8 to S11 present "NOE distance" plots extract from the PMFs. To identify if the PMFs were sampling experimentally relevant conformations NOE distance plots, based on published NMR experiments on different lesions and sequences were created. These plots involved determining the average distances between the atom pairs identified in the NMR experiments for the individual windows in the PMFs, such that regions where shorter distances in the PMFs occur correspond to conformations that are consistent with the experimental data.

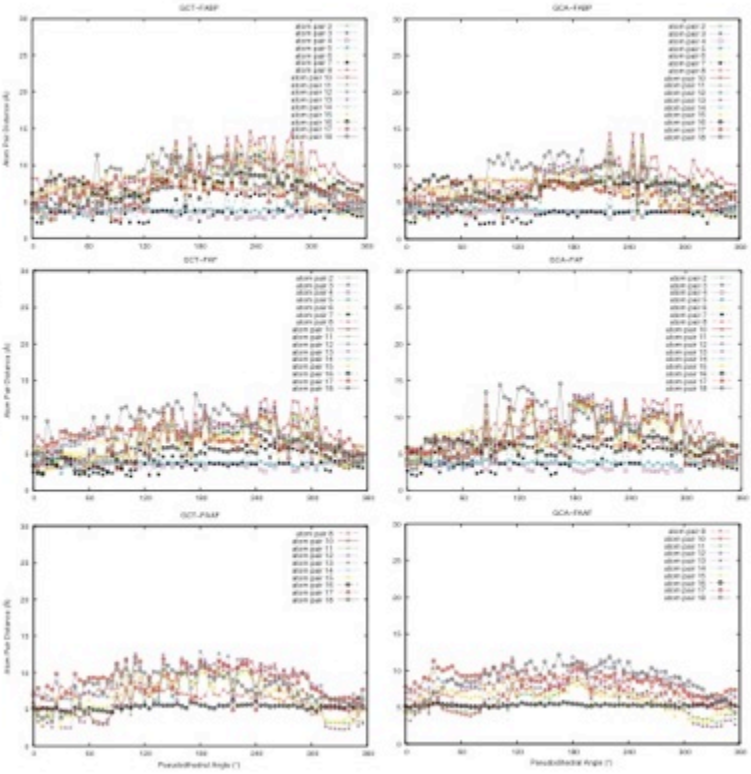


Figure S8: Distances of atom pairs for comparison to those listed in Table 3 of Mao et al. Biochemistry,1998, 37, 95. Atom pairs are defined in Table S3. Distances are averages calculated over last 3 nanosecond trajectory of anti simulations. 330~360° region is corresponding to B-state.

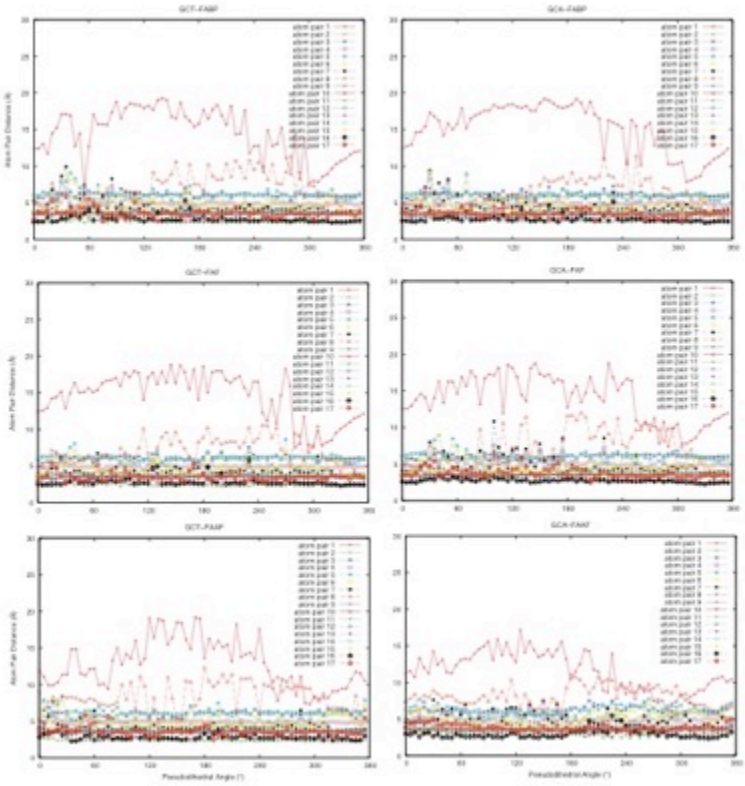


Figure S9: Distances of atom pairs for comparison to those listed in Table IV of O'Handley et al. Biochemistry,1993, 32, 2481. Atom pairs are defined in Table S4. Distances are averages calculated over last 3 nanosecond trajectory of syn simulations. 330~360° region is corresponding to S-state.

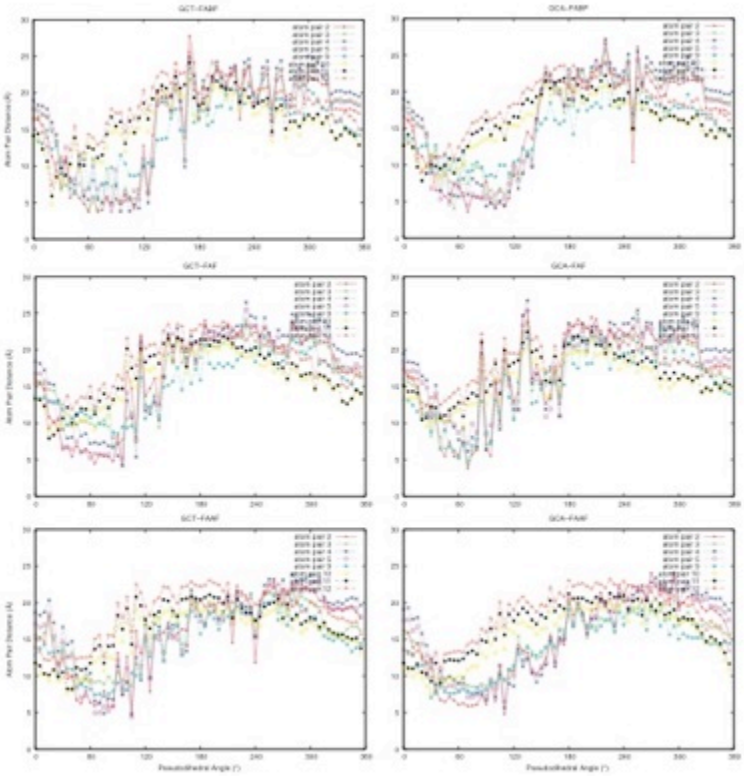


Figure S10: Distances of atom pairs for comparison to those listed in Table 4 of Eckel and Krugh. Biochemistry,1994, 33, 13611. Atom pairs are defined in Table S5. Distances are averages calculated over last 3 nanosecond trajectory of syn simulations. 330~360° region is corresponding to S-state.

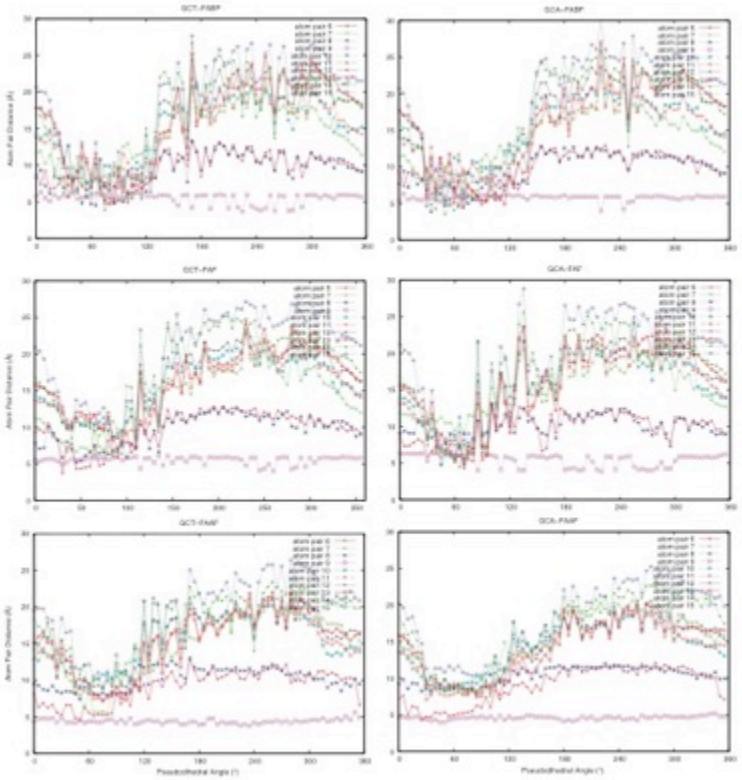


Figure S11: Distances of atom pairs for comparison to those listed in Table 5 Abuaf et al, Chem. Res. Toxicol. 1995, 8, 369. Atom pairs are defined in Table S6. Distances are averages calculated over last 3 nanosecond trajectory of syn simulations. 60~90° region is corresponding to W-state.

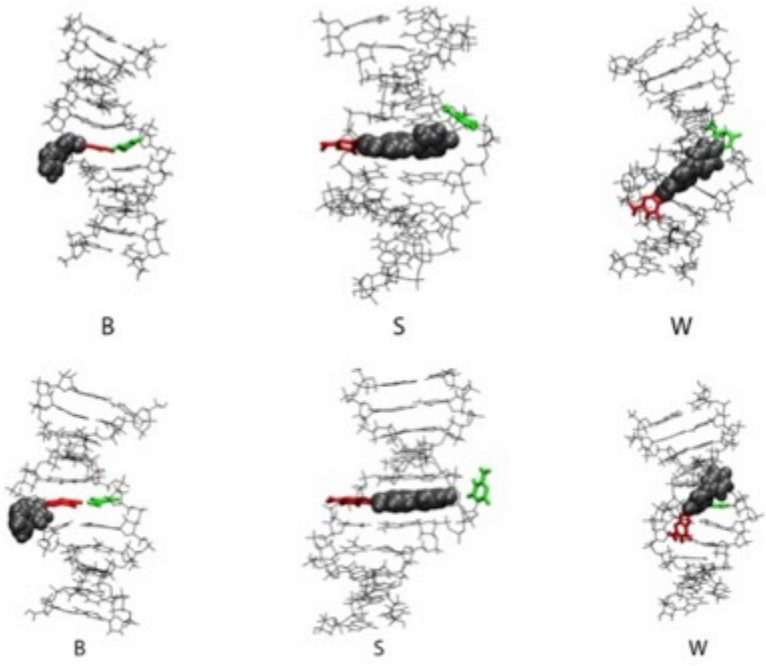


Figure S12: FABP B, S and W conformations for G*CT (upper) and G*CA (lower)

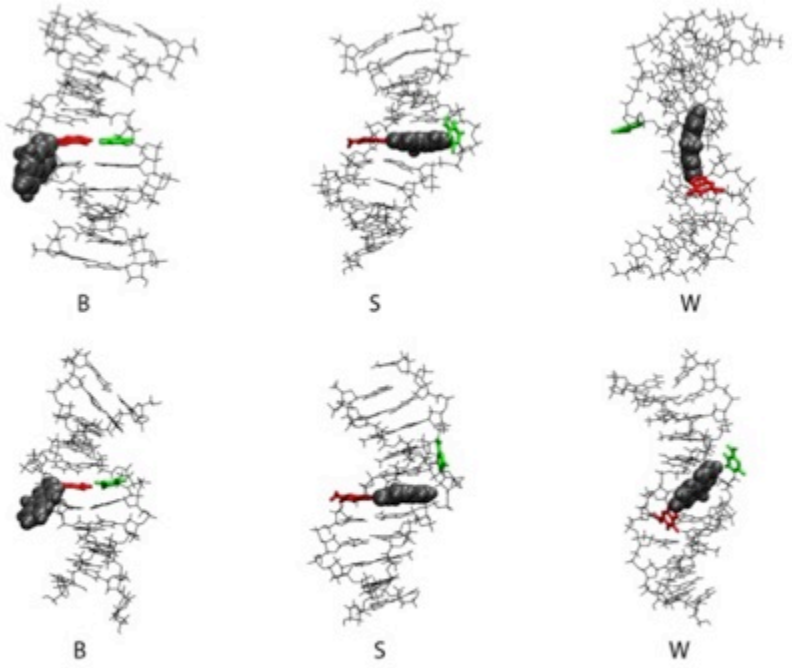


Figure S13: FAF B, S and W conformations for G*CT (upper) and G*CA (lower)

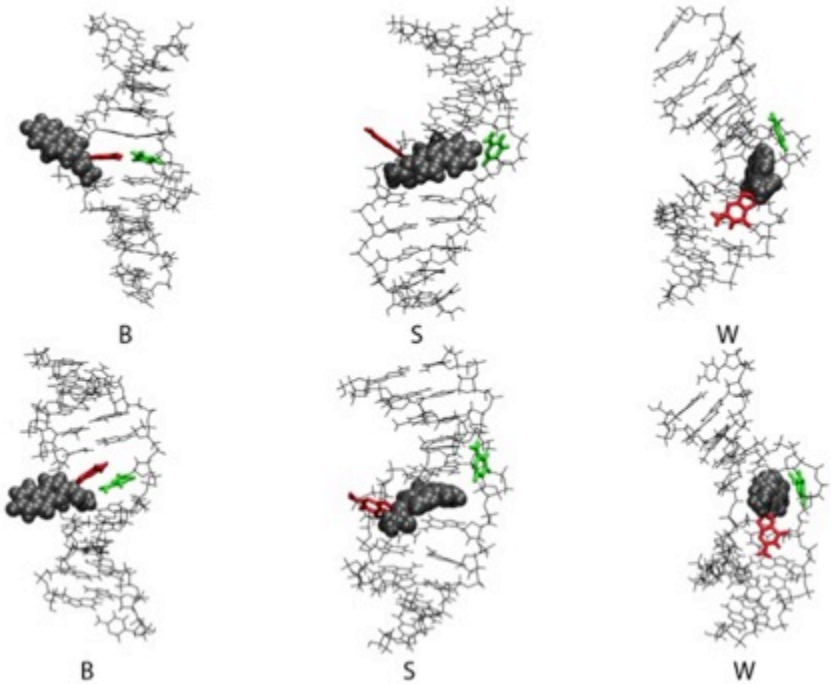


Figure S14: FAAF B, S and W conformations for G*CT (upper) and G*CA (lower)

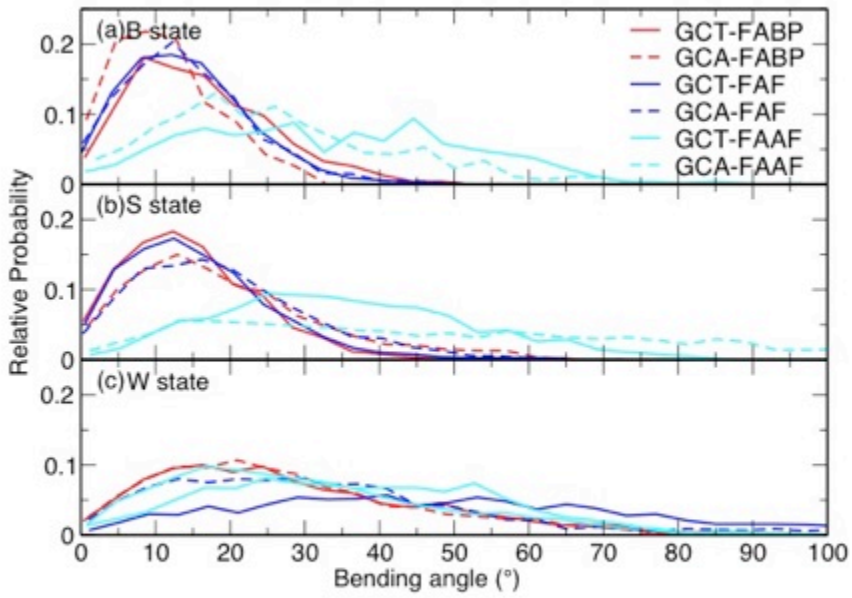


Figure S15: Probability distributions of the helical bending angles for the DNA-adducts in both the GCT and GCA sequences and for the (a) B-, (b) S- and (c) W-states. In the B- and S-states the extent of bending is significantly larger with FAAF (cyan) versus FABP (red) and FAF (blue). Moreover, the extent of bending in G*CA sequence is more when the lesions are in the S-state.

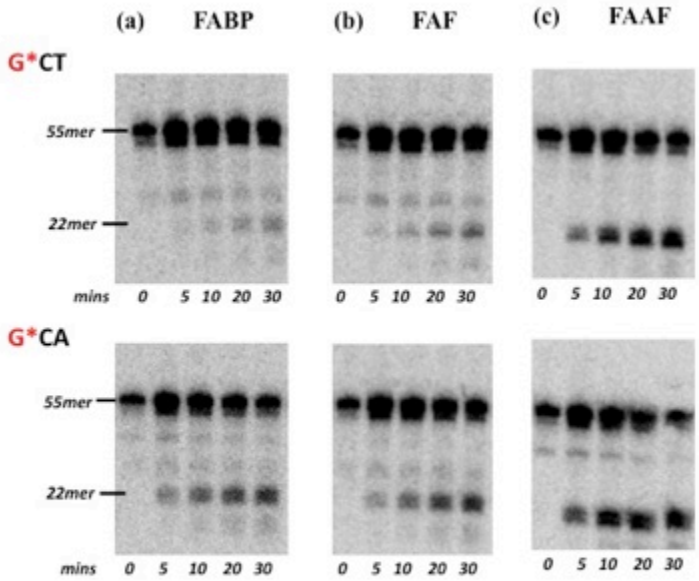


Figure S16: The 5'-terminally labeled G*CT and G*CA DNA substrates (2 nM) modified by either a) FABP, b) FAF or c) FAAF were incubated with UvrABC (UvrA, 10 nM, UvrB, 250 nM, and UvrC, 100 nM) in UvrABC reaction buffer at 37°C for the time period mentioned above. The incision products were then analyzed on a 12% polyacrylamide sequencing gel under denaturing condition. The 55-mer represents the intact DNA substrates, and the 22mer represent the 5'-incised DNA fragments. The results are tabulated and relative incision efficiencies are plotted in Figure 6.

Table S1: Thermal and thermodynamic parameters from UV-melting curves

	CCATC*G*CA*ACC GGTAGCGTTGG					CCATC*G*CT*ACC GGTAGCGATGG			
	$-\Delta H$ *kcal/mol*	$-\Delta S$ *eu*	$-\Delta G_{37}$ *kcal/mol*	T_m[b] *°C*		$-\Delta H$ *kcal/mol*	$-\Delta S$ *eu*	$-\Delta G_{37}$ *kcal/mol*	T_m[b] *°C*
Control[a]	87.6	241.3	12.7	60.6	**Control**	90.9	251.9	12.7	59.8
FABP[a]	73.8	207.6	9.4	49.8	**FABP**	68.5	192.7	8.7	47.2
FAF[a]	68.5	189.8	9.6	51.5	**FAF**	70.1	196.0	9.3	49.7
FAAF	54.2	146.2	8.8	50.9	**FAAF**	52.6	144.0	8.0	45.7
	$\Delta\Delta H$[c] *kcal/mol*	$\Delta\Delta S$[d] *eu*	$\Delta\Delta G_{37}$[e] *kcal/mol*	ΔT_m[f] *°C*		$\Delta\Delta H$[c] *kcal/mol*	$\Delta\Delta S$[d] *eu*	$\Delta\Delta G_{37}$[e] *kcal/mol*	ΔT_m[f] *°C*
FABP	13.8	33.7	3.3	-10.8	**FABP**	22.4	59.2	4.0	-12.6
FAF	19.1	51.5	3.1	-9.1	**FAF**	20.8	55.9	3.4	-10.1
FAAF	33.4	95.1	3.9	-9.7	**FAAF**	38.3	107.9	4.7	-14.1

a) The results of curve fit and T_m-lnC_t dependence were within 15% of each other, and these numbers are averages of the two methods. The average standard deviations for $-\Delta G$, $-\Delta H$, and T_m are ±0.2, ±6.2, and ±0.4, respectively.

b) T_m values at 0.1mM extrapolated from these two methods.

c) $\Delta\Delta H$ = ΔH (modified duplex) - ΔH (control duplex).

d) $\Delta\Delta S$= ΔS (modified duplex) - ΔS (control duplex).

e) $\Delta\Delta G$ = ΔG (modified duplex) - ΔG (control duplex).

f) ΔTm = Tm (modified duplex) - Tm (control duplex).

Table S2: Kinetic and Thermodynamic parameters from dynamic ^{19}F NMR

	FABP		FAF	
Sequence[a]	-G*CA-	-G*CT-	-G*CA-	-G*CT-
Ratio B:S[b] (20°C)	~100:0	40:60	34:66	10:90
$\Delta G_{B\leftrightarrow S}$ (20°C) kcal/mol[c]	>2.0[j]	-0.24	-0.39	-1.28
$\Delta H_{B\leftrightarrow S}$ kcal/mol[d]	~	3.63[k]	4.04	6.18
$\Delta S_{B\leftrightarrow S}$ Cal mol^{-1}K^{-1}[e]	~	11.5[k]	15.6	25.3
k_c (s^{-1})[f]	~	850[k]	921	~
$k_{30°C}$ (s^{-1})[g]	~	570[k]	300	~
τ (ms)[h]	~	1.75[k]	3.33	~
$\Delta G^{\ddagger}_{B\leftrightarrow S}$ (20°C) kcal/mol[i]	~	14.1[k]	14.3	~
$\Delta H^{\ddagger}_{B\leftrightarrow S}$ kcal/mol[j]	~	11.5[k]	13.7	~

a. The trimer portion of duplexes used in the present study. See Figure 1 for full sequence contexts. G*= FABP or FAF. b B/S conformer ratios by integration of ^{19}F signals at 20 °C. c. The energy difference between the B-and S-conformers: ΔG°= -RT ln K [K_{eqv}= S/B]. d,e.Enthalpy and entropy differences between the B- and S-conformers were estimated from the van't Hoff plots of the B/S-ratios vs temperature. f. Rate constants at a coalescence temperature, a lower limit on the exchange rate between the two conformers. K_c =2.22$\Delta\nu$ (the chemical shift difference between the two signals in Hz at slow exchange, i.e., at 5 °C). g. Rate constant at 30°C. The data were obtained by complete line shape analysis of temperatura dependent ^{19}F NMR results using the WINDNMR-Pro (see Experimental Procedures). h. Exchange time (1/k) indicates the amount of time the adduct spends in one conformation before jumping into another conformation. i,j The S/B interconversion energy barrier at 20 °C obtained from Eyring plots. k. Measured with D_2O as medium for better resolution of the B/S 19F NMR signals. l. 3% detection limit of ^{19}F NMR signal was assumed.

Table S3: Atom pairs used in NOE experiment vs PMF calculations for B-state NOE distance plot (Figure S10).[a]

atom pair #	NMR atom pair names	PMF atom pair names	NMR distance	atom pair #	NMR atom pair names	PMF atom pair names	NMR distance
2	AF(NH)-dG4(H1')	FAF(H81)-C5(H1')	3.0-5.0	3	AF(NH)-dG4(H2')	FAF(H81)-C5(H2')	4.0-6.0
4	AF(NH)-[AF]dG5(H1')	FAF(H81)-G6(H1')	3.0-5.0	5	AF(NH)-[AF]dG5(H2'')	FAF(H81)-G6(H2')	2.5-5.0
6	AF(NH)-dG4(H2'')	FAF(H81)-C5(H2'')	2.5-5.0	7	AF(NH)-[AF]dG5(H3')	FAF(H81)-G6(H3')	2.0-4.0
8	dG4(H1')-[AF]dG5(H4')	G5(H1')-G6(H4')	2.2-4.5	10	dG4(H1')-AF(H3)	G5(H1')-FAF(H10)	2.8-5.3
11	dG4(H2')-AF(H3)	C5(H2')-FAF(H10)	2.1-4.5	12	dG4(H2'')-AF(H1)	G5(H2'')-FAF(H12)	2.0-4.0
13	dG4(H2'')-AF(H3)	G5(H2'')-FAF(H10)	2.1-3.9	14	dG4(H3')-AF(H1)	G5(H3')-FAF(H12)	2.8-5.3
15	dG4(H3')-AF(H3)	C5(H3')-FAF(H10)	2.4-4.5	16	[AF]dG5(H3')-AF(H3)	G6(H3')-FAF(H10)	3.0-5.6
17	[AF]dG5(H3')-AF(H1)	G6(H3')-FAF(H12)	2.8-5.2	18	dC6(H5)-AF(H1)	C7(H5)-FAF(H12)	3.0-5.7

a) Experimental results from Table 3 of Mao et al, Biochemistry 1998, 37, 81-94.

Table S4: Atom pairs used in NOE experiment vs PMF calculations for S-state NOE distance plot (Figure S11).[a]

atom pair #	NMR atom pair names	PMF atom pair names	NMR distance	atom pair #	NMR atom pair names	PMF atom pair names	NMR distance
1	G5(H1)-A3(H3')	G6(H1)-T4(H3')	≤4.5	2	AAF(H1)-G5(H1')	FAAF(H12)-G6(H1')	≤3.2
3	A3(H2') to C4(H1')	T4(H1')-C5(H3')	≤3.5	4	G13(H3')-C14(H6)	G16(H3')-C17(H6)	≤3.5
5	C14-(H4')-G15(H8)	C17(H4')-G18(H8)	≤4.0	6	A3(H2'')-C4(H6)	T4(H2'')-C5(H6)	≥3.5
7	G13(H8)-C14(H5)	G16(H8)-C17(H5)	≤5.0	8	AAF(H1)-C4(H2')	G6(H12)-C5(H2')	≤5.0
9	G13(H8)-G13(H2')	G18(H8)-G18(H2')	≤2.5	10	A3(H8)-A3(H2')	T8(H8)-T8(H2')	≤2.5
11	T12(H2')-G13(H8)	T4(H2')-G16(H8)	≥2.6	12	T12(H2'')-G13(H8)	A15(H2'')-G16(H8)	≥2.6
13	G13(H2')-C14(H6)	G16(H2')-C17(H6)	≥2.6	14	G13(H2'')-C14(H6)	G16(H2'')-C17(H6)	≥2.8
15	C14(H1')-G15(H8)	C17(H1')-G18(H8)	≥2.8	16	A3(H2')-C4(H6)	T4(H2')-C5(H6)	≥3.5
17	C6(H2'')-A7(H8)	C7(H2'')-T8(H8)	≥2.6				

a) Experimental results from Table IV of O'Handley et al. Biochemistry, 1993, 32, 2481.

Table S5: Atom pairs used in NOE experiment vs PMF calculations for S-state NOE distance plot (Figure S12).[a]

atom pair #	NMR atom pair names	PMF atom pair names	NMR distance	atom pair #	NMR atom pair names	PMF atom pair names	NMR distance
2	F(H7)-C15(H5)	G6(F22)-C17(H5)	3.6±0.7	3	F(H5)-C15(H1')	G6(H18)-C17(H1')	3.9±0.8
4	F(H7)-C15(H1')	G6(F22)-C17(H1')	3.0±0.6	5	F(H8)-C15(H1')	G6(H19)-C17(H1')	3.4±0.7
9	F(H4)-T16(H1')	G6(H11)-A18(H1')	3.1±0.6	10	F(H1)-A5(H1')	G6(H12)-G16(H1')	3.7±0.7
11	F(H1)-A5(H2'')	G6(H12)-G16(H2'')	3.1±0.6	12	F(H1)-A5(H3')	G6(H12)-G16(H3')	3.1±0.6

a) Experimental results from Table IV of Eckel et al, Biochemistry, 1994, 33, 13611.

Table S6: Atom pairs used in NOE experiment vs PMF calculations for W-state NOE distance plot (Figure S13).[a]

atom pair #	NMR atom pair names	PMF atom pair names	NMR distance	atom pair #	NMR atom pair names	PMF atom pair names	NMR distance
6	C7(H1')-AF(H1)	C7(H1')-FAF(H1)	3.0-5.0	7	G18(H1')-AF(H1)	G18(H1')-FAF(H12)	4.0-6.0
8	C7(H1')-AF(H3)	C7(H1')-FAF(H10)	3.5-5.5	9	G6(H1')-AF(H3)	G6(H1')-FAF(H10)	3.5-5.5
10	I17(H1')-AF(H4)	C17(H1')-FAF(H11)	4.0-6.0	11	G18(H1')-AF(H4)	G18(H1')-FAF(H11)	4.0-6.0
12	I17(H1')-AF(H5)	C17(H1')-FAF(18)	4.0-6.0	13	G18(H1')-AF(H5)	G18(H1')-FAF(H18)	4.0-6.0

a) Experimental results from Table IV of Abuaf et al, Chem. Res. Toxicol. 1995, 8, 369.

Table S7: Average Twist, Roll and Tilt Values for base pair 8 in the studied systems for the B- and S-states. Averages over the average results from the individual windows of the PMFs defining the B- and S-states as described in the text. Errors represent standard deviations.

	FABP		**FAF**		**FAAF**	
	G*CT	**G*CA**	**G*CT**	**G*CA**	**G*CT**	**G*CA**
Twist						
B-state	37.8±4.7	35.9±4.1	37.0±5.1	36.0±4.1	38.8±5.2	34.6±5.0
S-state	37.8±5.0	35.7±4.5	38.8±4.9	36.2±4.3	37.5±4.5	34.8±4.8
Roll						
B-state	3.4±7.6	4.2±5.5	5.6±7.8	4.7±5.8	1.4±7.8	6.5±6.7
S-state	2.7±7.7	5.1±6.1	2.4±7.7	4.2±5.7	1.6±7.1	4.8±6.2
Tilt						
B-state	-1.0±4.8	-3.9±4.2	-0.8±5.3	-4.2±4.0	-1.9±4.9	-4.6±4.1
S-state	-0.6±5.0	-4.6±4.0	-1.3±4.9	-4.3±4.0	-0.3±4.7	-4.5±4.2

CHAPTER 3

STRUCTURAL AND THERMODYNAMIC INSIGHT INTO *ESCHERICHIA COLI* UVRABC-MEDIATED INCISION OF CLUSTER DIACETYLAMINOFLUORENE ADDUCTS ON THE *NAR*I SEQUENCE

Abstract

Cluster DNA damage refers to two or more lesions in a single turn of the DNA helix. Such clustering may occur with bulky DNA lesions, which may be responsible for their sequence-dependent repair and mutational outcomes. Here we prepared three 16-mer cluster duplexes in which two fluoroacetylaminofluorene adducts (dG-FAAF) are separated by zero, one, and two nucleotides in the *Escherichia coli Nar*I mutational hot spot (5′-CTCTCG$_1$G$_2$CG$_3$CCATCAC-3′): 5′-CG$_1$*G$_2$*CG$_3$CC-3′, 5′-C- G$_1$G$_2$*CG$_3$*CC-3′, and 5′-CG$_1$*G$_2$CG$_3$*CC-3′ (G* = dG-FAAF), respectively. We conducted spectroscopic, thermodynamic, and molecular dynamics studies of these di-FAAF duplexes, and the results were compared with those of the corresponding mono-FAAF adducts in the same *Nar*I sequence [Jain, V., *et al.* (2012) Nucleic Acids Res. 40, 3939–3951]. Our nucleotide excision repair results showed the diadducts were more reparable than the corresponding monoadducts. Moreover, we observed dramatic flanking base sequence effects on their repair efficiency in the following order: *Nar*I-G$_2$G$_3$ > *Nar*I-G$_1$G$_3$ > *Nar*I-G$_1$G$_2$. The nuclear magnetic resonance, circular dichroism, ultraviolet melting, and molecular dynamics simulation results revealed that in contrast to the monoadducts, diadducts produced a synergistic effect on duplex destabilization. In addition, dG-FAAF at G$_2$G$_3$ and G$_1$G$_3$ destacks

the neighboring bases, with greater destabilization occurring with the former. Overall, the results indicate the importance of base stacking and related thermal and thermodynamic destabilization in the repair of bulky cluster arylamine DNA adducts.

Introduction

The human genome is susceptible to chemical assault from a range of internal and external sources, including UV-radiation and carcinogens. Adduct formation occurs throughout the genome: some sites showed higher levels of adduct formation yet little or no mutation induction and *vice versa* [1,2]. The so-called locally multiply damaged sites (LMDS) or clustered lesions are defined as two or more lesions occurring in a short stretch (<24 bp) of DNA [3]. In general, cluster lesions are considered to be more genotoxic than a single lesion. It also has been hypothesized that cluster lesions are less reparable than individual ones and that mutation of one lesion is synergized by the presence of another lesion in the vicinity [4,5]. The repair efficiency of clustered oxidative damage has been shown to depend on the nature of interlesion distance, sequence contexts, and their relative orientation [6]. For instance, Kalam *et al.* [4] have found a significant fraction of tandem mutations when a uracil was adjacent to 8-oxo-G. The presence of an apurinic site either 5′ or 3′ to 8-oxo-G increased the frequency of 8-oxo-G to T transversions, with the effect on the latter being much stronger.

Evidence suggests that a similar clustering occurs with bulky DNA lesions and may be responsible for their sequence-dependent repair and mutational outcomes. Fuchs and co-workers have shown that the binding spectrum of 2-acetylaminofluorene (AAF) is essentially random and all guanine residues exhibited equal reactivity toward *N*-AcO- AAF

[7,8]. Other factors such as repair and neighboring sequences may also play a role in the formation of mutation hot spots [9]. For example, 5-methylcytosine in CpG sites can modulate the distribution of guanine adduct formation by benzo[*a*]pyrene diolepoxides (BPDE), *N*-methyl-*N*-nitrosourea (MNU), and AAF [10]. As in oxidative damage, GC-rich regions and runs of contiguous guanines in the genome have a higher probability of producing bulky cluster adducts [10]. Cluster adducts also can be produced through a mixture of either an individual lesion or multiple lesions, and such adducts may have important ramifications in damage recognition and verification and repair outcomes [6].

Arylamines are implicated in the etiology of human breast, liver, and bladder cancers [11]. The prototype AAF is no longer a threat *via* human exposure; however, its amino and nitro derivatives are environmental pollutants. These compounds produce two major DNA adducts via *in vivo* activation: *N*-(2-deoxyguanosin-8-yl)-2-aminofluorene (dG-AF) and *N*-(2-deoxyguanosin-8-yl)-2-acetylaminofluorene (dG-AAF) (Figure 1a) (which have been used extensively as models for bulky lesions) [12]. We have shown that dG-AF exists in an equilibrium of external binding *anti* B-type and *syn* stacked S conformers. In the bulky *N*-acetylated AAF, however, the *syn* adduct also could exist in the minor groove binding "wedge (W) conformer" (Figure 1c) [13,14]. The population ratios of these different conformations were found to be strongly dependent on the nature of the flanking sequence context around the lesion [13-18].

A number of factors have been implicated as being important for the repair of bulky DNA lesions such as arylamines and polycyclic aromatic hydrocarbons [19-22]. They include disruption of Watson-Crick hydrogen bonding, DNA bending, thermo-dynamic destabilization, local conformational flexibility, *etc.* [20]. Striking sequence effects were

observed in the *Escherichia coli* UvrABC-mediated nucleotide excision repair (NER) of fluorinated dG-AF and dG-AAF [dG-FAF and dG-FAAF, respectively (Figure 1a)] when they were present at different guanines of the *Nar*I sequence (5'-$G_1G_2CG_3CC$-3'). dG-FAAF exhibited conformation-specific repair; i.e., the highly S/W conformeric G_3 and G_1 duplexes incise considerably more efficiently than the highly B conformeric G_2 duplex ($G_3 \sim G_1 > G_2$). The NER efficiency of the *N*-deacetylated dG-FAF was 2-3-fold lower, but it too displayed a strong sequence dependence [13]. Mu *et al.* have confirmed the earlier human NER results for dG-AF and dG-AAF published by Fuchs and co-workers and provided a molecular dynamics-based structural rationale for their sequence and lesion dependence in repair [23]. A similar sequence dependence of repair has been reported in the non-*Nar*I sequences as well [24-26].

Little is known about the structural and repair consequences of clustered bulky lesions. As elaborated above, previous structure-activity relationship studies have focused on singly modified sequences, presumably because of the methodological ease of dealing with a single lesion rather than multiple adducts [27,28]. We hypothesize that cluster arylamine lesions could be formed wherever multiple guanines are present in the genome, resulting in major structural and/or conformational consequences on repair and mutational outcomes. To that end, we prepared duplex NER substrates, in which the bulky dG-FAAF lesions are located on two (G_1G_2, G_2G_3, and G_1G_3) of the three guanines (G_1, G_2, and G_3) of the *Nar*I recognition sequence (Figure 1b). We conducted spectroscopic and melting experiments and molecular dynamics (MD) simulations for conformational and thermodynamic analyses and NER measurements using the *E. coli* UvrABC system. Our

results present the first structure-repair study of NER of bulky cluster lesions and its sequence dependence.

Materials and Methods

Caution: *2-Aminofluorene derivatives are mutagens and suspected human carcinogens and therefore must be handled with caution*.

Crude oligodeoxynucleotides (10 μmol scale) in desalted form were purchased from Eurofins MWG operon (Huntsville, AL). All HPLC solvents were purchased from Fisher Inc. (Pittsburgh, PA).

Preparation and Characterization of FAAF-Modified Diadduct Oligodeoxynucleotides

We reported previously the preparation of 16-mer *Nar*I oligodeoxynucleotides (5'-CTCTCG$_1$G$_2$CG$_3$CCATCAC-3'), in which each of the three guanine was modified by FAAF [13]. We also have documented that the incorporation of the fluorine atom at the longest axis does not affect the overall conformational and thermodynamic profiles of aminofluorene-modified duplexes (e.g., dG-FAF and dG-FAAF) [16,29]. FAAF-modified 16-mer diadduct oligodeoxynucleotides were prepared similarly with a longer reaction time (Figure S1 of the Supporting Information) [13,14,30]. Experimental details of FAAF modification and HPLC purification are provided in the Supporting Information.

The three di-FAAF-modified 16-mer sequences were each annealed with the corresponding complementary sequence (5'-GTGATGGC- GCCGAGAG-3') to form fully paired duplexes, designated *Nar*I- G$_1$G$_2$, *Nar*I-G$_2$G$_3$, and *Nar*I-G$_1$G$_3$ depending on the

location of the FAAF modification (Figure 1b). It should be noted that only one 16-mer oligodeoxynucleotide sequence was used with FAAF adducts at different G positions.

MALDI-TOF Characterization of Diadduct Oligodeoxynucleotides

The modified oligodeoxynucleotides were characterized by following the general procedures on 3'-5' and 5'-3' exonuclease digestion MALDI-TOF mass spectrometry [31,32]. In general, 1 μL of sample containing approximately 100-250 pmol of a modified oligodeoxynucleotide was used for digestion. In the 5'-3' digestion, 0.01 unit of bovine spleen phosphodiesterase (BSP) was added to a solution containing 1 μL of oligodeoxynucleotide and 7 μL of water. Similarly, 0.1 unit of snake venom phosphodiesterase (SVP) was added to a solution of 1 μL of oligodeoxynucleotide, 6 μL of ammonium citrate (100 mM, pH 9.4), and 6 μL of water for 3'-5' digestion. In each case, 1 μL of the reaction mixture was withdrawn at regular time intervals until the digestion rate was significantly reduced, and the reaction was quenched by mixing the aliquot with 1 μL of matrix (3-HPA and DHAC in a 1:1 ratio). The samples were then immediately spotted on a MALDI plate, and the spectra were recorded on the Shimadzu Axima Performance mass spectrometer equipped with a 50 Hz nitrogen laser in positive ion reflectron mode.
UV melting experiments were performed on a Cary100 Bio UV-vis spectrophotometer equipped with a 6 × 6 multicell block and a 1.0 cm path length. Circular dichroism (CD) measurements were taken on a Jasco J-810 spectropolarimeter equipped with a Peltier temperature controller. The details of UV melting and CD experiments are presented in the Supporting Information.

Dynamic ^{19}F NMR

Approximately 10 OD_{260} units (optical density per milliliter at 260 nm) of a di-FAAF 16-mer oligodeoxynucleotide was annealed with an equimolar amount of the complementary sequence (5'-GTGATGGCGCCGAGAG-3') to produce a fully paired duplex (Figure 1b). The duplex samples were then dissolved in 300 µL (~0.2 mM) of a pH 7.0 NMR buffer containing a 10% D_2O/90% H_2O mixture, 100 mM NaCl, 10 mM sodium phosphate, and 100 µM EDTA and filtered into a Shigemi tube through a 0.2 µm membrane filter. All ^{1}H and ^{19}F NMR results were recorded using an HFC probe on a Varian NMR spectrometer operating at 500.0 and 476.5 MHz, respectively, using the usual acquisition parameters described previously [16,29,33]. Imino proton spectra were recorded at 5°C using a WET-1D sequence and referenced relative to that of DSS. ^{19}F NMR spectra were acquired in the ^{1}H-decoupled mode and referenced relative to that of CFCl3 by assigning external C_6F_6 in C_6D_6 to −164.9 ppm. ^{19}F NMR spectra were measured between 5 and 65°C with an increment of 5-10°C. Temperatures were maintained by a Varian FTS control system.

Molecular Modeling

The diadducts were modeled in the BB, BS, SB, and SS conformations, where B and S represent the *anti* and *syn* orientations, respectively, of the glycosidic bond of the modified G base. The different B and S states will be termed glycosidic isomers. Their starting conformations were built in the following way. First, d[CTCTCG$_1$G$_2$CG$_3$CCATCAC] was built using 3DNA [34] as canonical B-DNA for the unmodified duplex, where the glycosidic torsion angles are in the *anti* conformation for G_1, G_2, and G_3. Coordinates were then read into CHARMM [35] and patched to attach the FAAF adduct to the C8 atoms in the appropriate

guanine bases. This yielded the BB conformations of *Nar*I-G_1G_3, *Nar*I-G_2G_3, and *Nar*I-G_1G_2. The S conformations were prepared by rotating the glycosidic torsion angle by 180° from their respective B-type models. This resulted in a total of 12 starting conformations of FAAF-modified diadducts, and an unmodified control DNA-duplex. Each system was solvated in a cubic box with TIP3P water [36] with the distance from DNA duplex to the edge of the box set to 10 Å. Thirty sodium ions were added to neutralize each system. All the models were subjected to extensive energy minimization and equilibration using GROMACS version 4.5.5 [37] with the latest CHARMM [35] nucleic acid force field [38]. The particle mesh Ewald [39] method was used to treat long-range electrostatics using a real-space cutoff of 10 Å, forth-order spline, and an approximately 1 Å grid spacing. Lennard-Jones interactions were switched off between 8 and 10 Å. The temperature was controlled at 300 K using velocity rescaling [40,41] with a stochastic term of 0.1 ps, and the pressure was maintained at 1 atm with the Parrinello-Rahman [42] barostat with a time constant of 2 ps. The LINCS algorithm [43,44] was employed to constrain all covalent bonds involving hydrogens, and the equations of motion were integrated with the Leapfrog algorithm with a time step of 2 fs. Each system was minimized with the steepest descent method for 1000 steps and then equilibrated by a 100 ps NVT simulation, followed by the 100 ns production runs with the Watson-Crick (WC) N1-N3 hydrogen bond restrained between 2.5 and 3.5 Å for the terminal base pairs. Coordinates were saved every 20 ps. The first 20 ns of each trajectory was discarded as equilibration as judged from the root-mean-square (rms) differences with respect to the starting structures. The remaining 80 ns was used for analyses. Curves+ [45] and CHARMM were used to perform bending, structure, and interaction energy analyses. Clustering analysis was conducted using the single-linkage

algorithm [46]. VMD [47] and PyMOL [48] were used for visualizing trajectories and preparing structural figures. Averages and standard errors were estimated by using eight block averages with a 10 ns duration.

Substrate Construction and UvrABC Protein Purification

DNA substrates of 55 bp containing two FAAF adducts simultaneously on two of the three guanine residues were constructed as previously described [49,50], and the details are presented in the Supporting Information. UvrA, UvrB, and UvrC proteins were overexpressed in *E. coli* and then purified as previously described [51]. The estimated purity of the three proteins was >95%. A Bio-Rad Protein Assay was used to determine the protein concentration with BSA as the standard based on the procedures recommended by the manufacturer.

Nucleotide Excision Assay and Quantification of Incision Products

The 5'-terminally labeled DNA substrates were incised by UvrABC as previously described [49,50]. The details of the incision assay and quantification are provided in the Supporting Information.

Results

Model Sequences

We previously reported that a simple mixing of the 16-mer *Nar*I oligodeoxynucleotide (5'-CTCTCG$_1$G$_2$CG$_3$CCATCAC-3') [13] with the chemically reactive FAAF

at room temperature produced an approximately equal mixture of mono-, di-, and triadducts. A longer reaction time increased the level of production of di- and triadducts, which were well separated via RP-HPLC as shown in Figure S1 of the Supporting Information. We grouped them into three peaks, 1-3 (28-35 min), 4-6 (42-60 min), and 7 (84 min), as mono-, di-, and triadducts, respectively, on the basis of the relative absorption intensity (1:2:3) of the adduct-induced shoulder in the 290-320 nm range [13]. Moreover, we assigned peaks 1-3 as G_1, G_3, and G_2, respectively, using the exonuclease-LC/MS methods. In the study presented here, we have characterized peaks 4-6 similarly using MALDI-TOF as described below. The diadducts were designated as *Nar*I- G_1G_2, *Nar*I-G_2G_3, or *Nar*I-G_1G_3, in which the numbers signify the positions of FAAF-modified guanines (Figure 1b).

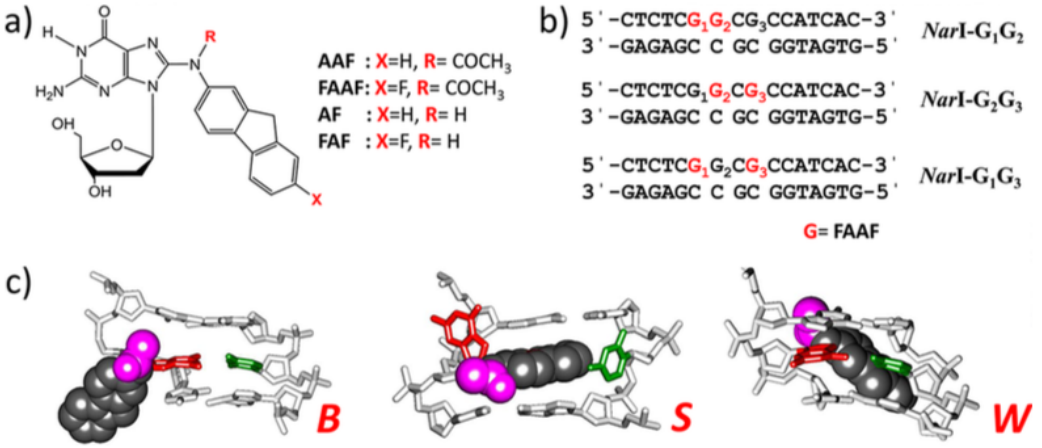


Figure 1. Adduct structures and sequences. (a) Chemical structures of AAF and FAAF adducts. (b) Fully paired 16-mer *Nar*I duplexes. (c) Major groove views of the central trimer segments of the B—S—W conformer equilibrium of FAAF-modified monoadduct duplexes. The modified dG and the complementary dC are shown as red and green sticks, respectively, and the aminofluorene moiety is highlighted with shiny gray CPK and the *N*-acetyl with pink CPK. In the B-type conformer, *anti* glycosidic FAAF-dG maintains Watson—Crick hydrogen bonds, thereby placing the carcinogen moiety in the major groove. The carcinogens in the S and W conformers stack into the helix and wedged into the minor groove, respectively, with the modified dG in the *syn* conformation.

MALDI-TOF Characterization

The molecular weights of di-FAAF modified 16-mers were measured by MALDI-TOF prior to sequence verification by exonuclease digestion. Figure 2 (see Figure S2 of the Supporting Information for complete time ranges) represents the 3' → 5' (Figure 2a) and 5'

→ 3’ (Figure 2b) exonuclease digestion fragments for peak 4 at different time intervals. In general, the 3’ → 5’ digestion was much faster than the 5’ → 3’ digestion. The ions observed at m/z 5257.2 at 30 s (Figure 2a) and m/z 5257.5 at 5 min (Figure 2b) represent $(M + H)^+$ ions of the doubly modified *Nar*I 16-mer oligodeoxynucleotide. An increase in the incubation time leads to the digestion of subsequent unmodified bases. However, 3’ and 5’ digestions were significantly slowed at m/z 3170.2 and 3781.0, respectively. The ion observed at m/z 3170.2 at 15 min of 3’ digestion represents the $(M + H)^+$ ion formed from the doubly modified 5’-CTCTCG$_1$G$_2$CG$_3$-3’ fragment (see the inset of Figure 2 for theoretical MW values). These results indicate a FAAF modification at the G_3 position of the *Nar*I oligodeoxynucleotide. In 5’ digestion, the ion observed at m/z 3781.0 at 80 min (Figure 2b) represents the $(M + H)^+$ ion formed from the doubly modified 5’-G$_1$G$_2$CG$_3$CCATCAC-3’ fragment, suggesting another FAAF at the G_1 position. The ion at m/z 4070.2 of 5’ digestion (Figure 2b) indicates that the digestion slowed one base prior to the modification. Taken together, peak 4 was assigned as *Nar*I-G$_1$G$_3$ (5’-CTCTCG$_1$[FAAF]G$_2$CG$_3$[FAAF]CCATCAC-3’). Similarly, peaks 5 and 6 were characterized as *Nar*I-G$_2$G$_3$ and *Nar*I-G$_1$G$_2$, respectively, and the results are shown in Figures S3 and S4 of the Supporting Information, respectively. The HPLC elution pattern and mass characterization of the FAAF diadduct 16-mer oligodeoxynucleotides in this study is similar to those of the FAF diadduct 12-mer *Nar*I sequence (5’-CTCG1G2CG3CCATC-3’) reported by Gao *et al.* [31].

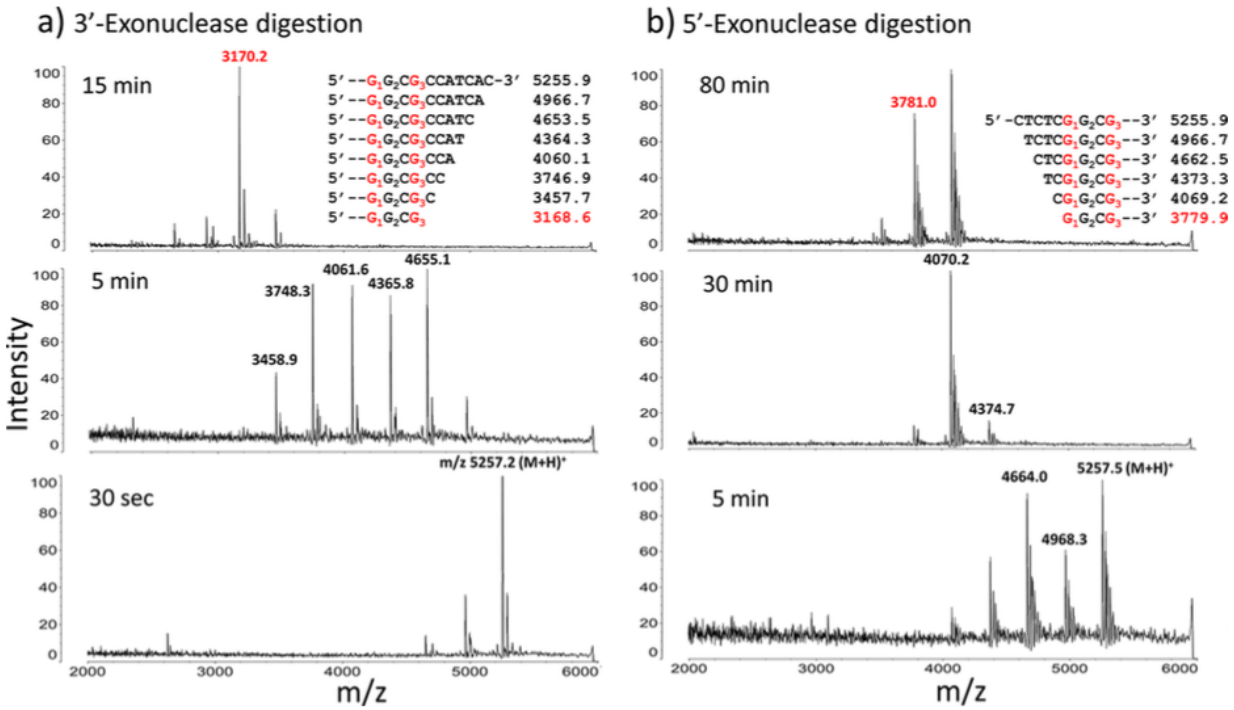


Figure 2. MALDI-TOF mass spectra of the hydrolysis fragments formed from exonuclease digestion of peak 4 (see Figure S1 of the Supporting Information). (a) 3′-Exonuclease digestion spectrum at 30 s (bottom), 5 min (middle) and 15 min (top). (b) 5′-Exonuclease digestion spectrum at 5 min (bottom), 30 min (middle), and 80 min (top). Insets show the theoretical molecular weights of the fragments that should form after the 3′- and 5′-exonuclease digestion of peak 4 if the oligodeoxynucleotide is modified at positions G_1 and G_3.

Circular Dichroism

Figure 3a shows an overlay of the CD spectra for the three FAAF diadduct *Nar*I-G_1G_3, *Nar*I-G_2G_3, and *Nar*I-G_1G_2 duplexes relative to the unmodified control (brown) at ~35 μM. Unlike that of FAF and FABP, the induced CD in the 290–320 nm range ($ICD_{290-320}$) for FAAF- induced B-S-W conformational heterogeneity has not been established. Nonetheless, comparative pattern analysis provided some valuable conformational insight. Unmodified and FAAF diadduct duplexes both displayed a positive and negative ellipticity at around 270 and 250 nm, respectively, which is characteristic of a regular B-form DNA double helix. The modified duplexes consistently exhibited a reduced intensity around 250 nm, signifying the alteration of the right-handed helicity. All modified duplexes displayed slight

blue shifts (2–3 nm) relative to the unmodified control duplexes, indicating an impact of the lesion on the DNA backbone [24]. The effect was most prominent in *Nar*I-G_1G_3. Also, the modified duplexes exhibited a change in the positive ellipticity at around 270 nm, indicating alterations of base stacking of B-form DNA. While the *Nar*I-G_1G_2 duplex exhibited a slight increase, the *Nar*I-G_2G_3 and *Nar*I-G_1G_3 duplexes actually displayed a decrease in intensity, with the impact being greater in the former (Figure 3a).

In addition, sequence-dependent $ICD_{290–320}$ patterns were seen in all three diadduct duplexes. Previously, we reported that the FAAF-modified *Nar*I monoadduct duplexes displayed a negative dip, with the *Nar*I-G_2 duplex exhibiting a greater dip than the G_1 or G_3 duplex [13]. Similar $ICD_{290–320}$ was observed in the diadduct duplexes. The *Nar*I-G_1G_3 and *Nar*I-G_2G_3 duplexes showed larger dips than the *Nar*I-G_1G_2 duplex.

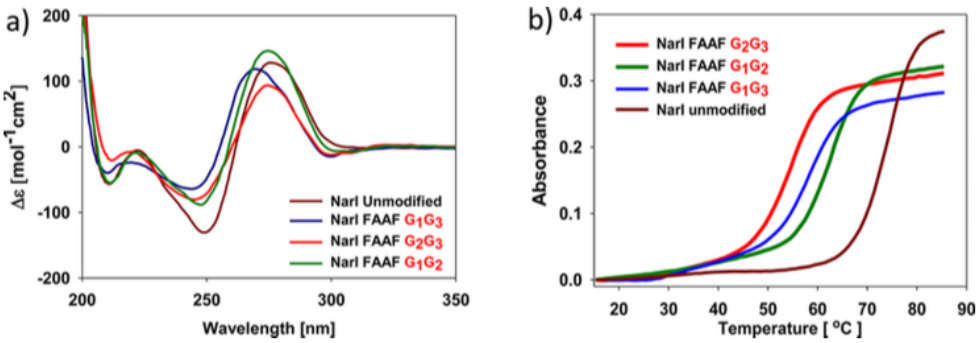


Figure 3. CD and UV melting profiles. (a) CD spectral overlays recorded at 20 °C and (b) UV melting curves in 20 mM phosphate buffer containing 0.1 M NaCl (pH 7) of fully paired 16-mer *Nar*I di-FAAF adduct duplexes: G_1G_2 (green), G_1G_3 (blue), and G_2G_3 (red).

UV Melting Experiments

Figure 3b shows the UV melting profiles of the three di-FAAF duplexes in comparison to the unmodified control at 4.8 μM. All duplexes exhibited typical monophasic, sigmoidal, helix-to-coil transitions with an excellent linear correlation ($R^2 > 0.9$) between T_m^{-1} and ln C_t. It is clear that modified duplexes were thermally and thermodynamically destabilized relative to the controls. Table 1 summarizes the thermal and thermodynamic

parameters calculated from the UV melting experiments. The results were converted into the thermodynamic chart shown in Figure 4. The *Nar*I-G_2G_3 duplex was most destabilized: $\Delta\Delta G_{37°C}$ = 7.8 kcal/mol, ΔΔH = 23.3 kcal/mol, and ΔT_m = −17.9°C. The ΔT_m value of *Nar*I-G_1G_3 was −14.1°C, approximately 4°C lower than that of *Nar*I-G_2G_3, but *Nar*I-G_1G_3 had a slightly higher ΔΔH of 24.1 kcal/mol. However, the large entropy (ΔΔS = 56.0 eu) compensated for the enthalpy (Figure 4 and Table 1), resulting in an overall free energy destabilization ($\Delta\Delta G_{37\,°C}$) of 6.8 kcal/mol. In comparison, the stability of the *Nar*I-G_1G_2 duplex was least affected ($\Delta\Delta G_{37°C}$ = 4.3 kcal/mol, and ΔT_m = −10.0 °C). The weak entropy compensation (ΔΔS = 6.0 eu) for the substantial enthalpy reduction (ΔΔH = 6.2 kcal/mol) resulted in the free energy destabilization ($\Delta\Delta G_{37°C}$ = 4.3 kcal/mol) (Figure 4 and Table 1). Overall, the magnitude of thermal (ΔT_m) and thermodynamic (ΔΔG) destabilization decreased in the following order: *Nar*I-G_2G_3 > *Nar*I-G_1G_3 > *Nar*I-G_1G_2.

Table 1. Thermal and Thermodynamic Parameters of FAAF-Modified *Nar*I Diadduct Duplexes Obtained from UV Melting Experiments

5'-CTCTCG$_1$G$_2$CG$_3$CCATCAC-3'
3'-GAGAGC C GCGGTAGTG-5'

	$-\Delta H$ (kcal/mol)	$-\Delta S$ (eu)	$-\Delta G_{37\,°C}$ (kcal/mol)	T_m[b] (°C)	$\Delta\Delta H$[c] (kcal/mol)	$\Delta\Delta S$[d] (eu)	$\Delta\Delta G_{37\,°C}$[e] (kcal/mol)	ΔT_m[f] (°C)
control[a]	121.9	324.7	21.2	70.6	–	–	–	–
*Nar*I-G_1G_3-FAAF[a]	97.8	268.7	14.4	56.5	24.1	56.0	6.8	−14.1
*Nar*I-G_2G_3-FAAF[a]	98.6	274.6	13.4	52.7	23.3	50.1	7.8	−17.9
*Nar*I-G_1G_2-FAAF[a]	115.7	318.7	16.9	60.6	6.2	6.0	4.3	−10.0

[a]The average standard deviations for $-\Delta G$, $-\Delta H$, and T_m are ±0.4, ±3.0, and ±0.4, respectively. [b]T_m is the temperature at midpoint of the sigmoidal melting curve. [c]$\Delta\Delta H$ = ΔH(modified duplex) − ΔH(control duplex). [d]$\Delta\Delta S$ = ΔS(modified duplex) − ΔS(control duplex). [e]$\Delta\Delta G$ = ΔG(modified duplex) − ΔG(control duplex). [f]ΔT_m = Tm(modified duplex) − T_m(control duplex).

Dynamic ^{19}F NMR Spectroscopy

Figure 5 shows the ^{19}F NMR spectra of diadduct duplexes as a function of temperature (5−65°C) (see Figure S5 of the Supporting Information for complete temperature ranges). As expected, the diadduct duplexes displayed an array of ^{19}F signals in the −114 to −118 ppm range. This is the same chemical shift range in which mono-FAAF

adducts display their B (−115.0 to −115.5 ppm), S (−115.5 to −117.0 ppm), and W (−117.0 to −118.0 ppm) conformers [13,14]. The complexity observed in the ^{19}F NMR spectra of di- *versus* mono-FAAF adducts was expected because of the presence of two FAAF lesions, each of which is capable of exhibiting B-S-W heterogeneity. As a result, it was very difficult to assign signals as we have done for the monoadducts [13].

Nonetheless, all diadducts exhibited typical dynamic NMR characteristics, i.e., exchange broadening of ^{19}F signals with increasing temperatures, reaching an eventual coalescence at ≥60°C, indicating thermal denaturation of a duplex (Figure 5). In addition, overall signal patterns were found to be dependent on flanking sequence. Figure S6a of the Supporting Information compares the ^{19}F NMR spectra of *Nar*I-G_2G_3 with those of the mono-FAAF counterparts (*Nar*I-G_2 and -G_3) at two different temperatures (5 and 20°C).

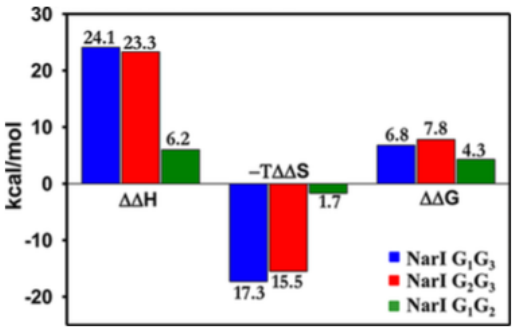


Figure 4. Comparative thermodynamic parameter histogram of *Nar*I di-FAAF adduct duplexes; G_1G_2 (green), G_1G_3 (blue), and G_2G_3 (red). $\Delta\Delta H = \Delta H$(modified duplex) − ΔH(control duplex). $T\Delta\Delta S = T\Delta S$(modified duplex) − $T\Delta S$(control duplex). $\Delta\Delta G = \Delta G$(modified duplex) − ΔG(control duplex).

Similar comparisons were made for *Nar*I-G_1G_3 and *Nar*I-G_1G_2 (Figure S6b, c of the Supporting Information). The first-order approximation was that the signal patterns of the diadduct duplexes would be the sum of their respective monoadducts, but that turned out not to be the case. This indicates that the two nearby FAAF adducts interact with each other, leading to conformational sampling not occurring with the monoadducts. It was noted that signals of *Nar*I-G_1G_3 and -G_2G_3 are broad even at 5°C (Figure 5), indicating some dynamic exchanges among conformers. This observation suggests a conformational flexibility at the lesion site, consistent with their high entropy

values (Figure 4 and Table 1). In contrast, the ^{19}F signals of *Nar*I-G_1G_2 were comparatively sharper, in line with their low entropy (Table 1). In all three duplexes, increasing temperatures resulted in a coalescence of signals into a single signal (approximately −115 ppm) that represents a free modified single strand. The *Nar*I-G_1G_2 duplex showed a higher-magnitude duplex melting signal at ~65°C compared to those at 60°C for the other two duplexes. These results indicate the greater thermodynamic stability of G_1G_2, consistent with the UV melting data given above. The ^{19}F conformational complexity was supported by their imino proton spectra (Figure S7 of the Supporting Information). The *Nar*I-G_1G_3 and *Nar*I-G_1G_2 duplexes exhibited multiple signals in the non-Watson-Crick base pairing 11-12 ppm region, which is known to occur from the imino protons of modified guanines. In *Nar*I-G_2G_3, however, the signals were shifted downfield to 12 ppm.

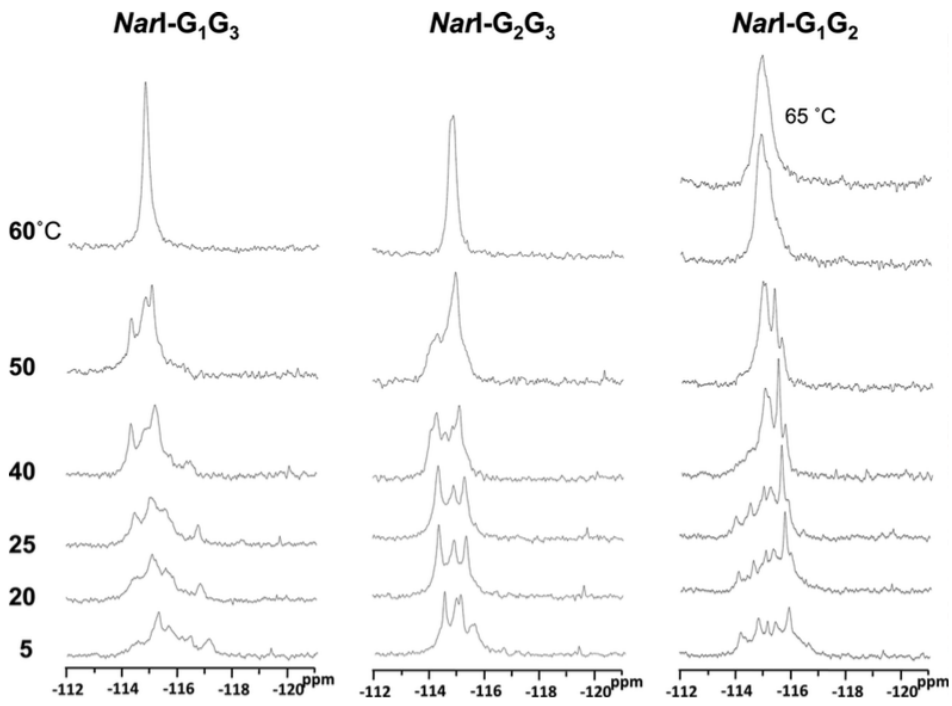


Figure 5. ^{19}F NMR spectra of di-FAAF-modified duplexes at different temperatures.

MD Simulations

To provide molecular explanations for the experimental results described above, we constructed all the FAAF-modified DNA diadducts as well as an unmodified duplex as controls and subjected those models to 100 ns MD simulations in explicit solvent with neutralizing counterions. Each diadduct lesion site was modeled in both B-type and S-type conformers, resulting in 12 possible diadducts. Detailed analyses were performed on the MD trajectories between 20 and 100 ns, during which the simulations were stable as shown by their rms differences as a function of time (Figure S8 of the Supporting Information). However, the rms differences for all the diadducts are significantly larger than that of the unmodified control duplex, and in many of the systems, brief excursions to larger rms differences occur during the simulations. The larger rms differences are associated with an increased level of bending in the diadducts compared to that in the control duplex, and the excursions suggest that a range of conformations are being sampled during the simulations. The sampling of a range of conformations is consistent with the ^{19}F NMR analysis discussed above. Please see the Supporting Information for torsion angles at the nucleotide-adduct linkage site.

MM/GBSA Analysis

As each diadduct has four possible glycosidic isomers (BB, BS, SB, and SS), we undertook MM/GBSA analysis [52] to determine the most highly populated isomer for each diadduct (Figure S9 and Table S1 of the Supporting Information). We note that additional terms may be included in the overall free energy estimate [53], such as configurational entropy; however, those terms have been omitted here, and we note the magnitude of the

standard deviations. We note that given the changes in the chemical structure of the duplexes, the energies between the control duplex and different diadducts cannot be directly compared. For each of the diadducts, one or two of the isomers have the most favorable free energies (Figure S9 of the Supporting Information). These include BB and BS for G_1G_3, BB for G_2G_3, and BB for G_1G_2, though with the latter the BS and SB conformations are within 20 kcal/mol of the minimum while the separation is 25 kcal/mol or more for the G_1G_3 and G_2G_3 diadducts. Notably, none of the SS diadducts was observed to be stable (see below). On the basis of this analysis, we assume that these lower-energy isomers contribute to the experimental results and interpret experimental and computational data on the basis of this assumption.

Some mix of B and S isomers is indicated that could contribute to the conformational heterogeneity observed in the ^{19}F NMR spectra. With G_2G_3, three relatively well-defined peaks are observed in the ^{19}F spectra at 25°C (Figure 5), consistent with the BB isomer dominating. Broader distributions of conformations are observed in the G_1G_3 and G_1G_2 NMR spectra, consistent with presence of one and three low-energy geometric isomers, respectively, being accessible based on the MM/GBSA calculations. In addition, the wide distributions in the G_1G_3 and G_1G_2 spectra are consistent with predicted sampling W conformations in some of the S states (see below).

Table 2. Selected Calculated Data of the MM/GBSA Low-Energy Diadduct Glycosidic Isomers[a]

	ΔT_m (°C)	$\Delta\Delta H$ (kcal/mol)	$-T\Delta\Delta S$	base pairing[b] (%)	base stacking (kcal/mol)	duplex bending (deg)	SASA (Å^2)	hydration energy (kcal/mol)
*Nar*I-WT	–	–	–	99.9	−178.2 ± 1.0	20.5 ± 11.8	4176 ± 34	−4286 ± 22
*Nar*I-G_1G_3-BB	−14.1	24.1	17.3	79.8	−152.4 ± 2.3	37.4 ± 1.9	4718 ± 41	−4605 ± 17
*Nar*I-G_1G_3-BS				83.3	−155.6 ± 3.6	42.6 ± 4.0	4624 ± 35	−4610 ± 30
*Nar*I-G_2G_3-BB	−17.9	23.3	15.5	93.5	−160.0 ± 10.0	41.3 ± 8.3	4591 ± 107	−4500 ± 139
*Nar*I-G_1G_2-BB	−10.0	6.2	1.7	99.7	−170.2 ± 1.5	46.3 ± 5.4	4691 ± 44	−4441 ± 12
*Nar*I-G_1G_2-BS				85.8	−169.1 ± 9.3	38.1 ± 15.1	4699 ± 56	−4573 ± 17
*Nar*I-G_1G_2-SB				83.9	−155.4 ± 6.8	34.3 ± 5.5	4661 ± 91	−4397 ± 49

[a] ΔT, $\Delta\Delta H$, and $\Delta\Delta S$ experimental results from Table 1 and Figure 4, with the remaining results from analysis of the MD simulations. The two terminal nucleotides at each end are excluded from the calculations. [b] The percent WC base pairing was calculated by applying a cutoff of 3.5 Å for the N1−N3 distance for all base pairs excluding the two terminal nucleotides at each end of the duplexes.

Clustering Analysis

Clustering analysis was undertaken to identify representative conformations of the different diadducts. Clustering was based on the rms differences of the six central nucleotides, with the results listed in Table S2 of the Supporting Information. In G_1G_3, the BB and BS states are dominated by one conformation, with significant sampling of a second conformation in both cases. With G_2G_3, a single conformation dominates the low-energy BB state. A similar situation occurs with the BB and BS isomers of G_1G_2, though three highly populated conformations occur with SB. Overall, these results are consistent with the ^{19}F NMR results (Figure 5); a range of conformations are present with G_1G_3 and G_1G_2, while a more limited number of conformations occur with G_2G_3. It should be noted that additional conformations not sampled in the simulations may contribute to the heterogeneity of the ^{19}F NMR spectra.

Representative structures from the most populous clusters are shown in Figure 6. For G_1G_3, there were two highly populated clusters for both BB and BS, with the four structures shown. Consistent with the *anti* versus *syn* linkages, three of the FAAF adducts are exposed to solvent, while in the BS 240 structure, the *syn* adduct is stacked in the duplex. Comparison of the BB structures shows the 955 structure to have the lower adduct partially stacked against a C base, an interaction that would be expected to impact the ^{19}F chemical shift. The two G_1G_3 BS structures are similar, with the exception of the C base flipping out of the duplex, thereby allowing for an increased level of hydration of the interior of the duplex, a phenomenon that could also contribute to the heterogeneity of the ^{19}F NMR spectra (Figure 5). With G_2G_3, the BB isomer was indicated to be the dominant species and a single conformation was sampled the majority of the duration of the

simulation. The structure in Figure 6 shows the Watson-Crick (WC) pairing to be well maintained, with the exception of the base pair between the two modified G residues. These results are consistent with the G_2G_3 ^{19}F NMR spectra being the most well-defined, although the presence of three peaks in that spectrum indicates that additional conformations are present that are not being sampled in these simulations. The ^{19}F NMR spectra of G_1G_2 (Figure 5) were quite complicated, which is consistent with the BB, BS, and SB glycosidic isomers indicated to be stable based on the MM/GBSA results (Table S2 of the Supporting Information). In the BB isomer, the WC pairing is maintained, but the duplex is distorted because of the FAAF adducts being adjacent to each other and, consequently, interacting directly with each other. The overall duplex is also well maintained on the BS

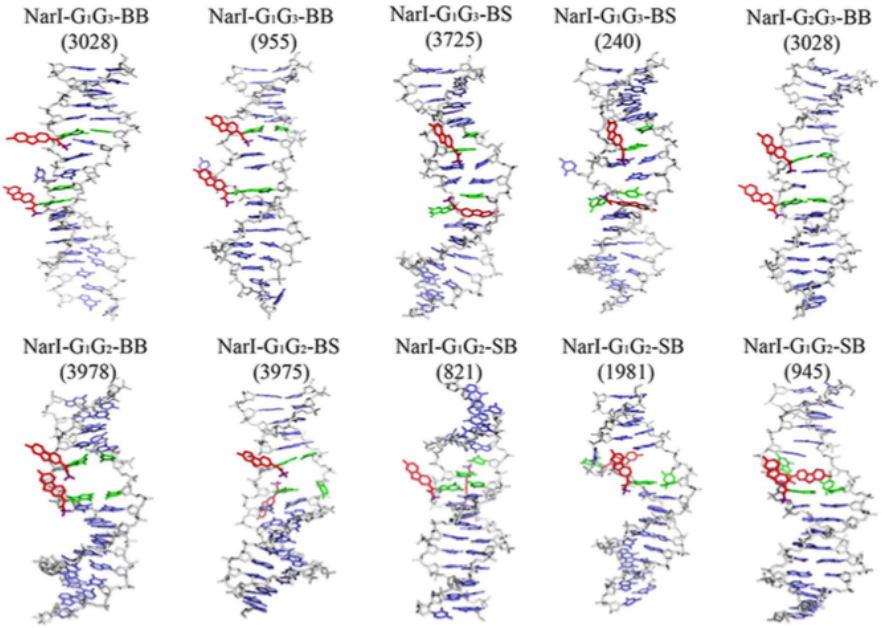


Figure 6. Representative conformations of the di-FAAF duplexes for the low-energy glycosidic isomers obtained from the large clusters based on rms differences of the six central nucleotides (B, B-type; S, stacked). FAAF adducts are colored red, and DNA bases are colored light blue except for G_1, G_2, G_3, and their complementary cytosine bases, which are colored green. Backbone and sugar rings are colored gray. The numbers of frames for each cluster are in parentheses as listed in Table S2 of the Supporting Information, which indicates the population of each cluster divided by the total number frames for each isomer (4000 in this study).

isomer, which appears to be associated with the *syn* FAAF adduct sampling a W-type conformation where it is interacting with the minor groove rather than being stacked in the duplex. Similarly, a W-type conformation of the *syn* FAAF adduct is observed in the 821 and

1981 conformation clusters of the SB isomer, while the FAAF adduct is stacked in the 945 conformation cluster. It should be noted that during the SB simulation, the S conformer was present at the beginning of the simulation, consistent with the initial model, yielding the 945 cluster, with the *syn* FAAF adduct subsequently converting to the W conformer that was maintained throughout the remainder of the simulation. This suggests that the W conformers are representative of the SB isomer of G_1G_2. Overall, the simulations indicate that the different diadducts sample different ranges of conformations, an observation that is consistent with the ^{19}F NMR data. However, especially in the case of G^2G^3, it is clear that the simulations are not sampling all the relevant conformations. This may be an outcome with limitations in the estimated stabilities from the MM/GBSA analysis as well as the limited simulation time in this study.

Base Pairing

Additional analysis focused on possible contributions to the observed stabilities of the diadducts, including base pairing, base stacking, bending, and solvation effects for the low-energy glycosidic isomers (Table 2). Table 2 also contains the relative T_m, enthalpy, and entropy values from the data in Table 1. For the unmodified control duplex, its WC hydrogen bonding is almost fully maintained, while with WC pairing is lost to varying degrees in the diadducts. This suggests that a loss of base pairing in the diadducts contributes to their lowered stability. However, as the least stable diadduct, G_2G_3, maintains a significant amount of WC interactions versus the more stable G_1G_3 diadducts, it appears that this term does not dominate the relative stabilities of the diadducts.

Base Stacking. Base stacking typically dominates the thermodynamic stability of DNA, and changes in adduct stability in deletion duplexes have been shown to have significant contributions from perturbations of base stacking [54]. Accordingly, we examined the relationship of base stacking energies to the measured thermodynamic enthalpies of the diadducts. Compared to the control, all diadducts have a significant loss of base stacking energy. The loss is smallest in the most stable G_1G_2 diadducts; however, the least stable adduct, G_2G_3, maintains base stacking to a larger degree than the relatively more stable G_1G_3 adduct. With relative experimental enthalpies of 24.1, 23.3, and 6.2 kcal/mol for G_1G_3, G_2G_3, and G_1G_2, respectively, it is evident that the calculated results are consistent with the relatively small loss in calculated stacking with G_1G_2 while the losses are significantly greater with G_1G_3 and G_2G_3. Thus, base stacking plays an important role in the relative stabilities of the diadducts.

We note that in all the SS glycosidic isomers the greatest loss of base pairing and base stacking occurred with one exception (Tables S3 and S4 of the Supporting Information). This is consistent with the unfavorable MM/GBSA energies of these species. This results suggest that the duplex is not able to accommodate the two *syn* adducts simultaneously stacking, leading to SS isomers not being stable. In the presence of two FAAF-modified G bases in the *syn* orientations, both of the lesions would be right next to each other in the same strand, leading to a decrease in stacking energy that is much larger than in the other diadduct geometric isomers. Thus, we consider it unlikely that this conformer has a noticeable population in the studied diadducts.

Bending and Solvation

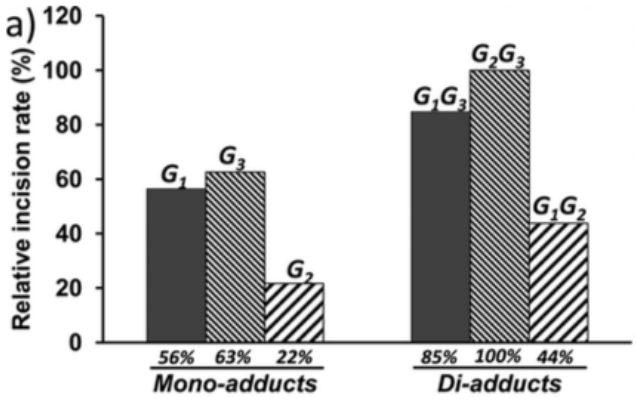
DNA bending analyses of the duplexes are presented in Table 2. Compared to that in the unmodified duplex, the presence of two FAAF adducts induces significant bending in all the diadducts, leading to the significant rms differences noted above. The extent of bending is evident in the *Nar*I-G_1G_2-BS (3975) and *Nar*I-G_1G_2-SB (1981) structures shown in Figure 6 and is consistent with the decrease in the extent of base pairing and stacking in the adducts. In addition, the increased level of bending is consistent with the more favorable entropy of the diadducts, which partially compensates for the enthalpy loss, especially with G_1G_3 and G_2G_3.

Solvation can also impact both enthalpic and entropic aspects of duplex stability. To analyze the possible contribution of the effects of solvation on the bulge regions, we computed the solvent accessible surface area (SASA) and the hydration energies from the simulations (Table 2). The increase in the SASA values in the diadducts over the control duplex is significant, indicating that solvation effects are making a large contribution to the destabilization of the duplexes in the diadducts. The increased SASA is accompanied by a more favorable hydration energy in the diadducts, which would lead to a more favorable enthalpy, in contrast to the experimental data. This suggests that the loss of base pairing and stacking in the diadducts is dominating the change in enthalpy, with solvation offsetting those unfavorable contributions. Interestingly, the MM/GBSA contribution to the free energy of solvation is the most favorable in the control (Figure S9 and Table S1 of the Supporting Information), indicating that the entropy component plays a significant role in the change in solvation. While qualitative in nature, these results suggest a model in which the diadducts lead to a more favorable solvation enthalpy, but this is overcome by a larger

unfavorable entropy contribution leading to the solvation contribution contributing to destabilization of the diadduct duplexes. With regard to the different diadducts, the solvation properties do not appear to make a significant contribution to the relative stabilities.

UvrABC Incisions of *Nar*I FAAF Diadducts

The NER kinetic assay results are shown in Figure 7, in which the 55-mer di-FAAF adduct duplex substrates were incised by UvrABC nuclease. These substrates were radioactively labeled at the 5'- end of the adducted strand. The major incision products can be seen as ~18- and 19-mer fragments separated on a urea-polyacrylamide gel electrophoresis gel under denaturing conditions (Figure S10 of the Supporting Information). The incision occurred at the eighth phosphate bond 5' to the modified nucleotide, which is consistent with the previously reported results of UvrABC incision [49,50]. Quantitative analysis of the incision products suggests that the diadduct substrates were incised 1.5-2 times more efficiently than the previously reported monoadducts (Figure 7) [13]. Moreover, the substrates were incised at different efficiencies, depending upon where the two damage sites were located in the sequence (Figure 7). Specifically, di-FAAF-adducts in *Nar*I-G_2G_3 (100%) displayed the fastest rate of incision, followed by *Nar*I-G_1G_3 (85%), whereas *Nar*I-G_1G_2 (44%) had a much lower rate of incision [G_2G_3 (100%) > G_1G_3 (85%) > G_1G_2 (44%) (Figure 7)].



b)

Duplex	*Nar*I-G_1G_3	*Nar*I-G_2G_3	*Nar*I-G_1G_2
Incision rate	0.344±0.013	0.406±0.020	0.178±0.008

Incision rate (fmol/min)

Figure 7. UvrABC nuclease incision kinetic results. (a) Comparative relative incision rate histogram of FAAF-modified mono- and diadducts located at different positions calculated by considering the *Nar*I-G_2G_3 FAAF incision rate as 100%. (b) Absolute incision rates of three diadduct duplexes.

Discussion

A combined experimental and computational study of selected di-FAAF-adducts of the *Nar*I sequence was undertaken. Results show the diadduct duplexes to be destabilized to a greater extent than the monoadducts. In addition, they sample a wider range of conformations as indicated by ^{19}F NMR spectra, with those conformations maintaining significant B-form DNA characteristics as indicated by CD and MD analyses. Notably, the diadducts lead to enhanced incision rates, as discussed below.

Impact of Sequence Context on Diadduct Duplex Structure and Thermodynamics

The CD results indicate that the overall DNA structure was influenced by how the two bulky FAAF lesions are arranged in the *Nar*I sequence (Figure 3a). In particular, base stacking represented by changes in the positive ellipiticity at 270 nm was dependent on flanking sequence. While a significant decrease was seen in *Nar*I-G_2G_3 and *Nar*I-G_1G_3, an increase in intensity was noted for *Nar*I- G_1G_2. Our MD simulation data indicate the perturbations of base stacking as a major contributor to the observed sequence effect. In line with CD, the base stacking was maintained to a larger extent in *Nar*I-G_1G_2, whereas severe disruption of base stacking in *Nar*I-G_1G_3 and -G_2G_3 resulted in a significant loss of duplex stability (Table 2 and Table S4 of the Supporting Information). However, in contrast to the impact of CD, the major impact on base stacking was observed in *Nar*I-G_1G_3. This difference could be due to the MD simulations not sampling all the relevant conformations contributing to the experimental results, though complications in the interpretation of the CD spectra cannot be excluded. A similar sequence effect was also observed in the thermal and thermodynamic instability (Figure 4 and Table 1): *Nar*I-G_2G_3 > *Nar*I-G_1G_3 > *N*arI-G_1G_2. This trend is not surprising because base stacking interactions are one of the most important factors that contribute to the thermal and thermodynamic stability of the DNA duplexes [16,17].

Assuming that each lesion can exist in two major conformations about the glycosidic linkage [*anti* (B) and *syn* (S)], there are four possible combinations of geometric isomers, BB, BS, SB, and SS. All four were modeled and simulated. The rms difference analysis (Figure S8 of the Supporting Information) indicates that the diadduct structures are perturbed to a greater extent than the unmodified control duplex, though all are reasonably

stable. The larger rms differences are associated with an increased level of bending in the diadducts compared to the nonadducted control because of a range of conformations being sampled during the simulations. Consequently, the ^{19}F NMR spectra of the diadduct duplexes produced very complex signal patterns (Figure 5). Interestingly, MM/GBSA analysis indicates the geometric isomers containing B conformers, BB and BS for *Nar*I-G_1G_3, BB for *Nar*I-G_2G_3, and BB, BS, and SB for *Nar*I-G_1G_2, are favored, with the SS geometric isomer being unfavorable in all cases. Cluster analysis of the conformations from the simulations of the favored conformations indicates each of the isomers samples a range of conformations, contributing to the heterogeneity of the ^{19}F NMR spectra.

We reported previously that the single dG-FAAF adduct at the G_3 position of the *Nar*I sequence exhibits a preference for the S conformation (61%) (Table S5 of the Supporting Information) [13]. Similar S conformer dominance was found for dG-AF [16] and related dG-C8 adducts derived from the heterocyclic amines 2-amino-3-methylimidazo(4,5-f)quinolone (IQ) [55] and 2-amino-1-methyl-6-phenylimidazo(4,5-b)pyridine(PHIP) [56]. Interestingly, in the study presented here, greater thermodynamic instability was found for the G_3-containing *Nar*I-G_2G_3 and *Nar*I-G_1G_3 duplexes that produce a higher combined S population (~76 and ~95%, respectively) than *Nar*I-G_1G_2 (~49%) (Table S5 of the Supporting Information). It is possible that the greater proximity of two FAAF motifs in *Nar*I-G_2G_3 (e.g., just one base apart) compared to that in *Nar*I-G_1G_3 (two bases) possibly could induce a greater DNA distortion. It should be noted, however, that the most thermodynamically stable *Nar*I-G_1G_2 has no base between the lesions. Thus, it appears that assessing the overall thermodynamic stability of diadducts is complicated by various conformational factors.

Conformation-Specific Nucleotide Excision Repair

The *E. coli* UvrABC system exhibited approximately 1.5-2.0-fold better efficiency in the reparability of the diadducts (G_2G_3, G_1G_3, and G_1G_2) than in the reparability of the monoadducts (G_1, G_2, and G_3) in the same *Nar*I sequence (Figure 7). This is not surprising because of the greater thermal and thermodynamic destabilization of the di-FAAF (Table 1) versus mono-FAAF duplexes [13]. It is well-known that NER efficiency is highly modulated by the damage recognition ($UvrA_2$) and verification and recognition (UvrB) steps [57]. Jaciuk et al. have shown that $UvrA_2$ does not interact with the DNA lesion directly but senses the structural and conformational perturbations induced by a lack of Watson-Crick hydrogen bonding and thermodynamic destabilization [20]. Moreover, the thermodynamic destabilization of the duplex assists UvrB in the insertion of a β-hairpin into the duplex, which is required for lesion verification and recognition [21]. We reported that the repair of mono dG-FAAF adducts in the *Nar*I sequence occurs in a conformation-specific manner; i.e., the highly S/W-conformeric G_3 and G_1 duplexes produced greater DNA distortion and thermodynamic destabilization and thus incised more efficiently than the B- type G_2 duplex ($G_3 \sim G_1 > G_2$) [13]. The presence of two dG-FAAF motifs in the same *Nar*I sequence is expected to synergize the destabilization effect because of the significant loss of base pairing and base destacking as observed in the MD simulations (Table 2). Indeed, the destabilization effect of diadducts ($\Delta\Delta G_{37\,^\circ C}$ = 4.3−7.8 kcal/mol, and ΔT_m = −10.0 to −17.9°C) is 1.5−2.0-fold greater than that of the monoadducts ($\Delta\Delta G_{37\,^\circ C}$ = 3.7−4.7 kcal/mol, and ΔT_m = −5.3 to −8.3°C) (Table 1) [13]. These results are in contrast to recent reports on oxidative DNA damage where the reparability is compromised because of clustering, leading to enhanced genotoxicity [6]. The rationale behind this discrepancy could lie in the complexity

of processing of oxidative cluster damage sites that involve base excision repair, nonhomologous end joining, and homologous recombination proteins.

In addition, we observed different repair efficiencies depending on the location of di-FAAF adducts (G_2G_3, G_1G_3, and G_1G_2). The *Nar*I-G_1G_2 duplex showed a considerably lower efficiency than *Nar*I-G_2G_3 and -G_1G_3 [G_2G_3 (100%) > G_1G_3 (85%) > G_1G_2 (44%)] (Figure 7). These NER results are in good agreement with the extent of the loss of base stacking and thermal and thermodynamic destabilization (G_2G_3 > G_1G_3 > G_1G_2). The *Nar*I-G_2G_3 duplex displayed the highest repair rate and also underwent the most severe loss of base stacking (Figure 3a and Table 2) and duplex destabilization (for *Nar*I- G_2G_3, $\Delta\Delta G_{37\,^\circ C}$ = 7.8 kcal/mol and ΔT_m = −17.9°C). By contrast, the *Nar*I-G_1G_2 duplex showed the weakest reparability, consistent with the minimal base stacking (Figure 3a and Table 2) and duplex destabilization (for *Nar*I-G_1G_2, $\Delta\Delta G_{37\,^\circ C}$ = 4.3 kcal/mol and ΔT_m = −10°C). *Nar*I-G_1G_3 showed an intermediary reparability and loss of stacking (Figure 3a and Table 2) and destabilization (for *Nar*I-G_2G_3, $\Delta\Delta G_{37\,^\circ C}$ = 6.8 kcal/mol and ΔT_m = −14.1°C). These results confirm the sequence dependence repair of the diadduct duplexes upon duplex destabilization.

It should be noted that the extent of repair of AAF involving position G_3 is consistently greater, which suggested the uniqueness of the G_3 position of the *Nar*I sequence in cases of both monoadducts [13] and diadducts. Interestingly, comparison of the incision products indicated that *Nar*I-G_1G_2/*Nar*I- G_1G_3 and *Nar*I-G_2G_3 are incised at different sites in the sequence (Figure S10b of the Supporting Information). Specifically, *Nar*I-G_1G_2 and *Nar*I-G_1G_3 are incised to generate a major 18-mer product, while G_2G_3 is incised to produce a 19-mer. These interesting results suggest that the incision position of

the diadducts depends on the position of the 5'-adduct but is independent of that of the 3'-adduct.

In summary, NER of di-dG-FAAF adducts was found to occur approximately 2-fold more efficiently than NER of the corresponding monoadducts in the *Nar*I mutational hot spot sequence. The greater reparability of the diadducts was primarily due to their greater thermal and thermodynamic destabilization. In addition, the incision efficiency occurred in a sequence-specific manner; i.e., highly destacked and destabilized *Nar*I-G_2G_3 and *Nar*I-G_1G_3 are repaired more efficiently than the relatively most stable *Nar*I-G_1G_2. These results indicate the importance of carcinogen-induced destacking and related thermal and thermodynamic destabilization in the repair of cluster arylamine adducts.

References

(1) Boldt, J., Mah, M. C., Wang, Y. C., Smith, B. A., Beland, F. A., Maher, V. M., and McCormick, J. J. (1991) Kinds of mutations found when a shuttle vector containing adducts of 1,6-dinitropyrene replicates in human cells. *Carcinogenesis 12*, 119–126.

(2) Burnouf, D., Koehl, P., and Fuchs, R. P. (1989) Single adduct mutagenesis: strong effect of the position of a single acetylaminofluorene adduct within a mutation hot spot. *Proc Natl Acad Sci USA 86*, 4147–4151.

(3) Ward, J. F. (1988) DNA Damage Produced by Ionizing Radiation in Mammalian Cells: Identities, Mechanisms of Formation, and Reparability, in, pp 95–125. Elsevier.

(4) Kalam, M. A., and Basu, A. K. (2005) Mutagenesis of 8-oxoguanine adjacent to an abasic site in simian kidney cells: tandem mutations and enhancement of G-->T transversions. *Chem. Res. Toxicol. 18*, 1187–1192.

(5) Shikazono, N., Pearson, C., O'Neill, P., and Thacker, J. (2006) The roles of specific glycosylases in determining the mutagenic consequences of clustered DNA base damage. *Nucleic Acids Res 34*, 3722–3730.

(6) Magnander, K., and Elmroth, K. (2012) Biological consequences of formation and repair of complex DNA damage. *Cancer Lett. 327*, 90–96.

(7) Fuchs, R. P. P. (1984) DNA binding spectrum of the carcinogen N-acetoxy-N-2-acetylaminofluorene significantly differs from the mutation spectrum. *J Mol Biol 177*, 173–180.

(8) Roy, A., and Fuchs, R. P. P. (1994) Mutational spectrum induced in Saccharomyces cerevisiae by the carcinogen N-2-acetylaminofluorene. *Mol. Gen. Genet. 245*, 69–77.

(9) Mah, M. C., Boldt, J., Culp, S. J., Maher, V. M., and McCormick, J. J. (1991) Replication of acetylaminofluorene-adducted plasmids in human cells: spectrum of base substitutions and evidence of excision repair. *Proc Natl Acad Sci USA 88*, 10193–10197.

(10) Ross, M. K., Said, B., and Shank, R. C. (2000) DNA-damaging effects of genotoxins in mixture: modulation of covalent binding to DNA. *Toxicol. Sci. 53*, 224–236.

(11) Luch, A. (2005) Nature and nurture - lessons from chemical carcinogenesis. *Nat Rev Cancer 5*, 113–125.

(12) Heflich, R. H., and Neft, R. E. (1994) Genetic toxicity of 2-acetylaminofluorene, 2-aminofluorene and some of their metabolites and model metabolites. *Mutation Research/Reviews in Genetic Toxicology 318*, 73–174.

(13) Jain, V., Hilton, B., Patnaik, S., Zou, Y., Chiarelli, M. P., and Cho, B. P. (2012) Conformational and thermodynamic properties modulate the nucleotide excision repair of 2-aminofluorene and 2-acetylaminofluorene dG adducts in the NarI sequence. *Nucleic Acids Res 40*, 3939–3951.

(14) Patnaik, S., and Cho, B. P. (2010) Structures of 2-acetylaminofluorene modified DNA revisited: insight into conformational heterogeneity. *Chem. Res. Toxicol. 23*, 1650–1652.

(15) Cho, B. P. (2004) Dynamic conformational heterogeneities of carcinogen-DNA adducts and their mutagenic relevance. *Journal of Environmental Science and Health*.

(16) Jain, N., Li, Y., Zhang, L., Meneni, S. R., and Cho, B. P. (2007) Probing the sequence effects on NarI-induced -2 frameshift mutagenesis by dynamic 19F NMR, UV, and CD spectroscopy. *Biochemistry 46*, 13310–13321.

(17) Meneni, S. R. (2006) Sequence effects of aminofluorene-modified DNA duplexes: thermodynamic and circular dichroism properties. *Nucleic Acids Res 34*, 755–763.

(18) Meneni, S. R., Shell, S. M., Gao, L., Jurecka, P., Lee, W., Sponer, J., Zou, Y., Chiarelli, M. P., and Cho, B. P. (2007) Spectroscopic and theoretical insights into sequence effects of aminofluorene-induced conformational heterogeneity and nucleotide excision repair. *Biochemistry 46*, 11263–11278.

(19) Geacintov, N. E., Broyde, S., Buterin, T., Naegeli, H., Wu, M., Yan, S., and Patel, D. J. (2002) Thermodynamic and structural factors in the removal of bulky DNA adducts by the

nucleotide excision repair machinery. *Biopolymers 65*, 202–210.

(20) Jaciuk, M., Nowak, E., Skowronek, K., Tańska, A., and Nowotny, M. (2011) Structure of UvrA nucleotide excision repair protein in complex with modified DNA. *Nature Publishing Group 18*, 191–197.

(21) Liu, Y., Reeves, D., Kropachev, K., Cai, Y., Ding, S., Kolbanovskiy, M., Kolbanovskiy, A., Bolton, J. L., Broyde, S., Van Houten, B., and Geacintov, N. E. (2011) Probing for DNA damage with β-hairpins: Similarities in incision efficiencies of bulky DNA adducts by prokaryotic and human nucleotide excision repair systems in vitro. *DNA Repair (Amst) 10*, 684–696.

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

(23) Mu, H., Kropachev, K., Wang, L., Zhang, L., Kolbanovskiy, A., Kolbanovskiy, M., Geacintov, N. E., and Broyde, S. (2012) Nucleotide excision repair of 2-acetylaminofluorene- and 2-aminofluorene-(C8)-guanine adducts: molecular dynamics simulations elucidate how lesion structure and base sequence context impact repair efficiencies. *Nucleic Acids Res 40*, 9675–9690.

(24) Jain, V., Hilton, B., Lin, B., Patnaik, S., Liang, F., Darian, E., Zou, Y., Mackerell, A. D., and Cho, B. P. (2013) Unusual sequence effects on nucleotide excision repair of arylamine lesions: DNA bending/distortion as a primary recognition factor. *Nucleic Acids Res 41*, 869–880.

(25) Mekhovich, O., Tang, M. S., and Romano, L. J. (1998) Rate of incision of N-acetyl-2-aminofluorene and N-2-aminofluorene adducts by UvrABC nuclease is adduct- and sequence-specific: comparison of the rates of UvrABC nuclease incision and protein-DNA complex formation. *Biochemistry 37*, 571–579.

(26) Zou, Y., Shell, S. M., Utzat, C. D., Luo, C., Yang, Z., Geacintov, N. E., and Basu, A. K. (2003) Effects of DNA adduct structure and sequence context on strand opening of repair intermediates and incision by UvrABC nuclease. *Biochemistry 42*, 12654–12661.

(27) Basu, A. K., and Essigmann, J. M. (1988) Site-specifically modified oligodeoxynucleotides as probes for the structural and biological effects of DNA-damaging agents. *Chem. Res. Toxicol. 1*, 1–18.

(28) Guengerich, F. P. (2006) Interactions of Carcinogen-Bound DNA with Individual DNA Polymerases. *Chem. Rev. 106*, 420–452.

(29) Zhou, L., Rajabzadeh, M., Traficante, D. D., and Cho, B. P. (1997) Conformational Heterogeneity of Arylamine-Modified DNA: 19F NMR Evidence. *J. Am. Chem. Soc. 119*, 5384–5389.

(30) Cho, B. P., and Zhou, L. (1999) Probing the conformational heterogeneity of the

acetylaminofluorene-modified 2'-deoxyguanosine and DNA by 19F NMR spectroscopy. *Biochemistry 38*, 7572–7583.

(31) Gao, L., Zhang, L., Cho, B. P., and Chiarelli, M. P. (2008) Sequence verification of oligonucleotides containing multiple arylamine modifications by enzymatic digestion and liquid chromatography mass spectrometry (LC/MS). *Journal of the American Society for Mass Spectrometry 19*, 1147–1155.

(32) Zhang, L.-K., and Gross, M. L. (2000) Matrix-assisted laser desorption/ionization mass spectrometry methods for oligodeoxynucleotides: Improvements in matrix, detection limits, quantification, and sequencing. *Journal of the American Society for Mass Spectrometry 11*, 854–865.

(33) Jain, N., Meneni, S., Jain, V., and Cho, B. P. (2009) Influence of flanking sequence context on the conformational flexibility of aminofluorene-modified dG adduct in dA mismatch DNA duplexes. *Nucleic Acids Res 37*, 1628–1637.

(34) Lu, X.-J., and Olson, W. K. (2008) 3DNA: a versatile, integrated software system for the analysis, rebuilding and visualization of three-dimensional nucleic-acid structures. *Nat Protoc 3*, 1213–1227.

(35) Brooks, B. R., Brooks, C. L., MacKerell, A. D., Nilsson, L., Petrella, R. J., Roux, B., Won, Y., Archontis, G., Bartels, C., Boresch, S., Caflisch, A., Caves, L., Cui, Q., Dinner, A. R., Feig, M., Fischer, S., Gao, J., Hodoscek, M., Im, W., Kuczera, K., Lazaridis, T., Ma, J., Ovchinnikov, V., Paci, E., Pastor, R. W., Post, C. B., Pu, J. Z., Schaefer, M., Tidor, B., Venable, R. M., Woodcock, H. L., Wu, X., Yang, W., York, D. M., and Karplus, M. (2009) CHARMM: the biomolecular simulation program. *J Comput Chem* (Brooks, C. L., III, and Case, D. A., Eds.) *30*, 1545–1614.

(36) Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., and Klein, M. L. (1983) Comparison of simple potential functions for simulating liquid water. *The Journal of Chemical Physics 79*, 926–935.

(37) Hess, B., Kutzner, C., van der Spoel, D., and Lindahl, E. (2008) GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation. *Journal of Chemical Theory and Computation 4*, 435–447.

(38) Hart, K., Foloppe, N., Baker, C. M., Denning, E. J., Nilsson, L., and MacKerell, A. D., Jr. (2012) Optimization of the CHARMM Additive Force Field for DNA: Improved Treatment of the BI/BII Conformational Equilibrium. *Journal of Chemical Theory and Computation 8*, 348–362.

(39) Darden, T., York, D., and Pedersen, L. (1993) Particle mesh Ewald: An N·log(N) method for Ewald sums in large systems. *The Journal of Chemical Physics 98*, 10089–10092.

(40) Bussi, G., Donadio, D., and Parrinello, M. (2007) Canonical sampling through velocity rescaling. *The Journal of Chemical Physics 126*, 014101.

(41) Bussi, G., Zykova-Timan, T., and Parrinello, M. (2009) Isothermal-isobaric molecular dynamics using stochastic velocity rescaling. *The Journal of Chemical Physics 130*, 074101.

(42) Parrinello, M., and Rahman, A. (1980) Crystal Structure and Pair Potentials: A Molecular-Dynamics Study. *Physical Review Letters 45*, 1196–1199.

(43) Hess, B. (2008) P-LINCS: A Parallel Linear Constraint Solver for Molecular Simulation. *Journal of Chemical Theory and Computation 4*, 116–122.

(44) Hess, B., Bekker, H., Berendsen, H. J. C., and Fraaije, J. G. E. M. (1997) LINCS: A linear constraint solver for molecular simulations. *J Comput Chem 18*, 1463–1472.

(45) Lavery, R., Moakher, M., Maddocks, J. H., Petkeviciute, D., and Zakrzewska, K. (2009) Conformational analysis of nucleic acids revisited: Curves. *Nucleic Acids Res 37*, 5917–5929.

(46) Sibson, R. (1973) SLINK: An optimally efficient algorithm for the single-link cluster method. *The Computer Journal 16*, 30–34.

(47) Humphrey, W., Dalke, A., and Schulten, K. (1996) VMD: Visual molecular dynamics. *Journal of Molecular Graphics 14*, 33–38.

(48) (2010) *The PyMOL Molecular Graphics System*. Schrodinger, LLC.

(49) Luo, C., Krishnasamy, R., Basu, A. K., and Zou, Y. (2000) Recognition and incision of site-specifically modified C8 guanine adducts formed by 2-aminofluorene, N-acetyl-2-aminofluorene and 1-nitropyrene by UvrABC nuclease. *Nucleic Acids Res 28*, 3719–3724.

(50) Zou, Y., Liu, T. M., Geacintov, N. E., and Van Houten, B. (1995) Interaction of the UvrABC nuclease system with a DNA duplex containing a single stereoisomer of dG-(+)- or dG-(-)-anti-BPDE. *Biochemistry 34*, 13582–13593.

(51) Zou, Y., and Van Houten, B. (1999) Strand opening by the UvrA(2)B complex allows dynamic recognition of DNA damage. *The EMBO Journal 18*, 4889–4901.

(52) Feig, M., Onufriev, A., Lee, M. S., Im, W., Case, D. A., and Brooks, C. L. (2004) Performance comparison of generalized born and Poisson methods in the calculation of electrostatic solvation energies for protein structures. *J Comput Chem 25*, 265–284.

(53) Beveridge, D. L., and DiCapua, F. M. (1989) Free energy via molecular simulation: applications to chemical and biomolecular systems. *Annu Rev Biophys Biophys Chem 18*, 431–492.

(54) Yakovchuk, P., Protozanova, E., and Frank-Kamenetskii, M. D. (2006) Base-stacking and base-pairing contributions into thermal stability of the DNA double helix. *Nucleic Acids Res 34*, 564–574.

(55) Elmquist, C. E., Wang, F., Stover, J. S., Stone, M. P., and Rizzo, C. J. (2007) Conformational differences of the C8-deoxyguanosine adduct of 2-amino-3-methylimidazo[4,5-f]quinoline (IQ) within the NarI recognition sequence. *Chem. Res. Toxicol. 20*, 445–454.

(56) Brown, K., Hingerty, B. E., Guenther, E. A., Krishnan, V. V., Broyde, S., Turteltaub, K. W., and Cosman, M. (2001) Solution structure of the 2-amino-1- methyl-6-phenylimidazo[4,5-b]pyridine C8-deoxyguanosine adduct in duplex DNA. *Proc Natl Acad Sci USA 98*, 8507–8512.

(57) Truglio, J. J., Croteau, D. L., Van Houten, B., and Kisker, C. (2006) Prokaryotic nucleotide excision repair: the UvrABC system. *Chem. Rev. 106*, 233–252.

Supplemental Data

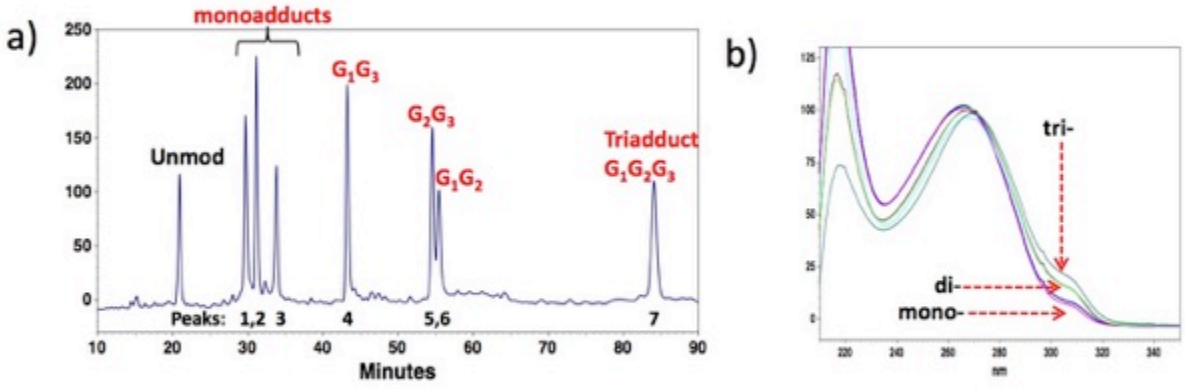


Figure S1: HPLC profile of FAAF-modified ODN. (a) Chromatogram of a reaction mixture between 16-mer *NarI* sequence (5'-CTCTCG$_1$G$_2$CG$_3$CCATCAC-3') and an activated FAAF (N-acetoxy-N-2-(acetylamino)-7-fluorofluorene). The mono- (G_1, G_3, G_2), di- (G_1G_3, G_2G_3 and G_1G_2), and tri-FAAF ($G_1G_2G_3$) adducts eluted in the 28-35, 42-60, and 84 min ranges were purified by reversed-phase HPLC. (b) on-line photodiode array UV/Vis spectra of mono-, di- and tri-FAAF adducts.

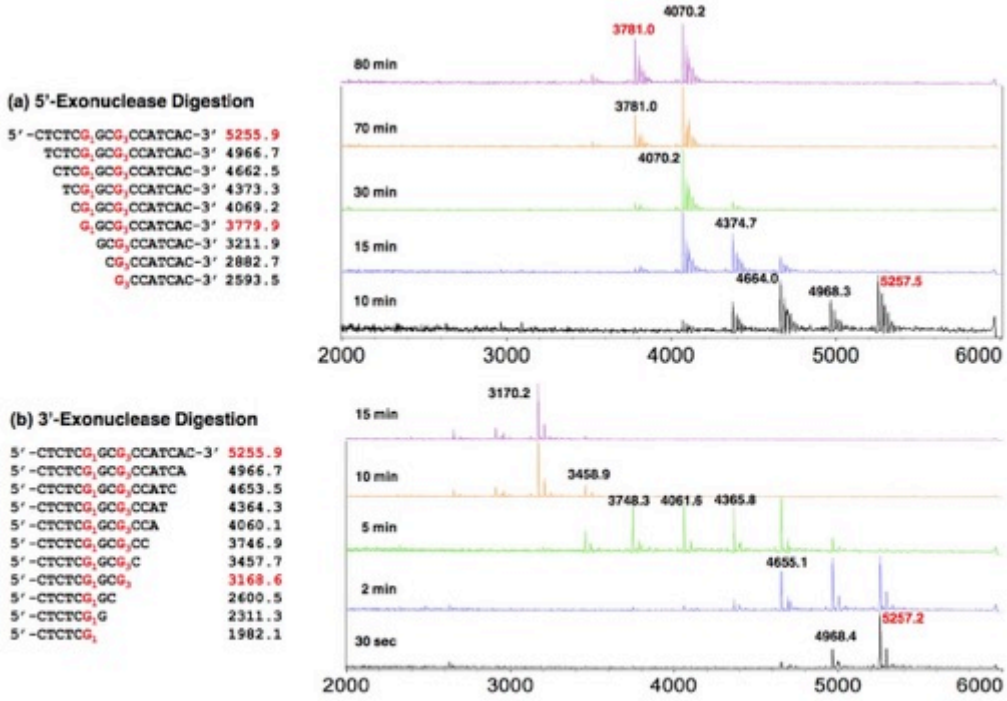


Figure S2: MALDI-TOF mass spectra of the hydrolysis fragments formed from exonuclease digestion of peak 4 (**a**) the 5'- exonuclease digestion spectrum and (**b**) the 3'-exonuclease digestion spectrum at different time points. Left side of the figure shows the theoretical molecular weights of the fragments that should form after the 3'- and 5'-exonuclease digestion of peak 4 if ODN is modified at G_1 and G_3 positions.

S7

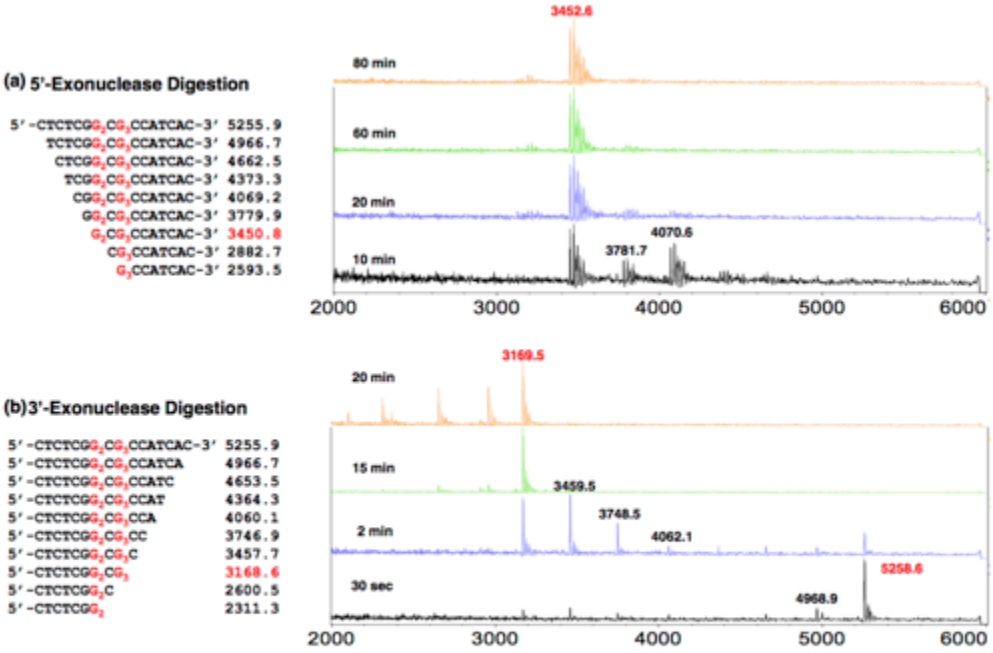


Figure S3: MALDI-TOF mass spectra of the hydrolysis fragments formed from exonuclease digestion of peak 5 (**a**) the 5'- exonuclease digestion spectrum and (**b**) the 3'-exonuclease digestion spectrum at different time points. Left side of the figure shows the theoretical molecular weights of the fragments that should form after the 3'- and 5'-exonuclease digestion of peak 5 if ODN is modified at G_2 and G_3 positions.

S8

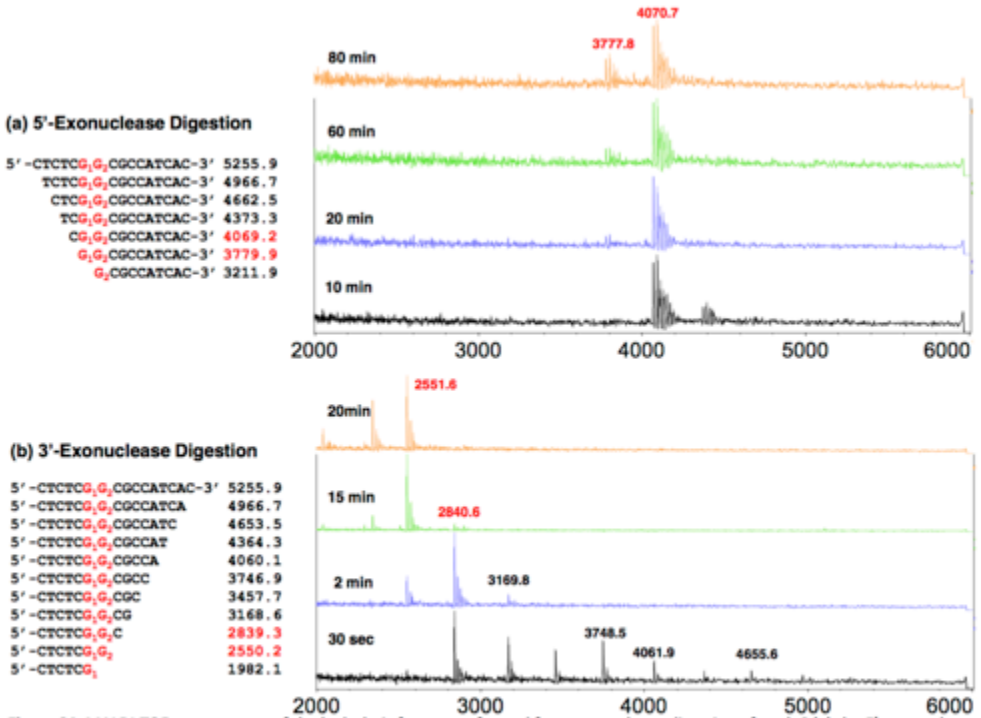


Figure S4: MALDI-TOF mass spectra of the hydrolysis fragments formed from exonuclease digestion of peak 6 (**a**) the 5'- exonuclease digestion spectrum and (**b**) the 3'-exonuclease digestion spectrum at different time points. Left side of the figure shows the theoretical molecular weights of the fragments that should form after the 3'- and 5'-exonuclease digestion of peak 6 if ODN is modified at G_1 and G_2 positions.

S9

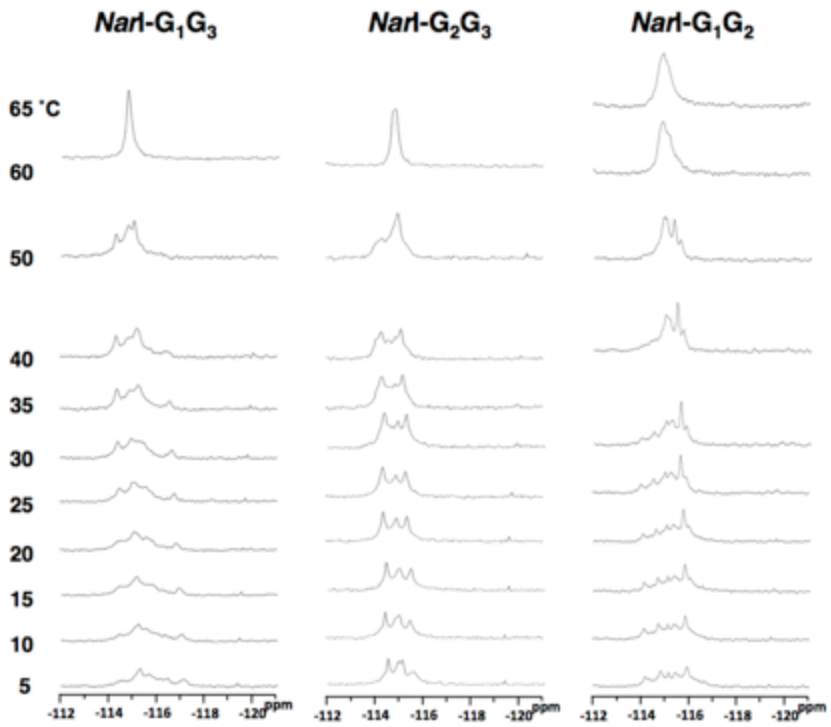


Figure S5: ^{19}F NMR spectra of di-FAAF-modified duplexes; *Nar*I-G_1G_3 (left), *Nar*I-G_2G_3 (middle), and G_1G_2 (right) at different temperatures.

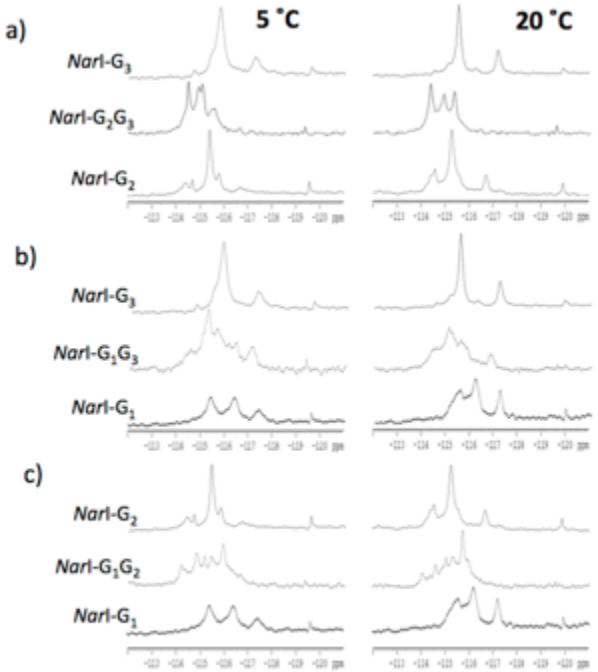


Figure S6: Comparison of ^{19}F NMR spectra of di-FAAF-modified duplexes with their respective mono-adducts at 5 °C (left) and 20 °C (right).

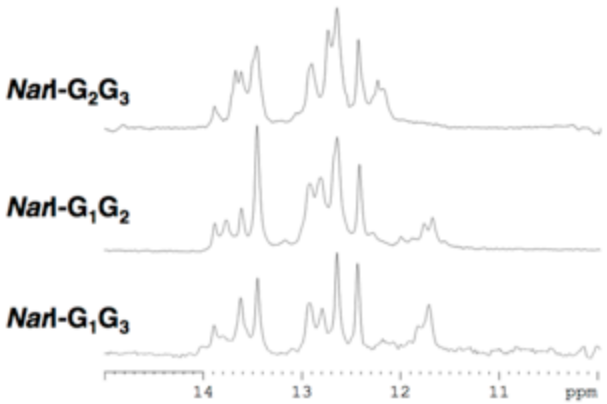


Figure S7: Imino proton spectra of FAAF-modified di-adduct duplexes; *Nar*I-G_1G_3 (bottom), *Nar*I-G_2G_3 (top), and G_1G_2 (middle) at 20˚C

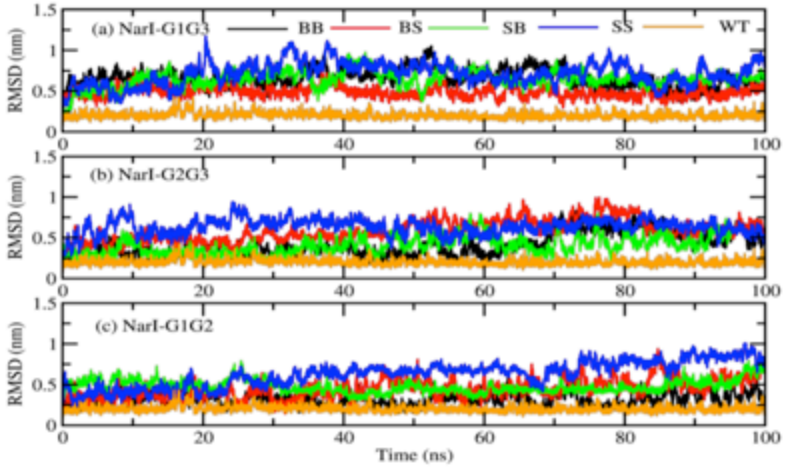


Figure S8: RMS differences (RMSD) of the DNA di-adducts simulations. RMSD were calculated based on the heavy atoms by excluding the terminal two nucleotides at each end of DNA duplexes. Structures were least-squares aligned to the initial model geometries prior to calculation of RMSD values.

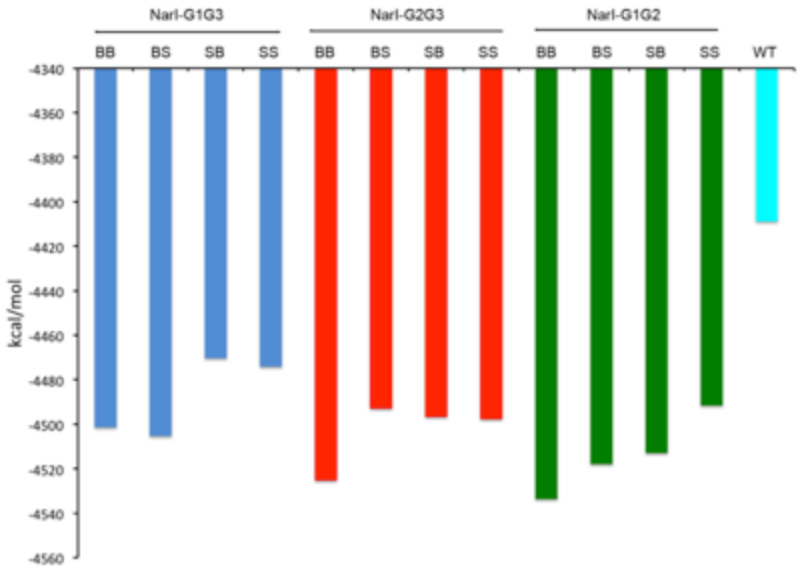


Figure S9: Total energies calculated from MM/GBSA. See Supporting Table S1 for individual energy components.

S14

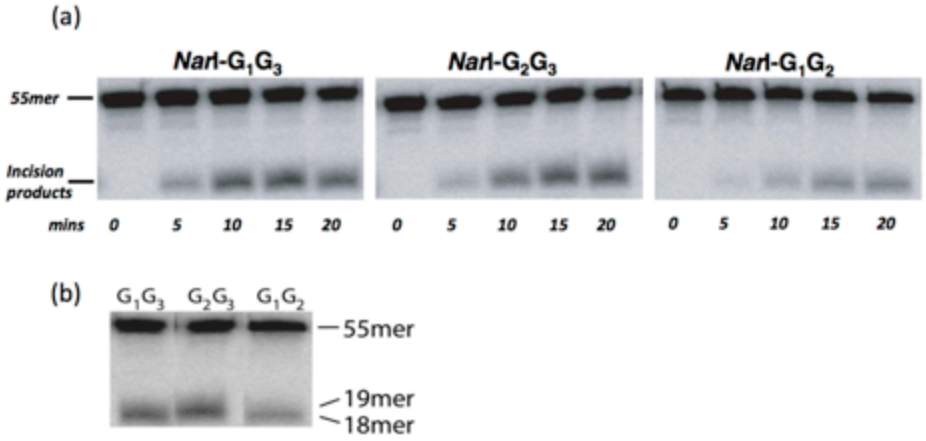


Figure S10: (a) The 5'-terminally labeled FAAF modified *Nar*I di-adduct substrates (2 nM) were incubated with UvrABC (UvrA, 10 nM, UvrB, 250 nM, and UvrC, 100 nM) in UvrABC reaction buffer at 37°C for the time period mentioned above. The incision products were then analyzed on a 12% polyacrylamide sequencing gel under denaturing condition. The 55-mer represents the intact DNA substrates, and the incision products represent the 5'-incised DNA fragments. (b) Side-by-side comparison of incision product at 20 min incubation with UvrABC nuclease. Incision products are observed as 18 or 19-mer with the determination of product size being dependent on the position of the 5' adduct. The results are tabulated and relative incision efficiencies are plotted in Figure 7.

S15

Table S1: MM/GBSA energetic analysis of the MD simulations. Individual contributions of the molecular mechanics (MM), solvent accessible surface area (SASA), and Generalized Born (GB) energies are shown (kcal/mol). Standard error were computed from block averages by using 8 blocks.

	MM	SASA	GB	total
***Nar*I-WT**	5241.9±25.4	45.2±0.2	-9695.9±25.4	-4409±25
***Nar*I-G_1G_3**				
BB	4639.3±83.9	49.0±0.2	-9189.7±83.9	-4501±84
BS	4713.6±32.7	48.5±0.4	-9267.4±30.3	-4505±33
SB	5126.0±213.5	51.0±1.1	-9647.2±205.2	-4470±214
SS	4699.9±92.0	48.7±0.7	-9222.6±97.2	-4474±98
***Nar*I-G_2G_3**				
BB	4873.9±114.3	48.1±0.8	-9447.3±121.6	-4525±122
BS	4957.2±152.3	49.4±0.9	-9499.6±146.8	-4493±153
SB	4928.5±78.1	49.0±0.2	-9474.2±75.5	-4497±80
SS	4594.1±146.6	48.8±0.4	-9140.6±152.5	-4498±153
***Nar*I-G_1G_2**				
BB	4868.4±49.7	48.8±0.3	-9450.8±49.4	-4534±50
BS	4835.6±25.8	48.9±0.4	-9402.4±23.3	-4518±30
SB	4925.3±84.4	48.6±0.7	-9486.8±86.5	-4513±90
SS	5066.3±138.1	48.7±1.4	-9606.6±127.8	-4492±140

Table S2: Clustering analysis of the 12 di-adducts MD simulations from the 20 to 100 ns regions of the MD simulations. Highlighted structures are included in Figure 6 of the main text.

	Number of frames		
***Nar*I-G1G3**	**cluster 1**	**cluster 2**	**cluster 3**
BB	3028	955	
BS	3725	240	
SB	3323	622	
SS	387	3559	
***Nar*I-G2G3**			
BB	3975	20	
BS	1789	1511	
SB	3996	4	
SS	2245	1738	
***Nar*I-G1G2**			
BB	3978	20	
BS	3975	18	
SB	821	1981	945
SS	201	306	3040

Table S3: Base pairing for all the DNA di-adducts.

NarI-WT	99.9%
NarI-G1G3	
NarI-G1G3-BB	79.8%
NarI-G1G3-BS	83.3%
NarI-G1G3-SB	66.5%
NarI-G1G3-SS	49.6%
Average	69.8%
NarI-G2G3	
NarI-G2G3-BB	93.5%
NarI-G2G3-BS	76.4%
NarI-G2G3-SB	91.1%
NarI-G2G3-SS	66.2%
Average	81.8%
NarI-G1G2	
NarI-G1G2-BB	99.7%
NarI-G1G2-BS	85.8%
NarI-G1G2-SB	83.9%
NarI-G1G2-SS	71.0%
Average	85.1%

Table S4: Base stacking energies (kcal/mol)

	vdw	elec	total
NarI-G1G3			
BB	-141.9±5.4	-10.5±4.1	-152.4±2.3
BS	-151.0±2.7	-4.6±2.4	-155.6±3.6
SB	-134.2±9.7	-11.3±4.7	-145.5±12.5
SS	-118.6±4.0	-27.75.4	-146.4±6.3
NarI-G2G3			
BB	-154.8±8.8	-5.2±1.3	-160.0±10.0
BS	-136.4±6.0	-13.1±3.4	-149.5±6.7
SB	-154.7±3.3	-7.8±4.2	-162.5±2.8
SS	-129.3±10.4	-15.1±5.5	-144.4±15.0
NarI-G1G2			
BB	-161.3±1.3	-9.0±0.6	-170.2±1.5
BS	-152.1±1.9	-17.1±9.2	-169.1±9.3
SB	-145.9±5.3	-9.4±2.0	-155.4±6.8
SS	-129.9±6.0	-3.4±2.8	-133.3±8.5
NarI-WT	-173.2±0.6	-5.0±0.5	-178.2±0.8

Table S5: Conformer population ratios (%)

*Nar*I duplexes	Percent Conformers (%)		
	B	***S***	***W***
***Nar*I-G_1**	46	34	20
***Nar*I-G_2**	57	15	9
***Nar*I-G_3**	13	61	26
***Nar*I-G_1 + *Nar*I-G_2**[a]	103	49	29
***Nar*I-G_2 + *Nar*I-G_3**[a]	70	76	35
***Nar*I-G_1 + *Nar*I-G_3**[a]	59	95	46

[a] The sum of percent population ratios of respective mono-adducts to estimate *Nar*I di-adduct duplexes; *Nar*I-G_1G_2, -G_2G_3 and -G_1G_3.

CHAPTER 4

DISSOCIATION DYNAMICS OF XPC-RAD23B FROM DAMAGED DNA IS A DETERMINING FACTOR OF NER EFFICIENCY

Abstract

XPC-RAD23B (XPC) plays a critical role in human nucleotide excision repair (hNER) as this complex recognizes DNA adducts to initiate NER. To determine the mutagenic potential of structurally different bulky DNA damage sites, various studies have been conducted to define the correlation of XPC-DNA damage equilibrium binding affinity with NER efficiency. However, little is known about the effects of XPC-DNA damage recognition kinetics on hNER. Although association of XPC is important, our current work shows that the XPC-DNA dissociation rate also plays a pivotal role in achieving NER efficiency. We characterized for the first time the binding of XPC to mono- versus di-AAF-modified sequences by using the real time monitoring surface plasmon resonance (SPR) technique. Strikingly, the half-life ($t_{1/2}$ or the retention time of XPC in association with damaged DNA) shares an inverse relationship with NER efficiency. This is particularly true when XPC remained bound to clustered adducts for a much longer period of time as compared to mono-adducts. Our results suggest that XPC dissociation from the damage site could become a rate-limiting step in NER of certain types of DNA adducts, leading to repression of NER. The results also support exploiting this mechanism as a new therapeutic approach in cancer therapy.

Introduction

The human genome is constantly under assault from exogenous and endogenous causes of DNA damage. The formation and propagation of the resulting adducts can be particularly destructive when these mutations occur within tumor suppressing genes, leading to tumorigenesis [1-4]. Consequently, human cells have several effective DNA repair pathways to protect against the plethora of genotoxic bombardments to the genome [5]; however, the mechanism by which damage-recognition proteins distinguish damage sites remains uncertain. Mutations that arise in genes associated with the nucleotide excision repair pathway (NER) result in a multitude of genetic disorders such as *xeroderma pigmentosum*, which is characterized by sensitivity to sunlight and, ultimately, the development of carcinomas [6].

NER is utilized to remove primarily bulky adducts, plus cross-links, and various other lesions [7-9]. NER is either associated with transcription in transcription-coupled repair (TCR) or independently of transcription in global genome repair (GGR). GGR in *Escherichia coli* consists primarily of a collaborative effort of three proteins that both recognize and incise damaged bases: UvrA, UvrB, and UvrC [10]. Two UvrA molecules associate and then form a trimeric complex with UvrB. This trimeric complex is thought to be the DNA damage sensor. UvrA facilitates UvrB binding and positions UvrB to confirm the existence of a damage site. Once UvrB is in the correct position, UvrA utilizes its ATPase activity to dissociate from the preincision complex. UvrB then recruits UvrC endonuclease, which incises the damaged DNA strand by 3' and 5' cleavages flanking the damage site [11-14]. In human GGR the $UvrA_2B$ equivalent is *Xeroderma pigmentosum* group C (XPC) in complex with RAD23B (XPC-RAD23B, henceforth XPC). The XPC complex acts in the DNA

damage recognition step, thus initiating GGR [15]. XPC has been shown to bind at the site of many types of damage *in vitro* and in UV-treated cells arrives at damage sites before other NER factors [9,16-18]. Once at the damage site XPC recruits the multi-subunit transcription factor TFIIH, including the helicase subunits of XPB and XPD, followed by XPA for damage confirmation, fork binding and subsequent recruitment of replication protein A (RPA) for single-stranded DNA (ssDNA) stabilization, and XPG and XPF-ERCC1 for dual incisions [19-22].

Crystal structures of XPC in association with undamaged or damaged DNA revealed a mechanism by which XPC hops along DNA until a thermodynamically stable recognition complex is formed, which effectively distinguishes damaged from non-damage sites [23,24]. Further studies have supported this hypothesis by suggesting that residence time of XPC on DNA damage sites may play a role in the relationship between XPC binding and NER efficiency [25,26]. Binding affinity of XPC at the damage site has been suggested to be the rate-limiting step for NER [25,27,28]. Although various efforts have been made to correlate the equilibrium binding of damage recognition to overall NER efficiency, little is known about the role of the kinetics of damage recognition in NER.

Arylamines and heterocyclic amines are notorious environmental carcinogens. The DNA adduct-forming arylamines can be found naturally in the environment, in addition to a number of unnatural sources such as cigarette smoke and hair dyes. Heterocyclic amines are most notably abundant in meat that has been cooked at high temperatures. It is inevitable that a person will be exposed to one or both of these carcinogens in his/her lifetime. Each of these mutagens has been documented to cause many types of cancer, such as breast, liver, and bladder, to name a few [2]. Metabolic activation of these amines *in vivo*

produces C8-substituted dG as the major bulky DNA adduct [29]. A well-known example is the human bladder carcinogen 4-aminobiphenyl [30]. The prototype environmental arylamine 2-aminofluorene produces two major DNA adducts via *in vivo* activation: N-(2′-deoxyguanosin-8-yl)-2-aminofluorene (AF) and N-(2′-deoxyguanosin-8-yl)-2-acetylaminofluorene (AAF) (Fig. 1a) [31]. Their fluorine derivatives FAF and FAAF have been used extensively as ^{19}F NMR conformational models for these bulky arylamine lesions [32-34]. Conformational studies have shown that FAF in a fully paired duplex DNA can adopt an equilibrium between two prototype conformers, while FAAF incorporates an additional conformation due to a single bulky acetyl group on the central nitrogen. This equilibrium exists between major groove binding anti B-type conformers, base-displaced *syn* stacked S conformers, and minor groove binding of the *syn* adduct wedge (W) conformer (Fig. 1c) [33,35]. Usually, the damage produced is a single mono-adduct: however, cluster di-adducts can form *in vivo*, though less frequently than mono-adducts as discussed previously [36-38]. Past work has shown that adduct conformation is strongly dependent on the flanking sequence, which modulates mutational and repair outcomes [27,39-41]. One such sequence is the mutational hotspot known as the *Nar*I sequence (5'- . . . $CG_1G_2CG_3CC$ . . . -3') (Fig. 1b), which has been extensively studied [33,42].

The substrate reparability of adducts in the *Nar*I sequence has been tested in both the *E. coli* UvrABC and human endonuclease systems and were found to be sequence dependent [27,33,43]. In addition, different repair efficiencies of the same lesions were observed between the two systems [27,33]. Furthermore, recent work has attempted to correlate the binding affinities of repair proteins with adduct excision or NER efficiency [25-27]. Yeo and colleagues, implementing electrophoretic mobility shift (EMSA) and dual-

incision assays, concluded that increased DNA thermodynamic destabilization, XPC-RAD23B binding, and overall NER efficiency of AAF adducts are directly correlated [25]. In contrast, H. Mu *et al.* showed that NER efficiencies of the same AAF lesions are correlated with greater extents of base sequence-dependent local untwisting and minor groove opening together with weaker stacking interactions [27]. Lee *et al.* have found minimal differences in XPC binding affinities of lesions derived from bulky polycyclic aromatic hydrocarbons while observing dramatic differences in NER efficiency [26]. These three individual reports employed different bulky adducts in their studies; however, Shell *et al.* demonstrate that XPC acts as a general sensor for DNA damage, with a preferential binding to damage sites, but concluding that lesion identity is not a determinant of XPC binding affinity [44].

In the present study, we analyzed the kinetic aspects of human XPC protein interaction with AAF arylamine adducts in the *Nar*I sequence context by surface plasmon resonance (SPR) analysis and defined the relationship with NER. SPR has the significant advantage over other methods designed to observe protein-DNA interactions in that the interaction can be observed in a dynamic real-time environment, much closer to native conditions. We show that at lesion clusters the kinetic off-rate of XPC has an inverse correlation to repair efficiency. In other words, the $t_{1/2}$ (time required for 50% of the bound XPC to dissociate from the DNA, $t_{1/2}$ (s) = ln(2)/$k_{d,}$) of the damage recognition complex inversely correlates to NER efficiency. This work reveals the significance of the dynamics of XPC recognition of conformationally diverse DNA adducts in NER. Here, we describe a new model for XPC activation of NER where the off-rate of XPC from the lesion site, particularly

in the case of cluster-lesions, is the rate-limiting step of NER, and propose applying this finding to design a more efficiently targeted approach to cancer therapy.

Materials and Methods

Caution

2-Aminofluorene derivatives are mutagens and suspected human carcinogens and, therefore, must be handled with caution.

Crude desalted oligodeoxynucleotides (1 μmol) were purchased from Operon (Eurofin, Huntsville, AL) and purified by reverse phase HPLC. All HPLC solvents were purchased from Fisher Inc. (Pittsburgh, PA).

Substrate preparation and characterization

Modified duplexes of 55 bp DNA substrates containing mono-FAAF and di-FAAF adducts in the *Nar*I sequence context were constructed as previously described [32,35,38]. The HPLC purification system consisted of a Hitachi EZChrome Elite system with an L2450 diode array as a detector and a Clarity column (10 mm × 150 mm, 3 μm) (Phenomenex, Torrance, CA). Purification of oligodeoxynucleotides was carried out using a 20 min linear gradient mobile phase system from 5 to 20% (v/v) acetonitrile with 100 mM ammonium acetate buffer (pH 6.5) at a flow rate of 2.5 mL/min.

Preparation of arylamine-modified template

The modified 55-mer biotinylated DNA templates were prepared according to published procedures [45-47]. Mono- and di-adduct oligodeoxynucleotides in the *Nar*I modified strand were purified by HPLC (described above) and characterized by Shimazdu Axima MALDI-TOF mass spectrometry as previously reported [38] (Fig. 1 and Supplementary Fig. 3). 5'-Biotinylated 55-mer (1 OD) was annealed with 55-mer complementary strand (1.05 ODs) in 1x HBS-EP$^+$ buffer for 5 min at 95°C. Identical unmodified duplexes were concurrently prepared as controls. The annealed oligodeoxynucleotides were then used for SPR experiments.

Oligonucleotide sequence used for surface plasmon resonance

5'-biotin-CCACTCCTATCCACCATCCATCTTACTCTCG$_1$G$_2$CG$_3$CCATCACCACTCACCACCACA-3'

3'-GGTGAGGATAGGTGGTAGGTAGAATGAGAGC C GC GGTAGTGGTGAGTGGTGGTGT-5'

G$_1$, G$_2$, and/or G$_3$: dG or dG-FAAF

Purification of XPC-RAD23B protein complex

The XPC-RAD23B protein complex was purified as described previously [48,49]. XPC-RAD23B protein was prepared from Sf21 insect cells infected with recombinant baculovirus expressing XPC and RAD23B proteins (graciously provided by A. Sancar, University of North Carolina, Chapel Hill). Protein concentration was determined using the Bio-Rad protein assay. SDS-PAGE (10%) and Western blotting confirmed the purity of the XPC-RAD23B complex.

Immobilization of streptavidin on CM5 chip and DNA coating

SPR measurements were conducted with a Biacore T200 (GE Healthcare). Streptavidin (SA) was immobilized on a CM5 dextran chip using an amine-coupling method [45,47]. Four flow cells were immobilized with streptavidin amine to ~2,200 resonance units (RU). Flow cell 1 was used as a reference. Before the coating of biotinylated DNA templates over SA, the surface was washed with 50 mM NaOH five times, each with 60s pulses at 100 µl/min to remove any free SA until the change in response units was below 5 RU. The surface was further injected 3-4 times with HBS-P+ running buffer (10 mM HEPES, 150 mM NaCl, 0.05% non-ionic surfactant P20) to remove any residual NaOH in the microfluidics path and to stabilize the surface. Biotinylated unmodified and various FAAF-modified DNA duplexes (0.025 nM) were injected at 100 µl/min for 240-300s over the flow cells 2, 3, or 4 to achieve 2-5 RU relative to flow cell 1, which was a blank reference. Any unbound DNA was washed away with running buffer.

Kinetics analysis

The binding kinetics for the interaction of UvrA or XPC with DNA was determined by injecting the UvrA (0-500 nM) or XPC (0-5 nM) in HBS-P+ running buffer containing 5 mM $MgCl_2$, 1 mM DTT and BSA (100 µg/mL). The flow rate was 100 µL/min for 30 s followed by dissociation for 60 s. The DNA surface was conditioned by a sequential injection involving 1x HBS-P+ running buffer, 3x HBS-P+ buffer, and 4x HBS-P+ buffer (prior to addition of enzyme). The surface was regenerated with a 30 s injection of 0.05% SDS with a flow rate of 100 µL/min, followed by a wash with HBS-P+ running buffer. Experiments were repeated 3x with duplex injections of the indicated concentrations. UvrA binding studies

were performed in the absence and presence of ATP (0.5 mM). All SPR sensograms were double referenced and fitted using a simple 1:1 Langmuir model (Supplementary Fig. 1). Processing included zeroing and cropping data, aligning injection times, fitting of binding curves and off-rate analysis. The equilibrium dissociation constant (K_D) for ternary systems was calculated using the steady-state affinity analysis in the BIA-Evaluation software package v2.0 provided by the manufacturer, General Electric. The average of the data (with standard deviation) of K_D, k_a, and k_d is shown in Table 2 and Supplementary Tables 1 and 2. The Scrubber software package (BioLogic Software) was used to process off-rate analysis of raw XPC-H23B SPR binding sensograms (Fig. 4b, c). Curve fittings were not ideal for certain UvrA data (Supplementary Fig. 2), which affected the reliability of rate constants (see Results).

Electrophoretic Mobility Shift Assay (EMSA)

Binding of XPC to various DNA substrates was analyzed by a gel mobility shift assay as described previously[49]. Typically, DNA substrates (0.5-1 nM) were incubated with varying concentrations of protein at 30°C in 20 μL of binding buffer [20 mM Hepes-KOH, pH 7.9, 75 mM KCl, 5 mM $MgCl_2$, 1 mM DTT, 5% glycerol, 100 μg/mL acetylated BSA (Promega)]. Reactions then were placed on ice, 2 μL of 80% (v/v) glycerol was added, and the mixture was immediately loaded onto a 3.5% native polyacrylamide gel and electrophoresed at 80 V in 1× TBE buffer for 2 h at 4°C. The gels were dried and exposed to phosphoimage screens overnight. Quantification of the radioactivity was carried out using a Fuji FLA-5000 scanner with the ImageGuage software.

Construction of closed-circular plasmid with adducts

The following double-stranded oligonucleotide was cleaved by KpnI and Xba1 (both lowercase) for insertion into the multiple cloning site of the vector pTZ19U between the *XbaI* and *KpnI* restriction sites using the QuikChange II Site-Directed Mutagenesis Kit (Agilent Technologies). The 16-mer sequence containing the *Nar*I hotspot of Figure 1b is underlined.

NER_ins1.....5'-CGGgtaccCCGCTCTCGGCGCCATCACTTAGtctagaCTAG-3'

NER_ins2.....3'-GCCccatggGGCGAGAGCCGCGGTAGTGAATCagatctGATC-5'

The NER-pTZ19U plasmid was propagated in *E. coli* (DH5α) cells, which were infected with virus M13KO7 (NEB) to generate single-stranded plasmid which was purified using the M13 isolation maxi kit (Omega Biotek). Closed-circular double-stranded plasmid containing adduct was made by priming the single-stranded NER-pTZ19U plasmid with 5'-phosphorylated 16-mer oligos (Fig. 1b) containing adduct in the presence of Sequenase 2.0 (Affymetrix) and T4 DNA ligase (Promega) following the manufactures protocol. Closed-circular plasmid DNA containing adduct then was purified by agarose gel electrophoresis and elution.

HeLa whole-cell extract preparation

Whole cell extracts were prepared from HeLa cell pellets purchased from the National Cell Culture Center. The thawed cell pellet was resuspended in four packed-cell volumes (PCV) of 10 mM Tris-HCl pH 8.0, 1 mM EDTA, 5 mM DTT, then incubated on ice for 20 min. The cells were lysed by homogenization in a Dounce homogenizer using eight strokes of the B pestle. Four PCV of 50 mM Tris-HCl pH 8, 10 mM $MgCl_2$, 2 mM DTT, 25%

sucrose (w/v), 50% glycerol (v/v) were then added and the mixture was stirred gently. One PCV of saturated $(NH_4)_2SO_4$ (pH 7.0) was added slowly, then the mixture was stirred for 20 min at 4°C. The lysate was clarified by centrifugation at 11,500xg for 30 min at 4°C. The supernatant was transferred to a fresh tube and solid $(NH_4)_2SO_4$ (0.33g/ml of suspension) was added. The suspension was mixed for 30 min and 0.01 ml of 1M NaOH per 10 grams of $(NH_4)_2SO_4$ was added. The precipitated proteins were collected by centrifugation at 11,500xg for 30 min and resuspended in dialysis buffer (25 mM Hepes-KOH pH 7.9, 100 mM KCl, 12 mM $MgCl_2$, 0.5 mM EDTA, 2 mM DTT, 12% glycerol) and was dialyzed against the same buffer overnight. Following dialysis the extract was clarified by centrifugation at 10,000xg for 10 min and aliquots of the supernatant were stored at -80°C.

Dual incision assay

The dual incision assay was adapted from Shivji *et al*. [50]. All incisions were carried out in a total reaction volume of 10 µl. A reaction mixture with 100 µg HeLa whole-cell extract protein in 5x repair buffer [200 mM Hepes-KOH, 25 mM $MgCl_2$, 110 mM phosphocreatine (di-Tris salt, Sigma), 10 mM ATP, 2.5 mM DTT and 1.8 mg/ml BSA, (adjusted to pH 7.8), 0.2 µl 2.5 mg/ml creatine phosphokinase (rabbit muscle CPK, Sigma)] was preincubated at 30°C for 10 min. The incision reaction was started with the addition of 50 ng of adduct-containing plasmid DNA, and incubation was continued for another 45 min at 30°C. Samples were placed on ice for 10 min, then 0.5 µl of 1 µM 3'-phosphorylated primer (5'GGGGCAGGTGATGGCGCCGAGAGGGATCCCC-3') was added and the mixture was heated to 95°C for 5 min and allowed to cool to room temperature for 15 min. Then, a sequenase/[α-^{32}P]-dCTP mix was added at 0.25 U of sequenase and 2.5 µCi of [α-^{32}P]-dCTP

per reaction. The reaction mixture was incubated at 37°C for 3 min before addition of dNTP mix (50 μM dCTP and 100 μM dATP, dGTP, dTTP) followed by an additional 12 min incubation. The reaction was stopped by addition of loading dye and heated to 95°C for 5 min before electrophoresis through a 12% Sequagel (National Diagnostics). Reaction products were visualized with a Fuji Film FLA-5000.

Results

Model Systems

Mono- and di-FAAF adducted substrates were prepared within the *Nar*I core sequence (5' -CTCTCG$_1$G$_2$CG$_3$CCATCAC-3', Fig. 1b) as previously reported [33,38,39] and their corresponding duplexes were used to examine their structural and thermodynamic properties (see below). This common *Nar*I sequence was ligated to prepare 55 bp substrates for SPR and double-incision assays on closed circular plasmid DNA. Depending on the location of the FAAF, the mono-adducts were designated as *NarI*-G$_1$, *Nar*I-G$_2$ or *Nar*I-G$_3$, in which G$_1$, G$_2$ and G$_3$ signify the position of modified guanine. The di-adducts were designated as *Nar*I-G$_1$G$_2$, *Nar*I-G$_2$G$_3$, or *Nar*I-G$_1$G$_3$, in which the numbers signify the positions of FAAF-modified guanines (Fig. 1b). In the DNA duplex, the bulky *N*-acetylated AAF can exist in three conformers: B-type, in which the carcinogen occupies the major groove of a double-helical DNA without inducing any significant perturbations, stacked (S), in which the carcinogen is intercalated between neighboring nucleobases in the duplex and the glycosidic linkage to the modified guanine is *syn*, and a wedge (W) conformer, in which

the hydrophobic fluorene moiety resides in the narrow minor groove area (Fig. 1c) [34,51-53]. The population ratios of these different conformations were found to be strongly dependent on the nature of the flanking sequence context around the lesion within the *Nar*I sequence [35,52].

Thermodynamics of mono- and di-FAAF adducts and *E. coli* NER

We previously studied the structures and thermodynamics of mono- and di-FAAF duplex adducts derived from the afore-mentioned *Nar*I-16-mer sequence [38]. The thermodynamic results summarized in Table 1 revealed that di-adducts produced a synergistic effect on duplex destabilization relative to the mono-adducts. Briefly, the lesion-induced thermal instabilities (ΔTm, Table 1) were substantially greater for di- (-10.0 ~ -17.9 ^{o}C) over mono-adducts (-5.3 ~ -8.3 ^{o}C). Within the di- adducts, the thermal stability was generally in the order of *Nar*I-G_1G_2 > *Nar*I-G_1G_3 > *Nar*I-G_2G_3. Molecular dynamic simulation data indicated the perturbations of nucleotide base stacking are a major contributor to the observed sequence effect. The di-adducts were more reparable in *E. coli* than the corresponding mono-adducts [33,38]. Moreover, we observed a dramatic trend in repair efficiency in *E. coli* in the reverse order of thermal stability, i.e., *Nar*I-G_2G_3 > *Nar*I-G_1G_3 >> *Nar*I-G_1G_2 (Fig. 3a). Taken together, these results indicate the importance of base stacking and related thermal and thermodynamic instability in the repair of bulky cluster arylamine DNA adducts. However, assessing the overall thermodynamic stability of di-adducts was complicated by the fact that the di-adducts sample a complex range of S/B/W-conformational heterogeneity [38]. The single dG-FAAF adduct at the G_3 position of the

Figure 1

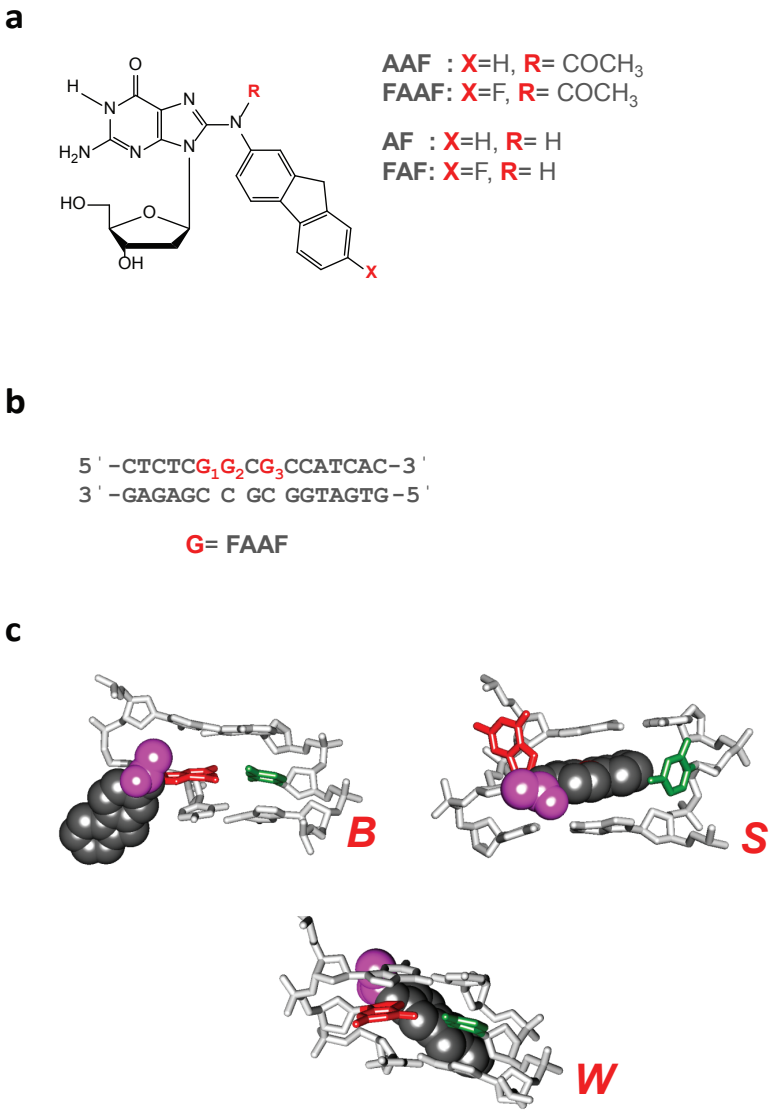


Figure 1. Adduct structures and sequences. **(a)** Structure of AAF [*N*-(2'-deoxyguanosin-8-yl)-2-acetylaminofluorene], AF [*N*-(2'-deoxyguanosin-8-yl)-2-aminofluorene] and fluoro models, FAAF [*N*-(2-deoxyguanosin-8-yl)-7-fluoro-2-acetylaminofluorene], FAF [*N*-(2'-deoxyguanosin-8-yl)-7-fluoro-2-aminofluorene] ; **(b)** Fully-paired 16-mer duplexes containing the central *Nar*I sequence (CGGCGCC) used in SPR, EMSA and *in vitro* NER constructs illustrating the placement of the adducted bases at G_1, G_2, and G_3 positions; **(c)** Major groove views of the B-, S-, and W-conformers of AAF. Modified-dG (red), dC (green) opposite the lesion site (orphaned C), fluorene (grey CPK), *N*-acetyl (magenta).

*Nar*I sequence exhibited a preference for the S conformation (61%). Interestingly, greater instability was observed for the G_3-containing *Nar*I-G_2G_3 and *Nar*I-G_1G_3 duplexes that produce a higher combined S population (~76 and ~95%, respectively) than *Nar*I-G_1G_2 (~49%). In addition, the greater proximity of two FAAF lesions in *Nar*I-G_2G_3 (e.g., just one base apart) compared to that in *Nar*I-G_1G_3 (two bases apart) possibly could induce a greater DNA distortion.

Figure 2

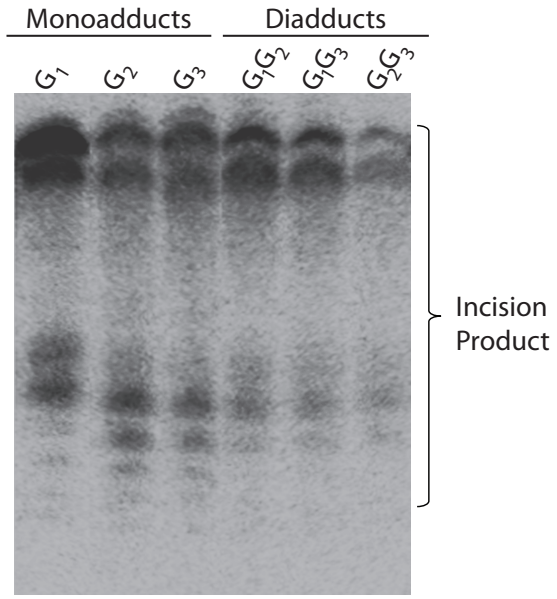


Figure 2. NER dual incision at adducts in the *Nar*I sequence in the human NER system. Plasmids containing site-specific mono-FAAF (lanes G_1, G_2, G_3) or di-FAAF adducts (lanes G_1G_2, G_2G_3, G_1G_3), were incubated with HeLa whole-cell extracts. Detection of excision products was by 3'-end labeling using a complementary oligonucleotide containing a 5'-GGGG base overhang. The reaction products were resolved on a 12% denaturing polyacrylamide gel run under constant current. The excision products are indicated on the right of the gel.

Human NER of mono- and di-FAAF adducted DNA

To examine the efficiency of NER in excising the mono- and di-FAAF lesions in the human system, oligonucleotides containing FAAF were generated within the *Nar*I sequence, as described above. These modified oligonucleotides then were incorporated into a closed circular plasmid and subjected to excision with extracts from untreated HeLa cells which contains the complete NER machinery. A complementary oligonucleotide with a 4-guanine overhang that matches the incision product then was used to generate a radioactive product for visualization and quantification of incision products. Incision experiments were carried out in triplicate under identical conditions, and then the relative incision was measured by separating the incision extension products in a denaturing urea gel as shown in Fig. 2. In contrast to the unmodified substrate the modified substrates generated products ranging in size from 25-34 bases, with the incision products indicated to the right.

Of the *Nar*I dG-FAAF adducts, G_1-FAAF was observed to have the maximum incision; therefore, all other *Nar*I adducts were normalized to *Nar*I-G_1-FAAF (Fig. 3b). The relative incision, therefore, was 1 for G_1-FAAF, 0.69 for both G_2-FAAF and G_1G_2-FAAF, 0.65 for G_3-FAAF, 0.30 for G_1G_3-FAAF, and finally 0.12 for G_2G_3-FAAF. The observed order for incision efficiency in the mono-adducts was $G_1 > G_2 \sim G_3$. The relative hNER efficiencies of the di-FAAF adducts in the *Nar*I sequence were $G_1G_2 > G_1G_3 > G_2G_3$. Previous work on these same mono-adduct substrates has yielded a variety of conclusions [25,27,28]; however, this is the first report to include hNER incision analysis of cluster di-AAF adducts. One interesting observation is that the most thermodynamically stable di-adduct, G_1G_2, has no base separating the lesions and, in this respect, is structurally comparable to the UV-induced

cyclopyrimidine dimer (CPD) adduct. Work from the Min lab has allowed a better understanding of how XPC binds undamaged and damaged DNA, specifically via the CPD adduct [23,24]. Other observations indicate comparability between the cluster di-adduct G_2G_3, which has just one base separating the lesions, and di-nuclear *cis*-platinum complexes [54]. These observations, discussed in further detail below, potentially have significant biological implications.

E. coli and human NER produce different incision efficiency patterns on common substrates

Interestingly, there is no direct correlation when comparing the NER incisions between the *E. coli* and human systems (Fig. 3a and 3b, respectively). Our previous work revealed that in the *E. coli* system mono-FAAF adducts were excised in the order of $G_3 \sim G_1 > G_2$ while di-FAAF adducts were excised in the order of $G_2G_3 > G_1G_3 > G_1G_2$. In the *E. coli* system, the di-adducts overall are more readily incised by the UvrABC system; in contrast, in the human system the mono-adducts have significantly greater incision versus the di-adducts. Another interesting observation is that the mono- and di-adducts that produced the most incision product in the *E. coli* system (G_3 and G_2G_3), led to the least incision product in the human system.

Although many processes in eukaryotic cells are conserved from prokaryotic systems and operate by similar mechanisms, in the case of NER the issue becomes more complex in that the UvrABC system is made up of only three proteins while in the human system nearly 30 proteins carry out the same function. Also, in humans multiple pathways arose to deal with the repair of a much more complex genome that also exists in a chromatin structure.

Figure 3

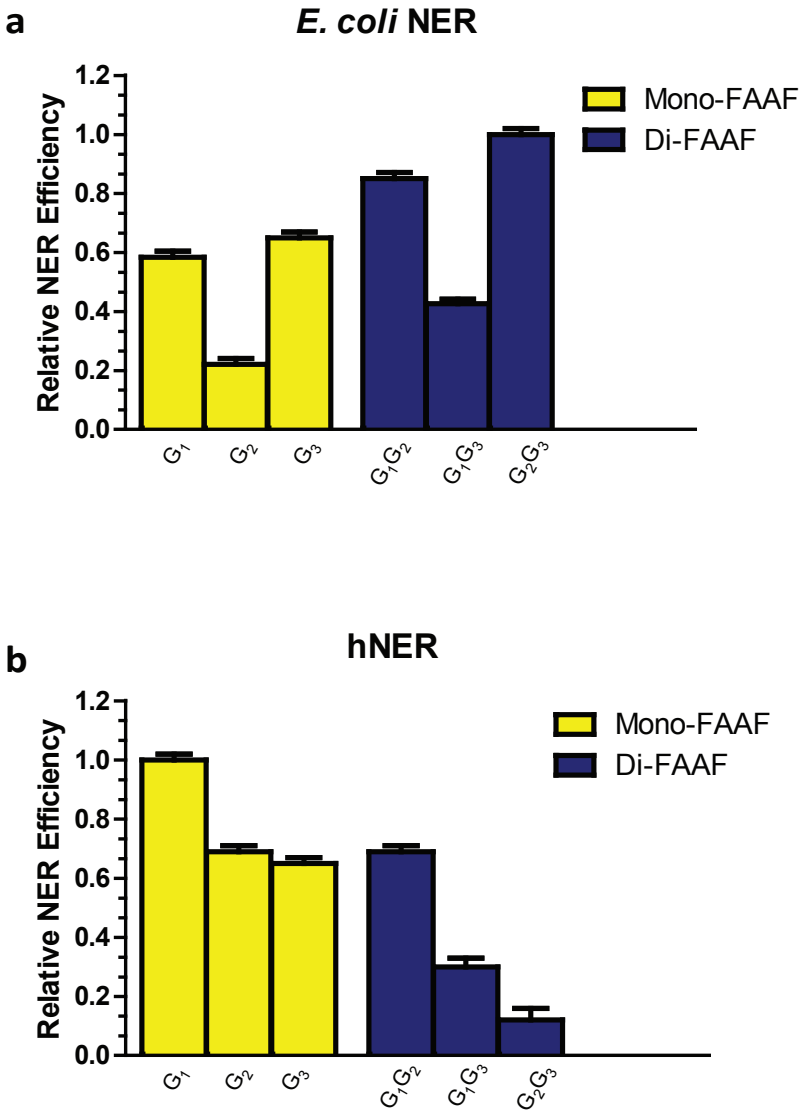


Figure 3. The efficiencies of NER at adducts in the *Nar*I sequence in the *E. coli* or the human NER system. The relative incision rates of mono-FAAF and di-FAAF adducts in the histogram were calculated by normalizing the mono- (gold) and di-adducts (blue) relative to the *Nar*I-G_2G_3 FAAF value for the *E. coli* system **(a)** or the *Nar*I-G_1 FAAF value for the human system **(b)**. Quantification of NER efficiencies was from at least three independent experiments.

On the contrary, adducts that destabilize and disorder the DNA are the best substrates for the *E. coli* system (Table 1) but are poorer substrates in the human system. This leads to the assertion that there are factors other than damage recognition that influences activation of NER in the human system.

Table 1.

	$\Delta\Delta H$	$\Delta\Delta G_{37^\circ C}$	ΔT_m
	kcal/mol	**kcal/mol**	**°C**
NarI-G_1-FAAF	21.9	3.7	-5.3
NarI-G_2-FAAF	18.6	4.1	-7.9
NarI-G_3-FAAF	24.7	4.7	-8.3
NarI-G_1G_3-FAAF	24.1	6.8	-14.1
NarI-G_2G_3-FAAF	23.3	7.8	-17.9
NarI-G_1G_2-FAAF	6.2	4.3	-10.0

Table 1. Thermal and thermodynamic parameters of mono- and di-FAAF modified *Nar*I duplexes. Comparative thermodynamic parameters are listed for the FAAF-modified substrates. This is a summary of previously reported data for mono- and di-FAAF substrates [38]. The average standard deviations for −ΔΔH, −ΔΔG, and ΔT_m are ±3.0, ±0.4, and ±4.0, respectively [38].
ΔΔH = ΔH(modified duplex) – ΔH (control duplex).
ΔΔS = ΔS (modified duplex) – ΔS (control duplex).
ΔΔG = ΔG (modified duplex) – ΔG (control duplex).
$\Delta T_m = T_m$ (modified duplex) – T_m (control duplex).

XPC binding to AAF adducts in the *Nar*I sequence context

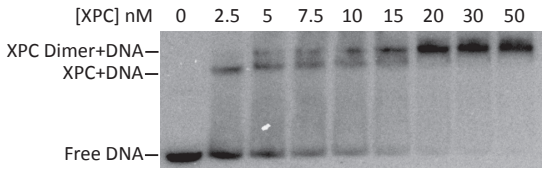
The XPC protein complex is widely believed to be involved in DNA damage recognition in the human NER pathway, much like the role of UvrA in *E. coli* [9]. To

determine the K_D, k_a, and k_d of the XPC interaction with the FAAF adducts we employed SPR molecular interaction analysis. By real-time monitoring of XPC binding a more informed conclusion can be drawn of its binding to and dissociation from adducted DNA. First, as a control, a traditional method, EMSA, was employed to demonstrate complex formation. Fig. 4a confirms previous reports that XPC can bind an adducted oligonucleotide as a multimer at increasing concentrations, as shown by the higher of the two slower migrating bands. This EMSA result is of XPC binding to the G_1 adduct and is representative of the results obtained for all other adducted substrates (data not shown).

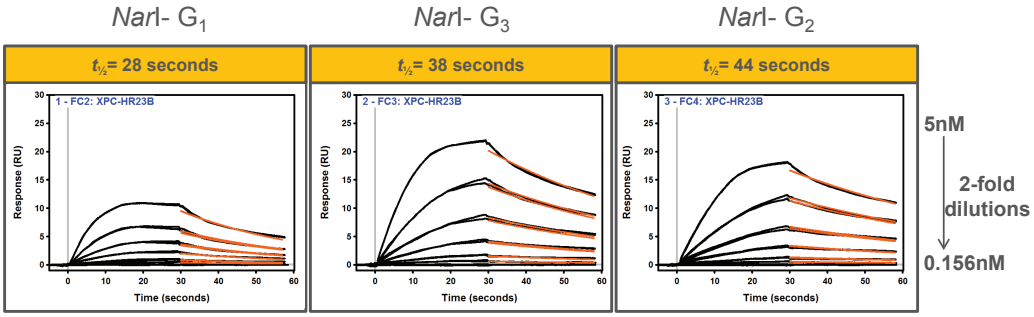
Next, mono- and di-FAAF-adducted DNA was used to determine XPC and UvrA association and dissociation rates by SPR. The SPR binding results are shown in Fig. 4 and Supplementary Fig. 1, in which the average of the triplicate data is displayed. Use of the usual simple 1:1 Langmuir-type binding model did not produce desirable closeness of fit between the protein and modified DNA (Supplementary Fig. 1). The fitting for XPC binding was improved somewhat (for di-adducts especially) by applying the 'heterogeneous ligand' model which is designed to accommodate the interaction of one protein analyte with two ligand sites on the surface. The usual remedy such as reducing the magnitude of the concentration gradient did not significantly mitigate the lack of ideal curve fittings, nor did decreasing the surface density of the immobilized DNA construct on the surface or increasing the flow rate. It is clear that the adduct-induced conformational heterogeneity for our substrates contributes many non-ideal molecular complexities, e.g., molecular diffusion affects, analyte heterogeneity [55,56], steric hindrance, and protein-DNA binding cooperativity and stoichiometry (e.g., see above Fig. 4a).

Figure 4

a XPC EMSA



b XPC:Mono-adducts



c XPC:Di-adducts

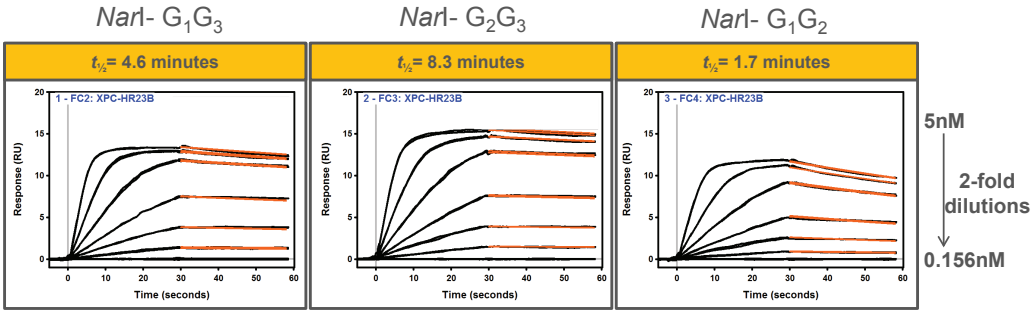


Figure 4. XPC binding to damaged DNA in the *Nar*I sequence. **(a)** Representative image of XPC binding to *Nar*I-G_1 in an EMSA assay. XPC protein, at increasing concentration, was incubated with a FAAF-damaged 55-bp oligo. Sensograms showing XPC binding kinetics to mono-FAAF **(b)** and to di-FAAF adducted substrates **(c)**. SPR responses were recorded for the binding of XPC NER protein (5, 2.5, 1.25, 0.625, 0.313, and 0.156 nM) to FAAF-modified full DNA duplexes. The recorded data are displayed as black lines while red lines represent curve fitting. The half-life ($t_{1/2}$) is indicated above the curves (in yellow box) and is defined as the time it takes for half of the XPC-DNA complex to dissociate. The fitted curves obtained from fittings using a one-independent site model ("Scrubber") are displayed (See Methods).

These factors affected the kinetic rate constants, which made them less reliable. However, we observed a dramatic difference in equilibrium dissociation constant (K_D) values with XPC and association and dissociation rates between mono- and di-adducts (Table 2). A similar trend was observed with UvrA (Supplementary Fig. 2), but here we will focus on the interactions between XPC and DNA containing FAAF-adducts.

Table 2.

FAAF on NarI sequence	SPR (K_D) (nM)	SPR (k_a) ($M^{-1}s^{-1}$)	SPR (k_d) (s^{-1})	t1/2 (s)	T_m(°C)(ΔT_m)	hNER
CCG*1G2CG3CC (G_1-mono)	2.9 (± 0.02) X 10^{-9}	2.3 (± 0.01) X 10^{7}	2.0 (± 0.02) X 10^{-2}	28 (± 0.5)	68.6 (-5.3)	1 (± 0.02)
CCG1G*2CG3CC (G_2-mono)	1.0 (± 0.01) X 10^{-9}	4.0 (± 0.17) X 10^{7}	3.8 (± 0.16) X 10^{-2}	43 (± 0.4)	66.0 (-7.9)	0.69 (± 0.03)
CCG1G2CG*3CC (G_3-mono)	1.3 (± 0.01) X 10^{-9}	2.4 (± 0.07) X 10^{7}	2.4 (± 0.08) X 10^{-2}	38 (± 0.1)	65.6 (-8.3)	0.65 (± 0.01)
CCG*1G*2CG3CC (G_1G_2-di)	7.6 (±0.26) X 10^{-11}	1.0 (±0.008) X 10^{8}	2.5 (± 0.01) X 10^{-3}	103 (± 1.3)	60.6 (-10)	0.69 (± 0.01)
CCG1G*2CG*3CC (G_2G_3-di)	4.2 (±0.34) X 10^{-12}	1.4 (±0.01) X 10^{8}	1.4 (± 0.01) X 10^{-3}	495 (± 4.5)	52.7 (-17.9)	0.12 (± 0.04)
CCG*1G2CG*3CC (G_1G_3-di)	1.6 (±0.76) X 10^{-11}	1.4 (±0.008) X 10^{8}	6.7 (± 0.05) X 10^{-3}	277 (± 2.9)	56.5 (-14.1)	0.30 (± 0.02)

Table 2. Correlation of XPC-RAD23B binding and dissociation parameters, melting temperature, and hNER efficiencies of FAAF-modified *Nar*I substrates. This is a comparison of binding kinetics and hNER efficiency. There is an inverse relationship between off-rate kinetics and human NER of the di-adducted dG-FAAF substrates. SPR (k_a), SPR (k_d) and SPR (K_D) are the association rate (k_a), dissociation rate (k_d) and equilibrium dissociation constant (K_D) values determined by SPR analysis of the interaction of XPC with mono- and di-adducts and t1/2 (s) is the calculated half-life of the protein-DNA complex. The T_m(ºC)ΔT_m data are the thermodynamic stability as previously reported [33,38]. The hNER efficiency is relative to the data displayed in Fig. 2.

As shown in Fig. 4, the binding of XPC to lesions in the *Nar*I sequence is more stable with the di-FAAF adducts than the mono-FAAF adducts. It is of note that our SPR binding results are similar to previous work demonstrating the binding affinity of XPC to cisplatin-damaged DNA [57,58]. Note also that the thermostability of these adducts varies and that the di-adducts have the lowest thermostability ($G_1 > G_2 > G_3 >> G_1G_2 > G_1G_3 > G_2G_3$) (Table 1). In addition, di-adducts bound with significantly slower dissociation (off) rates relative to the mono-adducts. These results are consistent with conformational changes accompanying the protein-DNA interactions which could provide the basis for deciphering the binding/repair mechanisms. Comparing K_D values (Table 2), the di-FAAF adducted DNAs show more stable XPC binding than the mono-adducts. The tightest XPC binding was to G_2G_3 with a dissociation rate of $4.2x10^{-12}$. The low K_D for XPC di-adduct binding signifies a much tighter association of XPC protein to clustered lesions. Table 2 summarizes the findings of XPC binding compared to the thermodynamic stability, the XPC-DNA complex half-life, and the incision efficiencies of FAAF adducts in the human NER system.

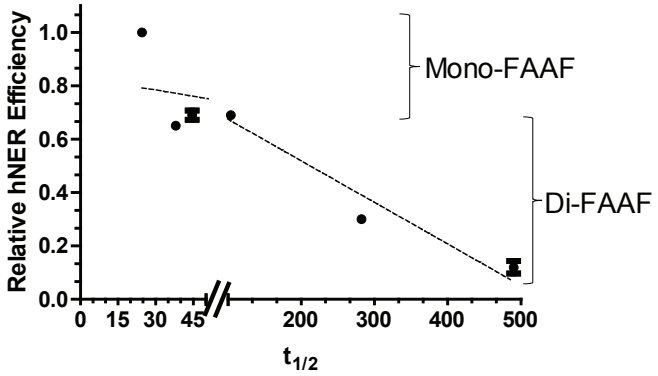


Figure 5. Comparison of hNER efficiency and half-life ($t_{1/2}$) of the XPC-DNA complex. The data points were analyzed independently as mono- or di-adduct groups and the two dashed lines indicates the group trends. Mono- and di-adducts are indicated on the right.

By calculating the half-life we observed that XPC-DNA complexes containing the di-FAAF adducts were on average at least 8 times more stable than complexes with the mono-FAAF adducts (Fig. 4b and c, Table 2). In the extreme case the half-life of XPC-DNA complexes containing *Nar*I di-G_2G_3-FAAF was 20 times longer than that of *Nar*I mono-G_1-FAAF. The half-life demonstrates an inverse relationship with the hNER efficiency (Fig. 5). This indicates that the longer XPC stays bound to the adduct site, the less productive the human NER process. These data are provocative as they suggest that a strong DNA damage recognition or binding itself does not necessarily guarantee efficient NER.

Discussion

Various studies have attempted to relate the protein binding interactions involved in DNA damage recognition with NER excision efficiency utilizing a wide array of damage to DNA [26,59]. This work proposes a novel mechanism beyond the conventional concept that the binding capability of DNA damage recognition proteins directly relates to the ability of the repair process to remove the damage from DNA. The presented SPR/hNER results suggest strongly that di-FAAF adducts fail to produce a productive complex for hNER even though the damage recognition binding is strong. In other words, robust XPC binding (K_D of $10^{-9} \sim 10^{-10}$ M) may be required for initiation of hNER. However, unusually strong XPC binding ($K_D < 10^{-11}$) and, more importantly, the extremely slow dissociation of di-FAAF adducts, and thus the long residence time ($t_{1/2}$) of XPC at the damage site, could be detrimental to recruitment of subsequent downstream proteins to complete the hNER process (Fig. 5). We demonstrate that in addition to the equilibrium binding affinity of XPC for DNA damage, kinetics and the off-rate of the interaction also play critical roles in

determining the NER efficiency. This is particularly true for certain types of DNA adducts, such as the di-FAAF examined here, which have a long XPC residence time during DNA damage recognition. Since dissociation of XPC from the damage site after initial recognition is necessary for subsequent binding of other repair factors in the mechanism of NER, it is possible that such an extended residence of XPC at the damage site would likely make the dissociation the rate-limiting step of NER. This could lead to inefficient DNA repair even though the XPC-damaged DNA binding affinity is high. Our findings may help us better understand the complex mechanisms relating protein binding and adduct clearance via NER.

For the mono-FAAF adducts we observed that the repair efficiency of the G_1 adduct was significantly higher that the G_2 and G_3 adducts, in agreement with the shorter half-life of XPC on DNA adducted at the G_1 rather than the G_2 or G_3 positions (Table 2). This contrasts with different reported preferences for repair in HeLa cell extracts of adducts at the G_2 position [27,28] or at the G_3 position [25]. These differences in NER preferences may stem from the nature of the DNA constructs used in the respective studies. The long-range sequence context of the DNA in which the adducted *Nar*I site embedded differs significantly between our construct and that of Yeo *et al*. as do the plasmids carrying these constructs (pTZ19U vs. pBluescript II SK+, respectively) [25]. In addition, the studies reporting a preference for repair of adducts at the G_2 position employed relatively short (<150 bp), linear, and internally labeled constructs that may display different torsional stresses on the DNA double helix in the HeLa extracts. Most studies on sequence context on NER efficiencies focus on the position effects within the *Nar*I sequence rather than the influence of the long-range sequence in which it is embedded or the effect of super-helical stresses

within the circular plasmids. Further studies are needed to resolve the influence of these experimental parameters on the NER efficiencies. However, the SPR, thermodynamic, EMSA and repair efficiency data reported here are internally consistent in highlighting the importance of the half-life of the XPC-DNA complex in determining NER efficiency.

Our finding that the increased residence time of XPC on damaged DNA containing di-adducts reduces repair efficiency is consistent with the *in vivo* studies of NER protein binding dynamics to chromatin and at damaged DNA sites, especially for XPC [60]. However, we analyzed reconstituted systems of short adducted dsDNAs of non-UV induced DNA damage plus purified XPC-RAD23B or adducted plasmid DNA plus a fractionated extract from HeLa cells not exposed to UV. Thus, our data may not directly correlate to cellular studies showing an observed influence of DDB2 (damaged DNA-binding 2) protein. This is also possibly true for the effects of the ubiquitin-dependent p97 segregase and centrin-2 on XPC binding/release dynamics from UV-damaged chromatin DNA [61]. Further studies are needed to test how such cellular factors might reduce the binding and retention time of XPC on DNA containing mono- and di-FAAF adducts, and to extend these analyses to UV-induced damage. However, our data are consistent with their observation that the prolonged retention of XPC in chromatin or an increase $t_{1/2}$ reduces the NER efficiency [61], though the prolonged retention in the previous UV-induced damage study was due to the deficiency of p97 segregase.

The *E. coli* NER system has been studied for decades, leading to many breakthroughs in our understanding of how cells can repair DNA damage. These studies have provided vital information that is applicable to our understanding of the human NER system; however, these two systems are not directly comparable when considering damage

recognition. Tight binding of recognition proteins may imply a better incision substrate, and this is true for *E. coli*, but in the human system tight binding can lead to a longer residence time, which results in a decrease in substrate processing. Analysis of previously reported structures of XPC- and $UvrA_2B$ (or $UvrA_2B_2$)-damaged DNA interactions may provide some understanding of our observed differences of the FAAF adduct incisions between hNER and UvrABC systems [62-64]. In both *E. coli* and human systems a damage recognition protein is required to initiate the NER. $UvrA_2B$ and XPC-RAD23B, respectively, fill this role in the GGR sub-pathway of NER. Interestingly, although both protein complexes recognize DNA damage, dissociation of these proteins from the damage site after recognition is quite different. In the case of UvrABC, $UvrA_2$ dissociates, while UvrB remains bound to the damage. In contrast, in hNER, RAD23B dissociates shortly after binding, however, XPC remains bound [65]. Interestingly, structural evidence revealed that $UvrA_2$ makes contact with the DNA that flanks the damage site and has no contact with the lesion itself [64], while UvrB establishes lesion contact, utilizing a β-hairpin domain to insert into the DNA strands and flip-out bases opposing the lesion [63,66]. On the other hand, XPC appears to be responsible for both roles carried out by $UvrA_2$ and UvrB of the $UvrA_2B$ complex since XPC also inserts its β-hairpin domain between DNA strands at the damage site [23,24]. The β-hairpin insertion is likely to make protein-DNA interaction more stable and, thus, the protein less likely to dissociate from DNA. Since $UvrA_2$ does not carry out β-hairpin insertion while XPC does, it is possible that $UvrA_2$ would be energetically easier to dissociate from DNA than XPC whose dissociation is more damage-type dependent. Thus, the type of di-FAAF adduct may increase the affinity of XPC at the damage site, leading to a longer residence time. For mono-FAAF adducts, in contrast, a normal $t_{1/2}$ keeps the

dissociation from being a rate-limiting step, remaining close to the regular binding-repair efficiency correlation. Our data supports the notion that XPC has an increased residence time at clustered lesion sites, making the dissociation from the lesion the rate-limiting step. Fig. 6b shows a 'hypothetical' 3D model of the G_1 mono-FAAF adduct that was printed (not simulated), derived from the published Rad4-CPD DNA structure (PDB ID# 2QSG) [24]. The G_1 mono-FAAF and G_2G_3 di-FAAF (not shown) adducts exhibit distinctive differences in hNER (1 *vs*. 0.12 relative efficiencies) (Fig. 3b) and XPC residence time (28 *vs*. 495 s) (Table 2), as well as DNA thermal stability (D*T*m -5.3 *vs*. -17.9°C) (Table 1). The CPD mismatch site was simply replaced by FAAF (yellow,*) for visualization of the G_1 mono-FAAF. The mono-lesion is expected to produce a similar NER complex as the CPD; that is, Rad4 inserts a b-hairpin into the damaged site with most of the modified strand being fully exposed and the mismatched bases (cyan) on the complimentary strand flipped out the double helix. The di-G_2G_3 adduct, however, contains two FAAF adducts separated by one nucleotide base and is likely to exhibit very different conformations that are responsible for its unusually strong binding and slow dissociation (Table 2). Possibilities include additional DNA interactions with the BHD3/BHD2 and TGD protein segments; the latter have been considered to be responsible for the highly kinked DNA conformations. These additional XPC-lesion interactions could provoke conformational alterations, which may affect the logistics of the subsequent verification step. One potential future study is to crystalize Rad4 or XPC complexed with the afore-mentioned mono-G_1 and di-G_2G_3-FAAF adducts used in this study.

Figure 6

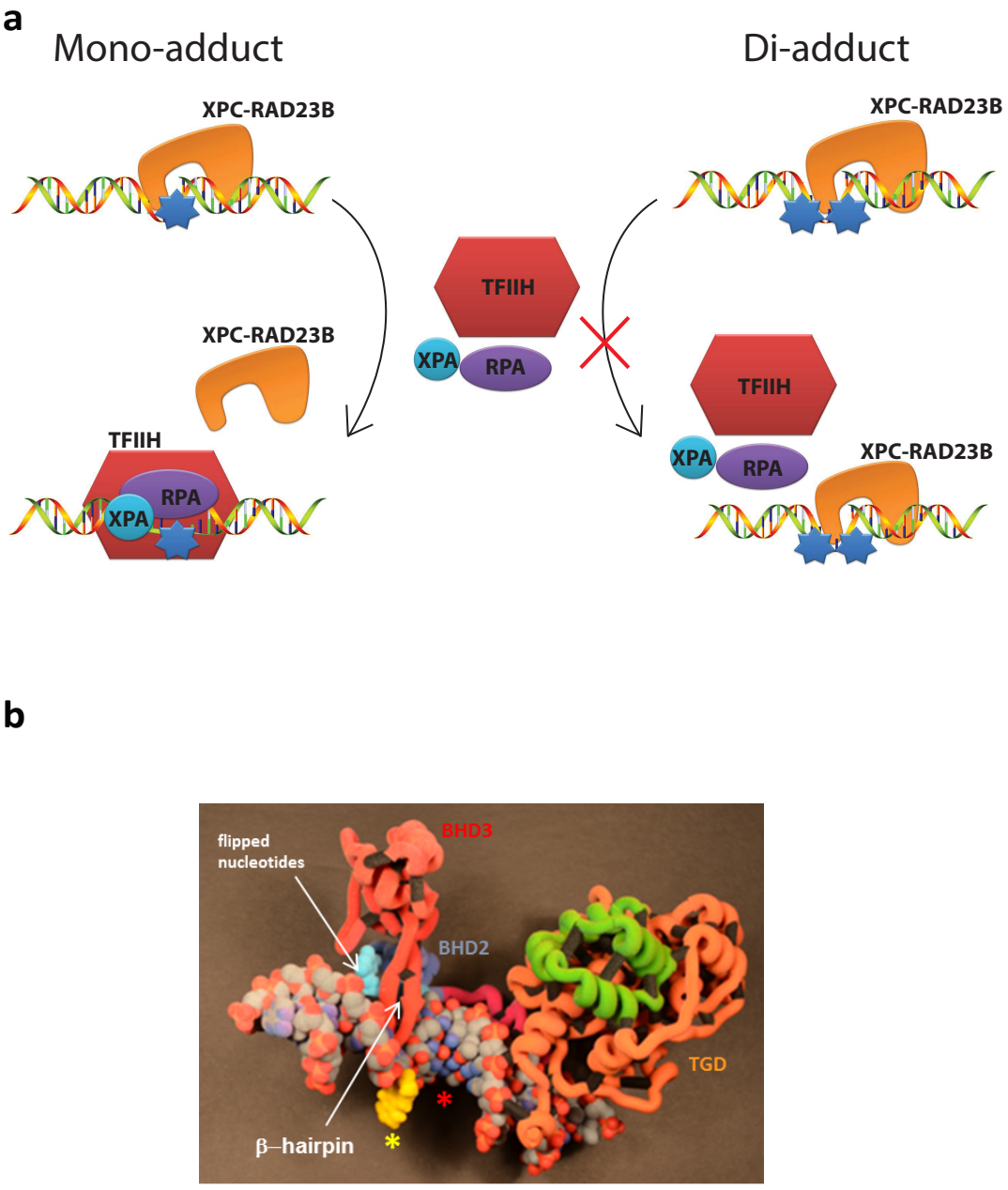


Figure 6. Proposed model for XPC interaction with DNA-adduct site. **(a)** A schematic illustrating the proposed mechanism of action where XPC is loosely bound to mono-FAAF adducted DNA (left) or tightly bound to di-FAAF adducted DNA (right). Following dissociation of XPC from the damage site subsequent NER factors are recruited to complete excision of the damaged base; however, in the di-adduct situation XPC is retained on the damaged DNA, delaying successful NER completion. **(b)** 3D-printed 'hypothetical' model (not simulated) to illustrate the potential binding of the yeast Rad4/Rad23 to the mono-G_1-FAAF duplex based on PDB ID 2QSG. The β-hairpin domains (BHD2 and BHD3) and the transglutaminase-homology domain (TGD), which are involved in protein-DNA interaction, are indicated. The domains were adapted from previous crystal structure analysis by Min *et al.* The duplex sequence used in this model is identical with that of Min's crystal work except that the CPD lesion was replaced by FAAF-G_1 (yellow *, as shown). The insertion of the BHD3 β-hairpin was accompanied with flipping of the mismatched bases (cyan) on the complimentary sequence.

If successful, the results are likely to provide valuable structural insights on XPC-DNA interactions that contribute to the large discrepancy in their residence time and reveal important clues regarding the structural requirements for recruitment of other NER proteins and subsequent lesion verification.

Tight binding during damage recognition may imply a better substrate for incisions, and this is true for *E. coli* NER, but in the human system binding too tightly can lead to a longer residence time, which decreases DNA repair (Fig. 6a). In addition to the new insights into understanding the mechanisms of hNER, the inverse relationship between the $t_{1/2}$ of tight binding and NER efficiency suggests a novel strategy as a new therapeutic approach in cancer therapy. Given that DNA damaging agents have been widely exploited for anticancer activities, targeting properly spaced di-adducts or a cluster-like drug that effectively stalls XPC or other damage recognition proteins could lead to strong resistance to repair and, thus, a higher efficiency in killing cancer cells. Recent studies have introduced new models that utilize residence time in drug design and to increase the efficacy of known drugs [67,68]. These models could be applied to our system allowing for design of novel drugs that could increase the residence time of XPC on damaged DNA for cancer therapeutics.

In summary, the present study suggests that dissociation of XPC from adducted DNA is the determining factor for successful NER elimination of adducts. In recognizing the types of adducts that are comparable to *Nar*1-G_2G_3 and *Nar*I-G_1G_3 cluster adducts, XPC can be stalled on these damage sites, preventing clearance of induced adducts. Exploiting the high proliferative rate of cancer cells and the slow dissociation rate of XPC from clustered adducts allows for a more efficiently targeted approach to cancer therapy. The current work also advances our understanding of the intricacies of the NER mechanism.

References

1. Melchior WB, Marques MM, Beland FA. Mutations induced by aromatic amine DNA adducts in pBR322. Carcinogenesis. 1994;15: 889–899.

2. Luch A. Nature and nurture - lessons from chemical carcinogenesis. Nat Rev Cancer. Nature Publishing Group; 2005;5: 113–125. doi:10.1038/nrc1546

3. Neumann H-G. Aromatic amines in experimental cancer research: tissue-specific effects, an old problem and new solutions. Crit Rev Toxicol. Informa UK Ltd UK; 2007;37: 211–236. doi:10.1080/10408440601028603

4. Poirier MC. Chemical-induced DNA damage and human cancer risk. Discov Med. 2012;14: 283–288.

5. Friedberg EC, Walker GC, Siede W, Wood RD. DNA Repair and Mutagenesis. American Society for Microbiology Press; 2005.

6. Lehmann AR. DNA repair-deficient diseases, xeroderma pigmentosum, Cockayne syndrome and trichothiodystrophy. Biochimie. 2003;85: 1101–1111.

7. Truglio JJ, Croteau DL, Van Houten B, Kisker C. Prokaryotic nucleotide excision repair: the UvrABC system. Chem Rev. American Chemical Society; 2006;106: 233–252. doi:10.1021/cr040471u

8. de Laat WL, Jaspers N. Molecular mechanism of nucleotide excision repair. Genes & …. 1999.

9. Gillet LCJ, Scharer OD. Molecular mechanisms of mammalian global genome nucleotide excision repair. Chem Rev. 2006;106: 253–276. doi:10.1021/cr040483f

10. Van Houten B. Nucleotide excision repair in Escherichia coli. Microbiol Rev. 1990;54: 18–51.

11. Zou Y, Van Houten B. Strand opening by the UvrA(2)B complex allows dynamic recognition of DNA damage. The EMBO Journal. 1999;18: 4889–4901. doi:10.1093/emboj/18.17.4889

12. Zou Y, Ma H, Minko IG, Shell SM, Yang Z, Qu Y, et al. DNA damage recognition of mutated forms of UvrB proteins in nucleotide excision repair. Biochemistry. 2004;43: 4196–4205. doi:10.1021/bi035992a

13. Zou Y, Luo C, Geacintov NE. Hierarchy of DNA Damage Recognition in EscherichiacoliNucleotide Excision Repair †. Biochemistry. 2001;40: 2923–2931. doi:10.1021/bi001504c

14. Zou Y, Bassett H, Walker R, Bishop A, Amin S, Geacintov NE, et al. Hydrophobic forces

dominate the thermodynamic characteristics of UvrA-DNA damage interactions. J Mol Biol. 1998;281: 107–119. doi:10.1006/jmbi.1998.1903

15. Sugasawa K, Ng JM, Masutani C, Iwai S, van der Spek PJ, Eker AP, et al. Xeroderma pigmentosum group C protein complex is the initiator of global genome nucleotide excision repair. Mol Cell. 1998;2: 223–232.

16. Scharer OD. Nucleotide excision repair in eukaryotes. Cold Spring Harb Perspect Biol. 2013;5: a012609–a012609. doi:10.1101/cshperspect.a012609

17. Ray A, Milum K, Battu A, Wani G, Wani AA. NER initiation factors, DDB2 and XPC, regulate UV radiation response by recruiting ATR and ATM kinases to DNA damage sites. DNA Repair (Amst). 2013;12: 273–283. doi:10.1016/j.dnarep.2013.01.003

18. Volker M, Moné MJ, Karmakar P, van Hoffen A, Schul W, Vermeulen W, et al. Sequential assembly of the nucleotide excision repair factors in vivo. Mol Cell. 2001;8: 213–224.

19. Wood RD. DNA damage recognition during nucleotide excision repair in mammalian cells. Biochimie. 1999;81: 39–44.

20. Hilton B, Shkriabai N, Musich PR, Kvaratskhelia M, Shell S, Zou Y. A new structural insight into XPA-DNA interactions. Biosci Rep. 2014;34: e00162–840. doi:10.1042/BSR20140158

21. Zou Y, Liu Y, Wu X, Shell SM. Functions of human replication protein A (RPA): from DNA replication to DNA damage and stress responses. J Cell Physiol. 2006;208: 267–273. doi:10.1002/jcp.20622

22. Missura M, Buterin T, Hindges R, Hübscher U, Kaspárková J, Brabec V, et al. Double-check probing of DNA bending and unwinding by XPA-RPA: an architectural function in DNA repair. The EMBO Journal. 2001;20: 3554–3564. doi:10.1093/emboj/20.13.3554

23. Chen X, Velmurugu Y, Zheng G, Park B, Shim Y, Kim Y, et al. Kinetic gating mechanism of DNA damage recognition by Rad4/XPC. Nat Commun. 2015;6: 5849. doi:10.1038/ncomms6849

24. Min J-H, Pavletich NP. Recognition of DNA damage by the Rad4 nucleotide excision repair protein. Nature. Nature Publishing Group; 2007;449: 570–575. doi:10.1038/nature06155

25. Yeo J-E, Khoo A, Fagbemi AF, Scharer OD. The efficiencies of damage recognition and excision correlate with duplex destabilization induced by acetylaminofluorene adducts in human nucleotide excision repair. Chem Res Toxicol. 2012;25: 2462–2468. doi:10.1021/tx3003033

26. Lee Y-C, Cai Y, Mu H, Broyde S, Amin S, Chen X, et al. The relationships between XPC binding to conformationally diverse DNA adducts and their excision by the human NER system: is there a correlation? DNA Repair (Amst). 2014;19: 55–63. doi:10.1016/j.dnarep.2014.03.026

27. Mu H, Kropachev K, Wang L, Zhang L, Kolbanovskiy A, Kolbanovskiy M, et al. Nucleotide excision repair of 2-acetylaminofluorene- and 2-aminofluorene-(C8)-guanine adducts: molecular dynamics simulations elucidate how lesion structure and base sequence context impact repair efficiencies. Nucleic Acids Res. Oxford University Press; 2012;40: 9675–9690. doi:10.1093/nar/gks788

28. Mu D, Bertrand-Burggraf E, Huang JC, Fuchs RP, Sancar A, Fuchs BP. Human and E.coli excinucleases are affected differently by the sequence context of acetylaminofluorene-guanine adduct. Nucleic Acids Res. 1994;22: 4869–4871.

29. Anderson KE, Hammons GJ, Kadlubar FF, Potter JD, Kaderlik KR, Ilett KF, et al. Metabolic activation of aromatic amines by human pancreas. Carcinogenesis. 1997;18: 1085–1092.

30. Juricek L, Bui L-C, Busi F, Pierre S, Guyot E, Lamouri A, et al. Activation of the aryl hydrocarbon receptor by carcinogenic aromatic amines and modulatory effects of their N-acetylated metabolites. Arch Toxicol. Springer Berlin Heidelberg; 2014;: 1–10. doi:10.1007/s00204-014-1367-7

31. Heflich RH, Neft RE. Genetic toxicity of 2-acetylaminofluorene, 2-aminofluorene and some of their metabolites and model metabolites. Mutation Research/Reviews in Genetic Toxicology. 1994;318: 73–174. doi:10.1016/0165-1110(94)90025-6

32. Cho BP, Zhou L. Probing the conformational heterogeneity of the acetylaminofluorene-modified 2'-deoxyguanosine and DNA by 19F NMR spectroscopy. Biochemistry. 1999;38: 7572–7583. doi:10.1021/bi990182d

33. Jain V, Hilton B, Patnaik S, Zou Y, Chiarelli MP, Cho BP. Conformational and thermodynamic properties modulate the nucleotide excision repair of 2-aminofluorene and 2-acetylaminofluorene dG adducts in the NarI sequence. Nucleic Acids Res. Oxford University Press; 2012;40: 3939–3951. doi:10.1093/nar/gkr1307

34. Meneni SR, Shell SM, Gao L, Jurecka P, Lee W, Sponer J, et al. Spectroscopic and theoretical insights into sequence effects of aminofluorene-induced conformational heterogeneity and nucleotide excision repair. Biochemistry. 2007;46: 11263–11278. doi:10.1021/bi700858s

35. Patnaik S, Cho BP. Structures of 2-acetylaminofluorene modified DNA revisited: insight into conformational heterogeneity. Chem Res Toxicol. 2010;23: 1650–1652. doi:10.1021/tx100341u

36. Kalam MA, Basu AK. Mutagenesis of 8-oxoguanine adjacent to an abasic site in simian

kidney cells: tandem mutations and enhancement of G-->T transversions. Chem Res Toxicol. 2005;18: 1187–1192. doi:10.1021/tx050119r

37. Shikazono N, Pearson C, O'Neill P, Thacker J. The roles of specific glycosylases in determining the mutagenic consequences of clustered DNA base damage. Nucleic Acids Res. 2006;34: 3722–3730. doi:10.1093/nar/gkl503

38. Jain V, Hilton B, Lin B, Jain A, Mackerell AD, Zou Y, et al. Structural and thermodynamic insight into Escherichia coli UvrABC-mediated incision of cluster diacetylaminofluorene adducts on the NarI sequence. Chem Res Toxicol. 2013;26: 1251–1262. doi:10.1021/tx400186v

39. Jain V, Hilton B, Lin B, Patnaik S, Liang F, Darian E, et al. Unusual sequence effects on nucleotide excision repair of arylamine lesions: DNA bending/distortion as a primary recognition factor. Nucleic Acids Res. 2013;41: 869–880. doi:10.1093/nar/gks1077

40. Mekhovich O, Tang MS, Romano LJ. Rate of incision of N-acetyl-2-aminofluorene and N-2-aminofluorene adducts by UvrABC nuclease is adduct- and sequence-specific: comparison of the rates of UvrABC nuclease incision and protein-DNA complex formation. Biochemistry. American Chemical Society; 1998;37: 571–579. doi:10.1021/bi971544p

41. Zou Y, Shell SM, Utzat CD, Luo C, Yang Z, Geacintov NE, et al. Effects of DNA adduct structure and sequence context on strand opening of repair intermediates and incision by UvrABC nuclease. Biochemistry. 2003;42: 12654–12661. doi:10.1021/bi034446e

42. Seeberg E, Fuchs RP. Acetylaminofluorene bound to different guanines of the sequence -GGCGCC- is excised with different efficiencies by the UvrABC excision nuclease in a pattern not correlated to the potency of mutation induction. Proc Natl Acad Sci USA. 1990;87: 191–194.

43. Liu Y, Reeves D, Kropachev K, Cai Y, Ding S, Kolbanovskiy M, et al. Probing for DNA damage with β-hairpins: Similarities in incision efficiencies of bulky DNA adducts by prokaryotic and human nucleotide excision repair systems in vitro. DNA Repair (Amst). 2011;10: 684–696. doi:10.1016/j.dnarep.2011.04.020

44. Shell SM, Hawkins EK, Tsai M-S, Hlaing AS, Rizzo CJ, Chazin WJ. Xeroderma pigmentosum complementation group C protein (XPC) serves as a general sensor of damaged DNA. DNA Repair (Amst). Elsevier B.V; 2013;: 1–7. doi:10.1016/j.dnarep.2013.08.013

45. Xu L, Vaidyanathan VG, Cho BP. Real-time surface plasmon resonance study of biomolecular interactions between polymerase and bulky mutagenic DNA lesions. Chem Res Toxicol. 2014;27: 1796–1807. doi:10.1021/tx500252z

46. Vaidyanathan VG, Xu L, Cho BP. Binding kinetics of DNA-protein interaction using

surface plasmon resonance. Nature Protocol Exchange. 2013. doi:10.1038/protex.2013.054

47. Vaidyanathan VG, Xu L, Cho BP. Binary and ternary binding affinities between exonuclease-deficient Klenow fragment (Kf-exo(-)) and various arylamine DNA lesions characterized by surface plasmon resonance. Chem Res Toxicol. American Chemical Society; 2012;25: 1568–1570. doi:10.1021/tx300289d

48. Reardon JT, Mu D, Sancar A. Overproduction, purification, and characterization of the XPC subunit of the human DNA repair excision nuclease. Journal of Biological Chemistry. 1996;271: 19451–19456.

49. Yang Z, Roginskaya M, Colis LC, Basu AK, Shell SM, Liu Y, et al. Specific and efficient binding of xeroderma pigmentosum complementation group A to double-strand/single-strand DNA junctions with 3“- and/or 5-”ssDNA branches. Biochemistry. 2006;45: 15921–15930. doi:10.1021/bi061626q

50. Shivji MK, Moggs JG, Kuraoka I, Wood RD. Dual-incision assays for nucleotide excision repair using DNA with a lesion at a specific site. Methods Mol Biol. 1999;113: 373–392. doi:10.1385/1-59259-675-4:373

51. O'Handley SF, Sanford DG, Xu R, Lester CC, Hingerty BE, Broyde S, et al. Structural characterization of an N-acetyl-2-aminofluorene (AAF) modified DNA oligomer by NMR, energy minimization, and molecular dynamics. 1993;32: 2481–2497. doi:10.1021/bi00061a005

52. Jain V, Vaidyanathan VG, Patnaik S, Gopal S, Cho BP. Conformational insights into the lesion and sequence effects for arylamine-induced translesion DNA synthesis: 19F NMR, surface plasmon resonance, and primer kinetic studies. Biochemistry. 2014;53: 4059–4071. doi:10.1021/bi5003212

53. Cho B. Structure–Function Characteristics of Aromatic Amine-DNA Adducts. The Chemical biology of DNA damage. 2010. doi:10.1002/9783527630110.ch10

54. Zou Y, Van Houten B, Farrell N. Sequence specificity of DNA-DNA interstrand cross-link formation by cisplatin and dinuclear platinum complexes. Biochemistry. 1994;33: 5404–5410.

55. Kortt AA, Gruen LC, Oddie GW. Influence of mass transfer and surface ligand heterogeneity on quantitative BIAcore binding data. Analysis of the interaction of NC10 Fab with an anti-idiotype Fab'. J Mol Recognit. John Wiley & Sons, Ltd; 1997;10: 148–158. doi:10.1002/(SICI)1099-1352(199705/06)10:3<148::AID-JMR360>3.0.CO;2-F

56. Cooper MA. Label-free screening of bio-molecular interactions. Anal Bioanal Chem. 2003;377: 834–842. doi:10.1007/s00216-003-2111-y

57. You J-S, Wang M, Lee S-H. Biochemical analysis of the damage recognition process in nucleotide excision repair. Journal of Biological Chemistry. American Society for Biochemistry and Molecular Biology; 2003;278: 7476–7485. doi:10.1074/jbc.M210603200

58. Trego KS, Turchi JJ. Pre-steady-state binding of damaged DNA by XPC-hHR23B reveals a kinetic mechanism for damage discrimination. Biochemistry. 2006;45: 1961–1969. doi:10.1021/bi051936t

59. Sugasawa K, Akagi J-I, Nishi R, Iwai S, Hanaoka F. Two-step recognition of DNA damage for mammalian nucleotide excision repair: Directional binding of the XPC complex and DNA strand scanning. Mol Cell. Elsevier; 2009;36: 642–653. doi:10.1016/j.molcel.2009.09.035

60. Luijsterburg MS, Bornstaedt von G, Gourdin AM, Politi AZ, Moné MJ, Warmerdam DO, et al. Stochastic and reversible assembly of a multiprotein DNA repair complex ensures accurate target site recognition and efficient repair. J Cell Biol. 2010;189: 445–463. doi:10.1083/jcb.200909175

61. Puumalainen M-R, Lessel D, Rüthemann P, Kaczmarek N, Bachmann K, Ramadan K, et al. Chromatin retention of DNA damage sensors DDB2 and XPC through loss of p97 segregase causes genotoxicity. Nat Commun. 2014;5: 3695. doi:10.1038/ncomms4695

62. Pakotiprapha D, Inuzuka Y, Bowman BR, Moolenaar GF, Goosen N, Jeruzalmi D, et al. Crystal structure of Bacillus stearothermophilus UvrA provides insight into ATP-modulated dimerization, UvrB interaction, and DNA binding. Mol Cell. 2008;29: 122–133. doi:10.1016/j.molcel.2007.10.026

63. Pakotiprapha D, Samuels M, Shen K, Hu JH, Jeruzalmi D. Structure and mechanism of the UvrA-UvrB DNA damage sensor. Nature Publishing Group. 2012;19: 291–298. doi:10.1038/nsmb.2240

64. Jaciuk M, Nowak E, Skowronek K, Tańska A, Nowotny M. Structure of UvrA nucleotide excision repair protein in complex with modified DNA. Nature Publishing Group. 2011;18: 191–197. doi:10.1038/nsmb.1973

65. Bergink S, Toussaint W, Luijsterburg MS, Dinant C, Alekseev S, Hoeijmakers JHJ, et al. Recognition of DNA damage by XPC coincides with disruption of the XPC-RAD23 complex. J Cell Biol. Rockefeller Univ Press; 2012;196: 681–688. doi:10.1083/jcb.201107050

66. Malta E, Moolenaar GF, Goosen N. Base flipping in nucleotide excision repair. Journal of Biological Chemistry. 2006;281: 2184–2194. doi:10.1074/jbc.M508901200

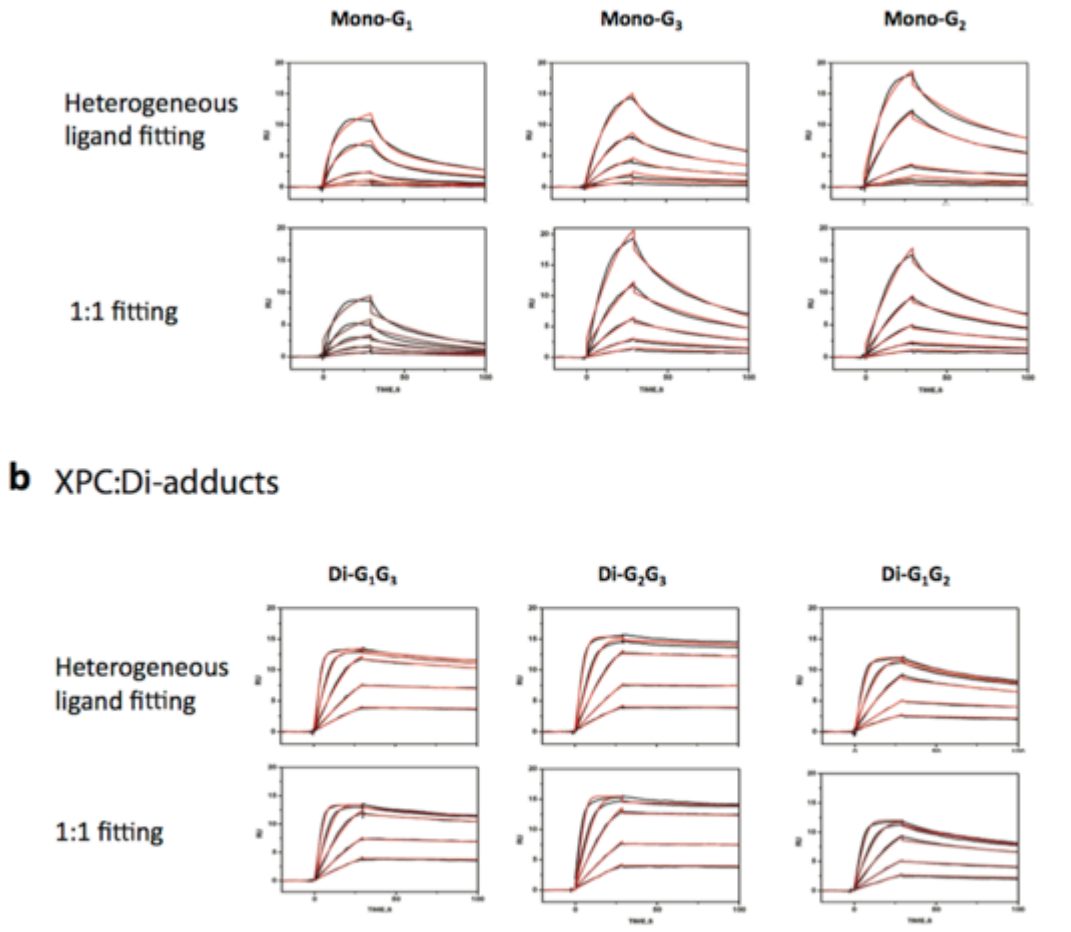
67. Bradshaw JM, McFarland JM, Paavilainen VO, Bisconte A, Tam D, Phan VT, et al. Prolonged and tunable residence time using reversible covalent kinase inhibitors.

Nat Chem Biol. 2015;11: 525–531. doi:10.1038/nchembio.1817

68. Walkup GK, You Z, Ross PL, Allen EKH, Daryaee F, Hale MR, et al. Translating slow-binding inhibition kinetics into cellular and in vivo effects. Nat Chem Biol. 2015;11: 416–423. doi:10.1038/nchembio.1796

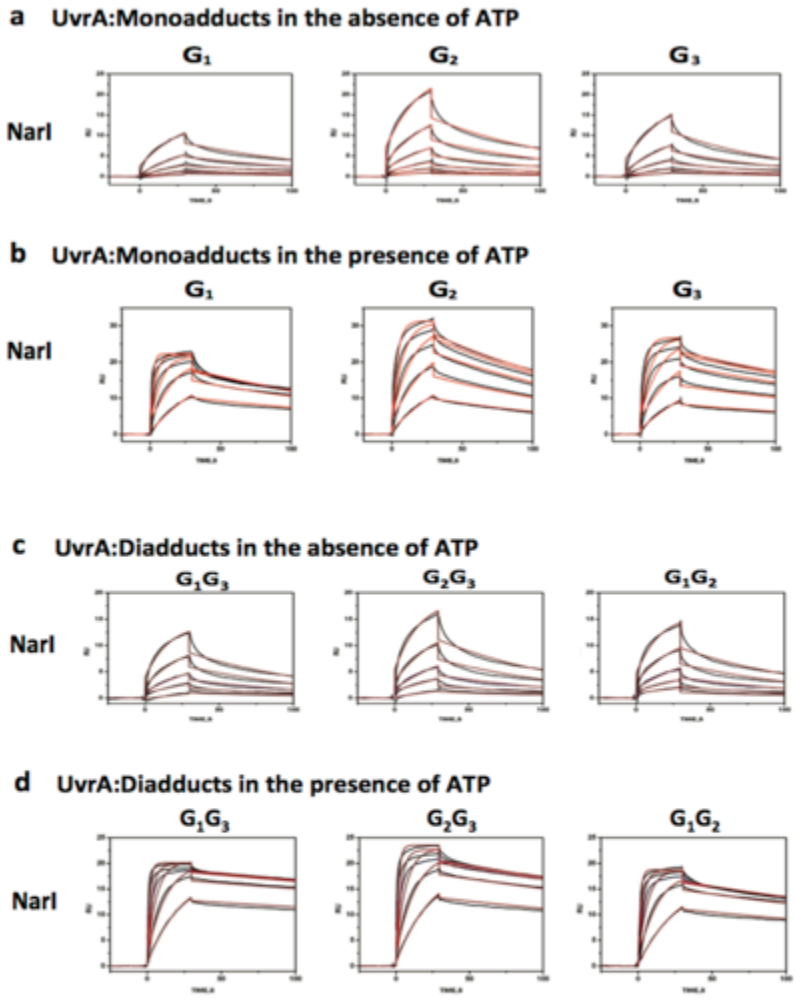
Supplemental Data

Supplementary Figure 1



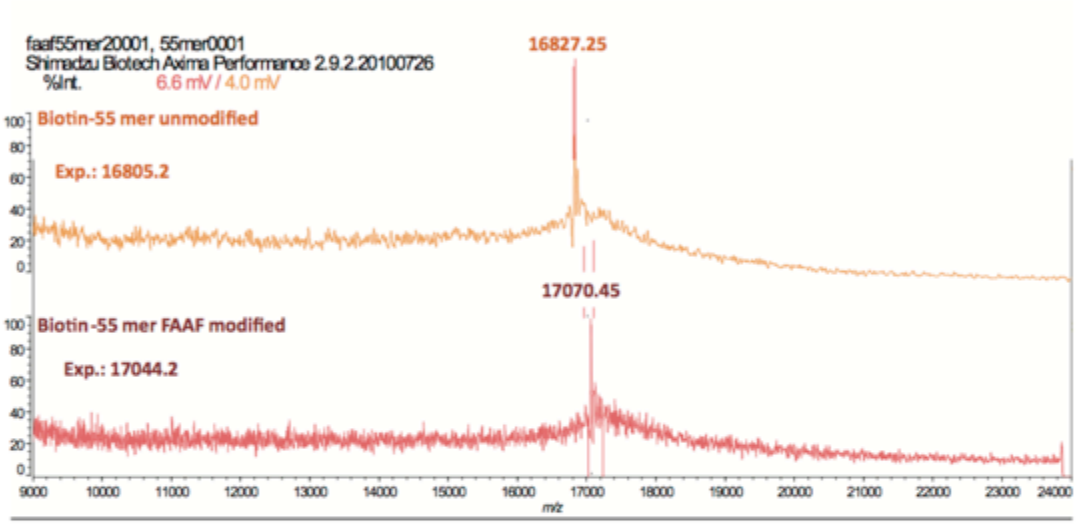
Representative SPR sensograms of mono- (**a**) and di-adducted (**b**) duplexes demonstrating the variation of curve fitting using the heterogeneous ligand fitting model (top) and the 1:1 Langmuir fitting model (bottom). The XPC protein concentrations used were 2.5, 1.25, 0.62, 0.31, and 0.15 nM.

Supplementary Figure 2



The effects of ATP on the UvrA binding to FAAF-damaged DNA in the NarI sequence. Sensograms showing UvrA binding kinetics to mono-adducted substrates in the absence of ATP (**a**) or in the presence of ATP (**b**) and to di-adducted substrates in the absence of ATP (**c**) or in the presence of ATP (**d**). SPR responses were recorded to of the binding of UvrA NER protein (250, 125, 62.5, 31.2, 15.6, and 7.8 nM) to modified full DNA duplexes. The recorded data are displayed as black lines while red lines represent curve fitting. The fitting curves obtained from fittings using a one-independent site model are displayed.

Supplementary Figure 3



Representative MALDI-TOF mass spectra of unmodified (orange) or FAAF-modified (red) substrates (55-mer).

Supplementary Table 1: The association rate (k_a), dissociation rate (k_d) and binding affinity constant (K_D) values of the interaction of XPC with FAAF-modified *Nar*I mono- and di-adducts.

		k_a (1/Ms)	k_d (1/s)	K_D (M)
Mono	G_1	2.25 (± 0.01) X10^{7}	2.00 (± 0.02) X10^{-2}	2.85 (± 0.02)X10^{-9}
	G_2	3.98 (± 0.17) X10^{7}	3.81 (± 0.16) X10^{-2}	1.04 (± 0.01) X10^{-9}
	G_3	2.37 (± 0.07) X10^{7}	2.38 (± 0.08) X10^{-2}	1.34 (± 0.01)X10^{-9}
Di	G_1G_3	1.43 (± 0.008) X10^{8}	6.7 (± 0.05) X 10^{-3}	1.61 (± 0.76) X10^{-11}
	G_2G_3	1.39 (± 0.01) X10^{8}	1.4 (± 0.01) X 10^{-3}	4.19 (± 0.34) X10^{-12}
	G_1G_2	1.00 (± 0.008) X10^{8}	2.5 (± 0.01) X 10^{-3}	7.56 (± 0.26) X10^{-11}

Supplementary Table 2: The association rate (k_a), dissociation rate (k_d) and binding affinity constant (K_D) values of the interaction of UvrA with FAAF-modified *Nar*I mono- and di-adducts in the absence of ATP.

		k_a (1/ms)	k_d (1/s)	K_D (M)
Mono	G_1	2.34(± 0.02) X10^{5}	0.99(± 0.006) X10^{-2}	4.34 (±0.03) X10^{-8}
	G_2	1.71(± 0.04) X10^{5}	1.05 (± 0.03) X10^{-2}	6.27 (±0.1) X10^{-8}
	G_3	2.30(± 0.05) X10^{5}	1.58(± 0.02) X10^{-2}	7.42 (±0.4) X10^{-8}
Di	G_1G_3	2.24 (± 0.04) X10^{5}	1.07(± 0.01) X10^{-2}	4.79 (±1.87) X10^{-8}
	G_2G_3	2.07 (± 0.05) X10^{5}	1. 06 (± 0.02) X10^{-2}	5.11(±1.66) X10^{-8}
	G_1G_2	2.45(± 0.03) X10^{5}	1.05 (± 0.09) X10^{-2}	4.67 (±1.89) X10^{-8}

Supplementary Table 3: The association rate (k_a), dissociation rate (k_d) and binding affinity constant (K_D) values of the interaction of UvrA with FAAF-modified *Nar*I mono- and di-adducts in the presence of ATP.

		k_a (1/ms)	k_d (1/s)	K_D (M)
Mono	G_1	4.93(± 0.05) X10^{6}	0.63(± 0.005) X10^{-2}	1.33 (±0.01) X10^{-9}
	G_2	1.80(± 0.005) X10^{6}	0.42 (± 0.002) X10^{-2}	2.43 (±0.04) X10^{-9}
	G_3	1.82(± 0.004) X10^{6}	0.68(± 0.002) X10^{-2}	3.94 (±0.05) X10^{-9}
Di	G_1G_3	5.17(± 0.01) X10^{6}	0.14 (± 0.01) X10^{-2}	2.73 (±0.46) X10^{-10}
	G_2G_3	4.76 (± 0.01) X10^{6}	0.24(± 0.02) X10^{-2}	5.03 (±1.33) X10^{-10}
	G_1G_2	6.23 (± 0.06) X10^{6}	0.30 (± 0.02) X10^{-2}	4.88 (±0.46) X10^{-10}

CHAPTER 5

A NOVEL STRUCTURAL INSIGHT INTO XPA-DNA INTERACTION

Abstract

Xeroderma pigmentosum group A (XPA) protein is an essential factor for nucleotide excision repair (NER) which is believed to be involved in DNA damage recognition/verification, NER nuclease recruiting, and stabilization of repair intermediates. Past studies on the structure of XPA have focused primarily on XPA interaction with damaged DNA. However, how XPA interacts with other DNA structures remains unknown though recent evidence suggest that these structures could be important for its roles in both NER and non-NER activities. Previously, we reported that the XPA monomer recognizes undamaged DNA double-strand/single-strand (ds-ssDNA) junctions containing ssDNA branches with a binding affinity much higher than its ability to bind to bulky base-adduct DNA damage. To understand how this interaction occurs biochemically we implemented a structural determination of the interaction using a mass spectrometry-based protein footprinting method and limited proteolysis. By monitoring surface accessibility of XPA lysines to NHS-biotin modification in the free protein and the DNA junction-bound complex we show that XPA physically interacts with the DNA junctions via two lysines, K168 and K179, located in the previously known XPA(98-219) DNA-binding domain (DBD). More importantly, we uncovered new lysine residues which residues were located outside of the known DBD also involved in formation of the nucleoprotein complex. We found that residues K221, K222, K224, K236 in the C-terminal domain participated in

DNA binding. Limited proteolysis analysis of XPA-DNA binding further confirmed the observation. Our results indicate that DNA binding domain of XPA should be expanded to include at least a portion of the C-terminal domain of XPA. A structural model of the redefined DNA binding domain of XPA has been generated and discussed.

Introduction

Human *xeroderma pigmentosum* group A (XPA) protein plays a critical role in both global genome repair (GGR) and transcription-coupled repair (TCR) subpathways of nucleotide excision repair (NER) [1-7]. The importance of XPA in NER is illuminated by evidence demonstrating that of all XP phenotypes, patients deficient in XPA are most severely affected, XPA-deficient mice show a complete deficiency in NER, and XPA-deficient cells are most sensitive to UV irradiation [8-11]. Also importantly, a growing body of evidence has link XPA to non-NER cellular processes. XPA has been showed to be involved in the mechanism of Hutchinson-Gilford progeria syndrome (HGPS) by binding to aberrant DNA replication forks in HGPS cells [12]; a recent finding suggests that XPA deficiency may lead to mitochondrial dysfunction [13].

XPA has long been thought to participate in DNA damage recognition in NER, working in cooperation with replication protein A (RPA) and the XPC-HR23B complex [2,4,5]. In the GGR pathway the XPC-HR23B complex is believed to be the factor that initially recognizes damage. In the TCR pathway the XPC-HR23B complex is not required and RNA polymerase is believed to initiate recognition, leaving the role of XPA unclear [14]. The involvement of XPA in both subpathways of NER and it being a limiting factor for NER and UV damage repair illustrate the importance of XPA in the NER machinery [15]. Taken

together, past studies imply that in addition to its DNA damage recognition role, XPA may have additional biochemical activities in NER and non-NER pathways.

Our previous work reported a potentially novel function of XPA in NER. We demonstrated that XPA has a much higher affinity for ds-ssDNA junctions with 3′- and/or 5′-ssDNA overhangs (including Y-structure junction and the junctions with either a 3′- or 5′-ssDNA branch) than for damaged dsDNA [16]. These are the same structures found as intermediates in NER. Furthermore, XPA recruits and anchors other NER proteins, especially the XPF-ERCC1 complex to repair sites for 5'-incision of damaged DNA [17-19]. The recruitment of the XPF-ERRC1 by XPA occurs near the ds-ssDNA junction of unwound DNA intermediate at the adduct [20]. Other studies confirm that XPA is found near ds-ssDNA junctions, initiating formation of pre-incision complexes [21]. The high binding affinity for junction structures provides an explanation for involvement of XPA in both subpathways of NER and for XPA remaining on the repair complex through the late stages of NER [9,22].

Junction structures also are present in many other DNA metabolic processes including non-NER DNA repair, replication and recombination. This implies that XPA may have additional roles in other cellular processes. An example is its involvement in HGPS via XPA-junction binding [12]. Also XPA was recently shown to interact with PCNA through its APIM domain [23]. These authors concluded that XPA is associated with the replisome in cells and is loaded onto newly synthesized DNA in undamaged cells. Therefore, XPA-junction interaction could play a role in this function. In addition, it has been previously reported that XPA is able to bind double-stranded three-way or four-way Holliday junctions which were used to mimic the helical kinks induced by DNA damage [24,25]. Together, these studies

demonstrate the important role of XPA-DNA junction interaction in DNA metabolisms. Nevertheless, the structural aspects of the XPA-junction binding have yet to be determined.

XPA contains three main domains: an RPA-interacting domain, a TFIIH-interacting domain, and a central or DNA-binding domain (DBD) [26]. The solution structures of the DBD domain of XPA were determined by NMR [27,28]. The interaction of XPA and its binding partners has been extensively studied; however, structures of the full-length XPA-DNA complex are unknown.

In the present study, we conducted a structural determination of the full-length XPA binding to DNA-junctions. By implementing a mass spectrometric protein footprinting method [29,30], specific protein-nucleic acid contacts of XPA-DNA interaction has been mapped. Then partial proteolysis coupled with protection through binding [31] were performed to confirm the identified contacts. We not only verify previously identified binding sites within the DNA-binding domain, but also and importantly, reveal new DNA-binding sites outside the known DNA binding domain. Together our data shed light on the solution structures of XPA-DNA complex with implication on the roles of XPA in NER and other non-NER cellular pathways.

Materials and Methods

Protein Purification and Substrate Labeling

Recombinant human XPA was expressed in insect cells and purified as previously described [32]. The protein concentration was determined by the Bradford assay (Bio-Rad). Oligonucleotides used to form the Y-shape, 5'-overhang and 3'-overhang substrates were

synthesized by Integrated DNA Technologies, then radiolabeled with [γ-32P]ATP (>5000 mCi/mmol; PerkinElmer) using T4-PNK (Ambicon), and unincorporated nucleotides were removed by a P6 spin column (Bio-Rad). The labeled substrate then was annealed to a complementary strand, followed by gel purification in an 8% native polyacrylamide gel.

Gel Mobility Shift Assays

Effects of the *N*-hydroxysuccinimidobiotin (NHS-biotin) (Pierce) modification on XPA-junction binding were determined by gel mobility shift assays. The labeled Y-shape, 5'-overhang, or 3'-overhang substrate (4 nM) was incubated with unmodified XPA in binding buffer (20 mM HEPES-KOH, pH 7.9, 75 mM KCl, 5 mM $MgCl_2$, 1 mM DTT, 5% glycerol, 1 mg/ml HSA) for 20 min at 30°C. The preformed complex then was exposed to modification with increasing concentrations NHS-biotin for another 15 min. The amino acid lysine was added to 10 mM to quench the reaction. In parallel experiments, XPA was modified first with NHS-biotin, then junction DNA was added to the reaction mix. Reactions were immediately loaded onto 3.5% native polyacrylamide gels and electrophoresed in TBE buffer at 4°C.

Biotin Modification and In-Gel Proteolysis

Purified XPA was modified with NHS-biotin in the presence and absence of each junction DNA. Typically, XPA (26.4 μM) was incubated with junction DNA (80 μM) in the binding buffer at room temperature for 30 min and then modified by adding NHS-biotin (400 μM of final concentration) for an additional 30 min incubation at room temperature.

Modifications were quenched by addition of 10 mM lysine, followed by separation of the XPA subunits by SDS-PAGE. The subunits were visualized by Coomassie Blue staining, excised from the gel, and extensively destained in 50% methanol/10% acetic acid. SDS was removed by washing the gel pieces with ammonium bicarbonate, dehydrated with 100% acetonitrile, and vacuum desiccated. Samples were digested with 1 μg of trypsin (Roche) in 50 mM ammonium bicarbonate overnight at room temperature. The supernatant then was recovered for MS analysis.

Mass Spectrometric Analysis

MS spectra were obtained using matrix-assisted laser desorption time-of-flight (MALDI- TOF) and quadrupole time-of-flight (Q-TOF) techniques. MALDI-TOF experiments were performed using a Kratos Axima-CRF instrument (Kratos Analytical Instruments) with an R-cyano-4-hydroxycinnamic acid matrix. MS analyses were performed on a Micromass Q-TOF-II instrument (Manchester, U.K.) equipped with an electrospray source and Micromass cap-LC. Peptides were separated with a Waters Symmetry 300 5 μm precolumn (Waters, Milford, MA) and a Micro-Tech Scientific (Vista, CA) ZC-10-C18SBWX-150 column using two sequential gradients of 5-40% acetonitrile for 35 min and 40-90% acetonitrile for 10 min. MS sequence data and the MASCOT automated peptide search engine (www.matrixscience.com) were used to identify XPA peptide peaks from the NCBInr primary sequence database, and matched peaks were then located in the primary MS spectra. Protection events were qualitatively assigned as the appearance of a peak corresponding to a biotin-modified peptide in the modified protein spectrum and the absence of the modification peak in the modified nucleoprotein complex spectrum. A

protection was considered to be significant only when the intensity of a modifiable peak was reduced by at least 85% in the nucleoprotein complex spectrum. To accurately identify protection events, at least two peaks that are not affected by procedures and present in all three spectra (unmodified protein, modified protein, and modified nucleoprotein complex) were used as control peaks. These control peaks served to standardize the peak intensities in each spectrum for accurate qualitative assignment of protection. Data were reproducibly compiled and analyzed from six independent experimental groups.

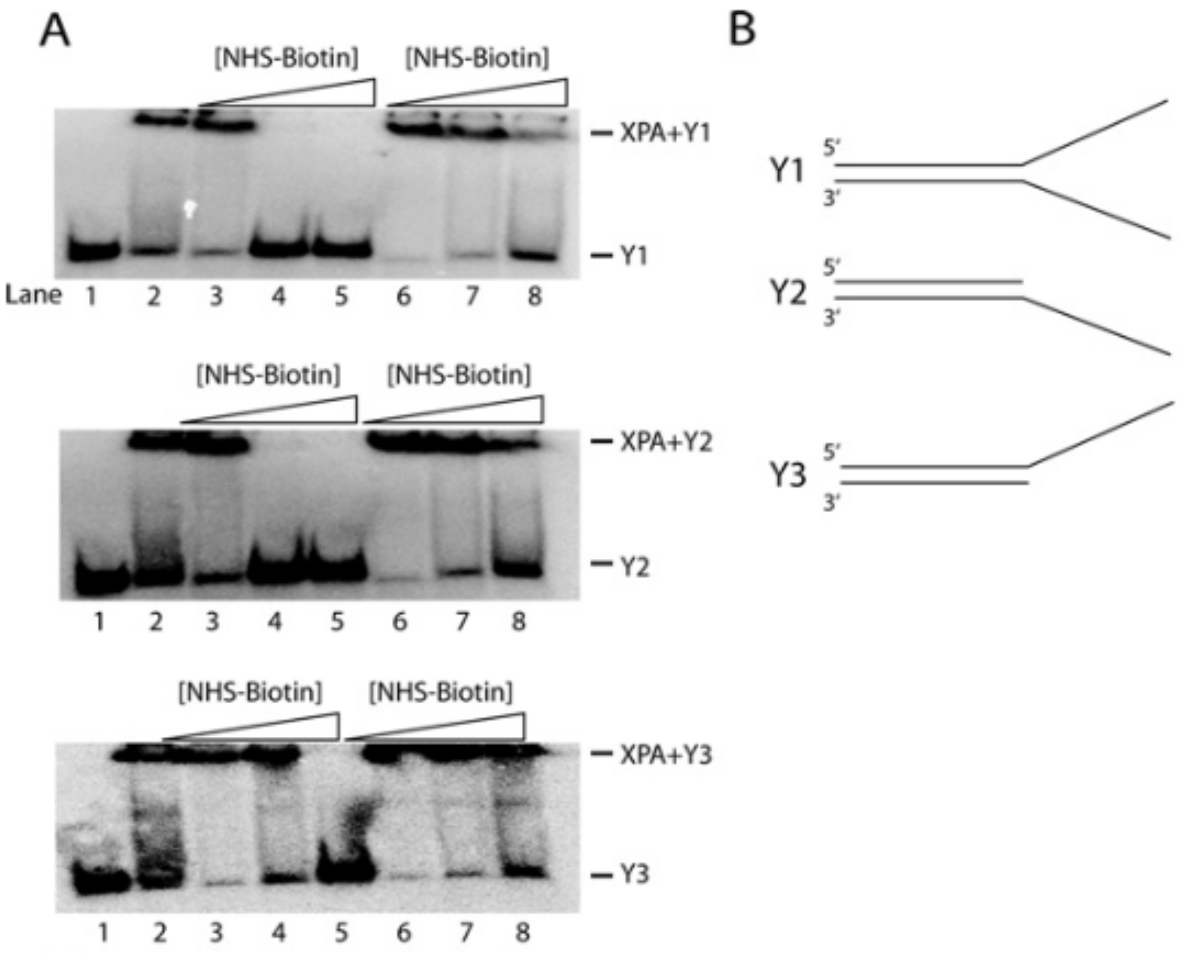
Partial Proteolysis and Identification of Proteolytic Fragments

4 μg of recombinant human XPA was incubated with chymotrypsin (1:80::Chymotrypsin:XPA) at 30°C for the indicated times in 20 μl XPA binding buffer (20 mM HEPES-KOH, pH 7.9, 75 mM KCl, 5 mM $MgCl_2$, 1 mM DTT, 5% glycerol, 1 mg/ml HSA). At each time point (5, 20, or 60 min), 5 μl of the reaction mixture was removed and terminated by the addition of Laemmli sample buffer and boiled for 5 min. The proteolytic products were resolved by 15% SDS-PAGE and stained with SYPRO Ruby protein gel stain (Bio-Rad). After destaining with a solution containing 10% methanol and 7% acetic acid, the gel was imaged in a phosphorimager (FLA-5000, FUJIFILM) using a 473-nm laser line. For automated N-terminal sequencing, the proteolytic fragments separated on the 15% SDS-PAGE were transferred to a polyvinylidene difluoride membrane followed by staining with Coomassie Brilliant Blue R-250. The protein bands of interest were excised and sent to the protein chemistry laboratory at Iowa State University for sequencing.

Results

Effect of lysine modification on XPA-DNA Interaction

We employed a protein footprinting [29,31-34] by which modification of solvent-accessible XPA lysine residues by the primary amine-specific reagent *N*-hydroxysuccinimidobiotin (NHS-biotin) is prevented at the XPA-DNA junction binding sites, while the surface lysines not involved in the binding are available for modification. NHS-biotin was used to insure minimal XPA-DNA junction complex disruption and maximum modification. And complex integrity and efficient modification by NHS-biotin must be confirmed. As shown in Figure 1A, gel mobility shift assays were implemented to demonstrate the effect of increasing concentration of NHS-biotin on the binding of XPA to DNA junction. XPA was incubated in the presence of NHS-biotin before or after addition of DNA junction substrates (lanes 3-5 and 6-8, respectively). The effect of modification on the binding was tested for three different types of DNA junction (Figure 1B). The data in Figure 1A showed that addition of increasing concentrations of NHS-biotin prior to nucleoprotein complex formation reduced and eventually eliminated the binding at 800 nM NHS-biotin (Figure 1A, lane 5) for all junction structures. This indicates that the lysine residues which can be modified by NHS-biotin are required for interaction of XPA with DNA junctions. Conversely, when NHS-biotin was added at increasing concentrations following nucleoprotein complex formation the binding was only slightly reduced (Figure 1A, lane 8). This indicates that the lysine residues



C

```
Y-Shape (Y1)
5' - GGA CCT GAA CAC GTA CGG AAT TCG ATT TCC TCG TGC CTG TTC TGC GCC TGC TGG C
3' - CCT GGA CTT GTG CAT GCC TTA AGC TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT T

5' Overhang (Y2)
5' - GCC AGC TGG CGC AGA TCT GGC TCG
3' - CGG TCG ACC GCG TCT AGA CCG AGC TCC TAG AGC TTA AGG CAT GCA CAA GTC CAG G

3' Overhang (Y3)
5' - GGA CCT GAA CAC GTA CGG AAT TCG AGA TCC TCG AGC CAG ATC TGC GCC AGC TGG C
3' - CCT GGA CTT GTG CAT GCC TTA AGC T
```

Figure 1. Effects of biotinylation on XPA–DNA junction binding (**A**): Lanes: 1, free junction DNA; 2, junction DNA–XPA complex; 3–5, junction DNA bound by XPA pretreated with 300, 500 and 800 μM NHS-biotin; 6–8, preformed XPA-junction DNA complex treated with 300, 500 and 800 μM NHS-biotin. Lanes 3–5 show that addition of increasing amounts of NHS-biotin prior to addition of junction DNA blocks critical lysine residues from junction DNA interaction, abolishing binding. Lanes 6–8 show formation of a XPA–junction DNA complex prior to adding NHS-biotin shielded critical lysine residues from subsequent modification and showed minor effects of biotin on the stability of the preformed XPA–junction DNA complex. Panel (**B**) shows substrate profiles used in all experiments: Y-shaped DNA (Y1); 5'-overhang DNA (Y2); 3'-overhang DNA (Y3). Panel (**C**) indicates the DNA sequence for each oligonucleotide.

involved in complex formation are shielded by the protein-DNA interaction. Based on these results a concentration of 500 nM NHS-biotin was chosen for the following studies as it is sufficient for modification of solvent-accessible lysine residues of XPA to prevent binding while having minimal effects on complex integrity. The three ds-ssDNA junction substrates (Figure 1B, C) yielded essentially identical results; therefore, all of the data presented hereafter is the representative data collected on substrate Y1.

MS Analysis of the XPA-DNA Complex

The NHS-biotin modified XPA, modified XPA-DNA junction complex, and unmodified XPA were subjected to SDS-PAGE prior to trypsin proteolysis. SDS-PAGE electrophoresis rendered the protein in a linear, denatured form that exposed all possible trypsin proteolytic sites, ensuring a complete and reproducible hydrolysis of the protein. The XPA lysines modified with NHS-biotin were detected by MS analysis of trypsin-generated peptide fragments. Figure 2A shows a representative MALDI-TOF spectrum for NHS-biotin-modified XPA with mass/charge peaks assigned to proteolysis-generated peptide fragments. This allowed for mono-isotopic resolution of the peaks reliably identifying the expected and biotin-modified tryptic fragments of XPA.

To focus on which lysines specifically interact with DNA in the nucleoprotein complex we analyzed the mass spectrum for *m/z* peaks that are present in the modified free protein but absent in both the modified nucleoprotein complex and the unmodified free protein spectra. These peaks corresponded to the mass of a trypsin-generated fragment plus the mass of the bound NHS-biotin molecules. The number of NHS-biotin molecules attached to any tryptic peptide was equal to the number of missed lysine

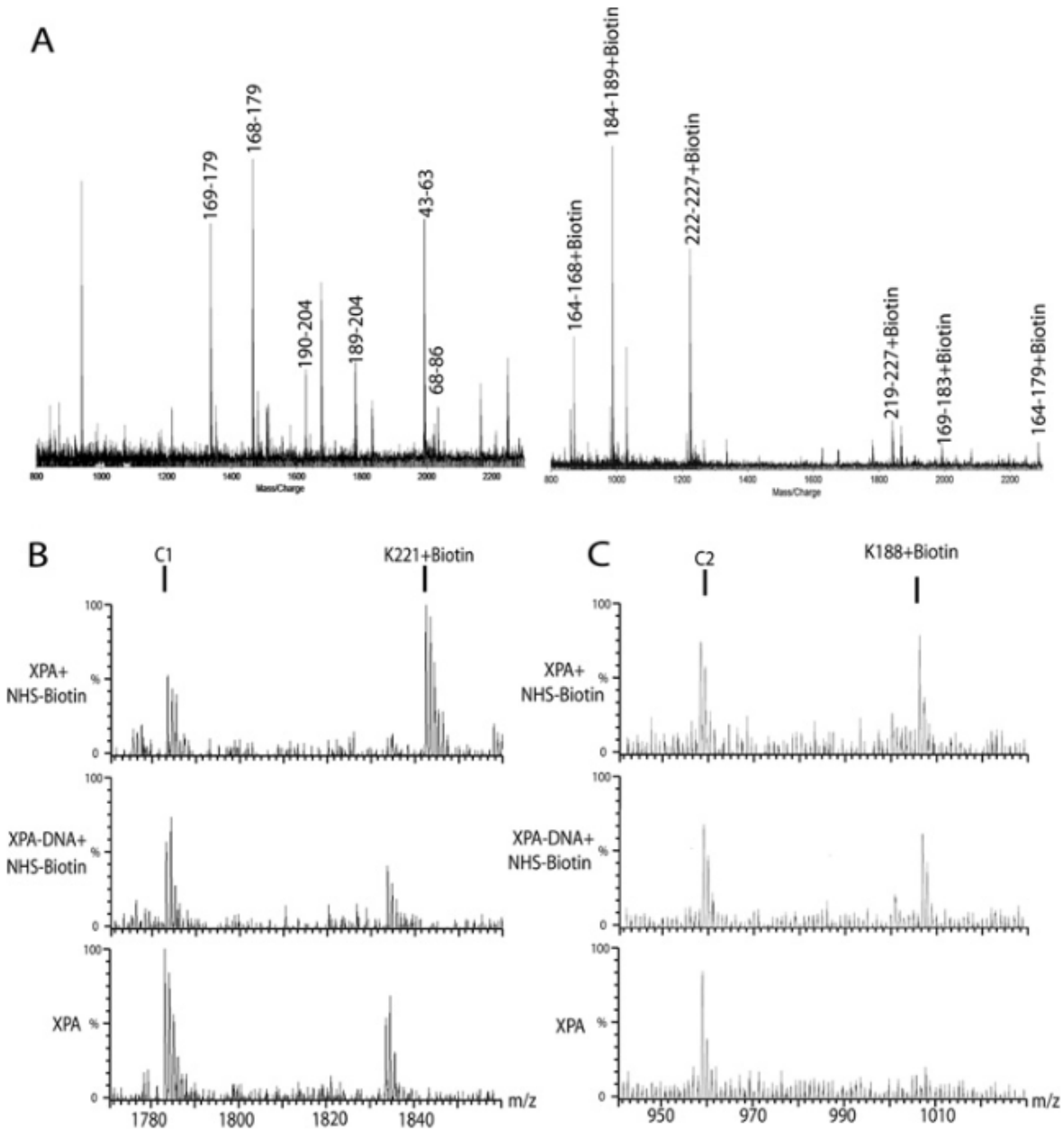
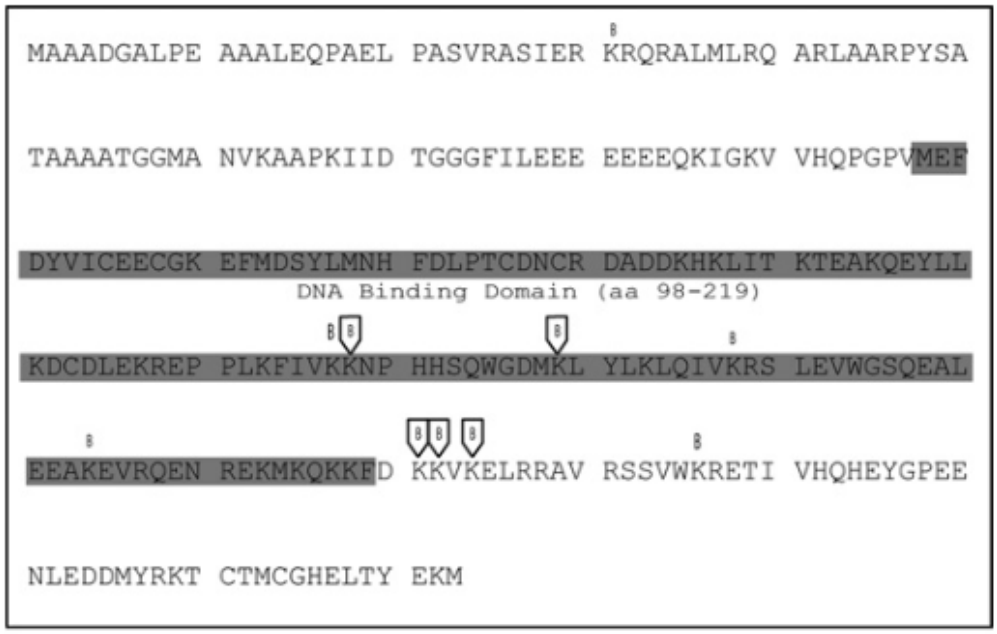


Figure 2. Identification of lysine residues involved in DNA binding (**A**) A typical MALDI-TOF spectrum of tryptic digests of XPA protein (left) or biotinylated XPA protein (right). Monoisotopic resolution for all of the peptide peaks were obtained, allowing unequivocal assignment of singly-charged unmodified and biotinylated peptide fragments. (**B**) The typical Q-TOF data illustrate the protection of a lysine residue in XPA. The peak corresponding to peptide fragment 221, 222, 224 + biotin in XPA is an example of protection from modification by binding of ds–ssDNA junction DNA. When junction DNA is not present, K221, K222 and K224 are readily modified by NHS-biotin; however, in the presence of junction DNA the modification peak disappears. (**C**) In contrast, K188 is a lysine residue not protected by junction DNA binding. A modification peak appears upon treatment with NHS-biotin and persists following addition of junction DNA. Each multiply-charged peptide resulted in clearly resolved peak clusters, indicating monoisotopic resolution; unmodified peaks C1 and C2 serve as controls.

cleavage sites within the fragment. Figure 2B shows MALDI-TOF analysis with a typical assignment of lysine protection. When free XPA was NHS-biotin modified, a peak corresponding to XPA peptide fragment 219-227 (K221, 222, 224+biotin) was observed. However, when XPA-DNA complex was modified, this peak was absent, indicating shielding of lysine 221, 222, and 224 in the nucleoprotein complex from modification. This NHS-biotin modified peak also was absent in the untreated free XPA protein. In contrast, lysine K188, however, was not protected from NHS-biotin modification by XPA-DNA junction binding as the peak corresponding to peptide fragment 184-189 (K188 + biotin) persists in both the free and DNA-bound XPA spectra (Figure 2C). Peaks C1 and C2 serve as internal controls to provide a reference for peak intensity for each spectrum.

In our analysis, five lysine residues in XPA were found to be protected from NHS-biotin modification by DNA binding (Figure 3). Two of these lysines, K167/168 and K179, were previously shown to be located in DNA-binding domain of XPA and thus may also serve as a control for our footprinting analysis. Interestingly, the three lysines, K221, K222 & K224, have not been previously identified to be involved in XPA interaction with DNA. As summarized in Figure 3B, two other lysines outside of the binding domain were partially protected, K31 and K236. These lysines may also be involved in DNA junction binding. These newly uncovered DNA binding residues is of particular interest as they are located outside of the DNA-binding domain (residues 98-219) and may play an important role in the structure-function relationship of XPA.

A



B

Proteolysis Fragments of XPA Containing Modified Lysines

Fragment	Modified K	Protection
31-32	31	Partial
164-168	167	Partial
164-179	167 or 168	+
169-183	179	+
184-189	188	-
190-207	204	-
219-227	221, 222, 224	+
222-227	222, 224	+
232-237	236	Partial

Figure 3. Summary of the footprinting results in the context of the XPA sequence (**A**) Biotinylation sites in XPA are indicated in the primary amino acid sequence either as protected (boxed B) or as unprotected residues (unboxed B). The location of the DBD is indicated by the shaded sequence plus the name and amino acid numbering of the structure. Panel (**B**) tryptic digest fragments of XPA are shown with the modified lysine residue indicated. Lysines that are shown to be protected from modification in the presence of ds–ssDNA junctions are indicated by (+) while residues readily modified in the presence and absence of ds–ssDNA junction are indicated by (−). Lysine residues that were only partially protected are indicated by (Partial). Even though the evidence was not as strong for 'Partial' residues, it is likely there is interaction of these lysines with ds–ssDNA junctions.

Limited Proteolysis of XPA-DNA junction

Partial proteolysis has been widely used to map protein's domains [31,35-37]. Accessible proteolytic cleavage sites are usually located in loops between domains on the surface of a protein. Protein binding to DNA could change the accessibility of some of the cleavage sites, which would alter the proteolytic cleavage pattern [38]. In this study the cleavage pattern of XPA, in the presence or absence of DNA junction substrates, was analyzed and compared to determine which regions interacted with DNA junction. XPA protein was partially digested before or after the addition of DNA junction with a fixed amount of chymotrypsin for the indicated time periods (Figure 4). In the case of XPA binding to junction DNA, several peptides were more sensitive while others were more resistant to cleavage as shown in Figure 4A (peptides 1 & 2 *vs*. peptides 3 & 4, respectively). Three N-terminal cleavage sites of these peptides were determined by protein microsequencing. These chymotrypsin-sensitive sites occurred at amino acid positions 26, 48 and 75 in XPA (Figure 4B, Table 1). The calculated C-terminal cleavages occurred approximately at three amino acid positions 246, 238 and 226 for peptides 1 & 2, 3 & 4 (Table 2), respectively. The determined peptides were summarized in Figure 4B. Upon XPA-DNA complex formation only peptides 3 & 4 persisted at 60 min of chymotrypsin digestion (Figure 4A). The shift from peptides 1 & 2 to peptides 3 & 4 suggests that XPA binding to junction DNA led to conformational changes, making peptide bonds at positions 48, 75, 226 and 238 significantly more sensitive to proteolytic cleavage. This was consistent with the NHS-biotin data that the novel DNA-binding site containing K221, K222 and K224 was protected from modification in the XPA-DNA junction complex. This region falls just outside the known DNA-binding domain of XPA.

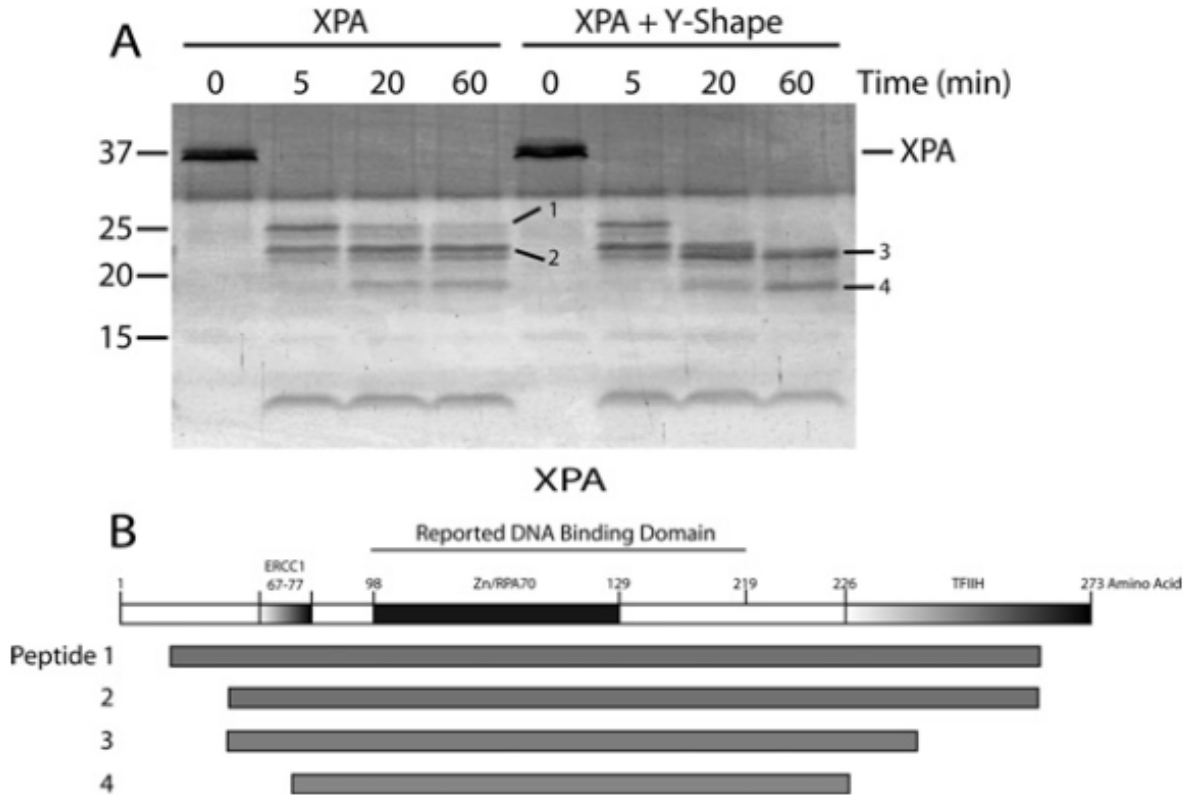


Figure 4. Partial proteolysis of XPA versus XPA–DNA junction complex by chymotrypsin (**A**) Four microgram of XPA was digested with chymotrypsin (1:80 molar ratio of chymotrypsin:XPA) at room temperature for the times indicated. The reactions were terminated and the cleavage products resolved by SDS–PAGE (15 % gel) then visualized using SYPRO Ruby stain. Untreated XPA was loaded as a control (0 time). The molecular mass markers are indicated on the left. Individual fragments of interest are designated by dashes with numbers. (**B**) Schematic map of XPA with its functional domains highlighted followed by a schematic of the four proteolytic peptides indicated in (A) that were generated upon XPA treatment with chymotrypsin.

Table 1 Identification of the peptides generated by chymotrypsin cleavage of XPA with and without bound junction DNAs

Peptide*	Size (kDa)†	N-terminal sequence	Location‡
1	26	SIER	27–246
2	23	SATA	49–246
3	22	SATA	49–238
4	18	ILEE	76–226

*Peptides are numbered as illustrated in Figure 4(A).
†Peptide sizes are estimated from their electrophoretic mobility relative to the Bio-Rad Precision Plus size standards.
‡The location of the peptide in the intact XPA sequence is calculated as the position of the first amino acid in the determined N-terminal sequence plus the number of amino acids corresponding to this size of the peptide as to the nearest likely chymotrypsin-cleavage site using Geneious Software.

These data could also suggest that K236, found to be partial protected, could also be involved in binding. Taken together, the partial proteolysis data confirmed the higher resolution results obtained from mass spectrometric footprinting.

Discussion

A growing body of recent studies has shown that XPA plays an important role not only in NER but also in some other DNA metabolic pathways, not necessarily related to DNA damage. This is consistent with the fact that XPA has a much higher affinity for DNA junctions than damaged dsDNA [16]. The former is a structure commonly present as the intermediate in many DNA metabolic pathways including NER. Given the importance of XPA, defining the structural aspects of XPA-DNA interaction is essential for understanding the mechanisms of XPA in DNA metabolism. The present study represents an effort to understand the structure-function of XPA by generating the structural information on the native full-length XPA interacting with DNA.

We have identified at least six lysine residues of XPA and the corresponding domain involved in the XPA-DNA junction interaction. Of them, K168 and K179 were located in the reported DNA binding domain of XPA and previously shown to be aligned with DNA in a modeling study [39]. It has been also showed previously that K179 interacted with damaged DNA in absence of RPA70 leading to the question of whether the absence of RPA70 freed K179 of XPA to interact with DNA [40]. In contrast to K169 and K179, the newly identified lysine residues K221, K222, K224, and K236 are all located in the C-terminal domain which is outside of the DBD. The C-terminal domain has not been shown previously to interact

with DNA. Our results expand the DNA-binding regions of XPA from the previously known DBD to the C-terminal domain, suggesting that redefining the DNA binding domain of XPA may be necessary.

To understand how the newly identified residues and domain are involved in XPA-DNA junction binding, a predicted structure of the C-terminal XPA$^{217-245}$ was generated based on sequence of XPA$^{217-273}$ using RaptorX software package [41-43]. Discovery Studio software was then used to tether this predicted structure with the published solution structure of XPA DNA binding domain (pdb: 1XPA) by overlapping 3 amino acids shared between the two structures. Thus a structural model of XPA$^{98-241}$ was generated (Figure 5). In this model, the newly identified K221, K222, K224 are located in a α-helical structure with K221, K222 and K224 open to the binding cleft of the known DBD. In contrast, K168 and K179 are located in the β-sheets of DBD facing the C-terminal α-helix. Interestingly, this structural model is featured with two linked arms with the two groups of lysines located separately in each of the arms. The space between the two arms appears to form a perfect path for docking DNA. Thus the structure serves more like a clamp in light of DNA binding. This structure model also shows some similarity to the "right hand" structure commonly found in many DNA polymerases. The two arms resemble the "thumb" and "fingers" while the region connecting the two arms resembles the "palm" although it contains no catalytic activity found in polymerases.

To determine whether XPA could adopt different structures in binding different types of DNA junction, three different DNA junction substrates were used (Figure 1). However, we observed no significant difference in the structural determination of XPA binding to these substrates. A possible explanation for the results is that besides binding to

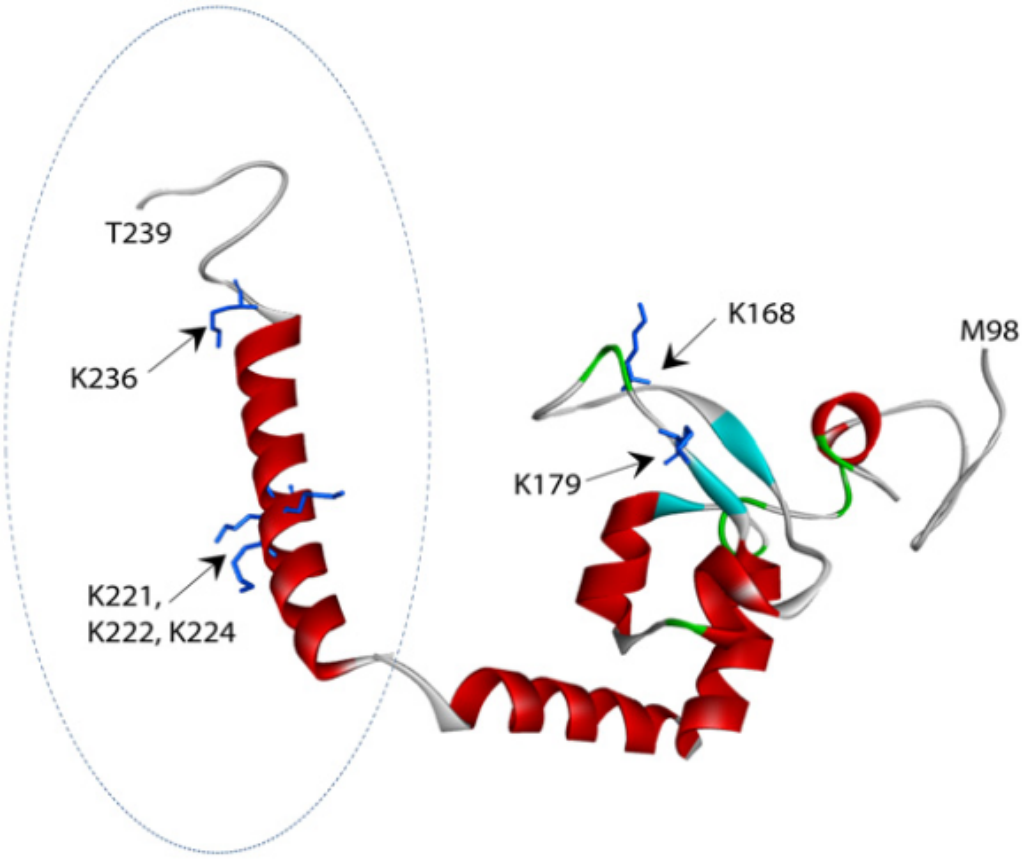


Figure 5. A structural model of redefined DBD of XPA

Structure of the previously reported DBD (aa98-219) of XPA is shown with the predicted extended-structure generated by RaptorX (aa217–239, indicated by dashed circle). Structural mode of the redefined DBD of XPA was generated by Discovery Studio tethering of the two structures together by overlapping three amino acids shared between the two structures. Lysine residues in the structure are presented in stick representation. K168 and K179 are found in the previously reported DBD of XPA, while K221, K222, K224 and K236 are found within the extended structure. The two structures resemble two arms of a clamp. Biotin-modified lysine residues protected from modification in the presence of ds–ssDNA junctions are shown in blue.

the double-stranded region at the junction, XPA may additionally bind only to one ssDNA sequence without preference for either 3'- or 5'-ssDNA overhang. However, it is also possible that some structural differences may occur in non-lysine containing regions of XPA but could be at the minimal level as our limited proteolysis analysis showed no differences among these substrates.

It is worth noting that the very recent work by Sugitani *et al.* analyzed an extended DNA binding domain of XPA. Using fluorescence spectroscopic method, they demonstrate that residues outside of the established DNA binding domain are not only involved in binding but increase affinity by five fold using y-shape DNA junctions [44]. Their NMR data confirms that the extended DNA binding domain binds to junction DNA much stronger than a construct with only the traditional binding domain length. Their finding that XPA_{98-239} binds at a higher affinity to junction DNA as compared to duplex DNA shows not only that the XPA binding domain is more extensive than previously thought but that XPA's main function in the cell could have been initially misinterpreted. These are in agreement with our findings in this study.

References

(1) Thoma, B. S.; Vasquez, K. M. *Mol. Carcinog.* **2003**, *38* (1), 1–13.

(2) Sancar, A.; Lindsey-Boltz, L. A.; Unsal-Kaçmaz, K.; Linn, S. *Annu. Rev. Biochem.* **2004**, *73* (1), 39–85.

(3) Reardon, J. T.; Sancar, A. *Prog. Nucleic Acid Res. Mol. Biol.* **2005**, *79*, 183–235.

(4) Wood, R. D. *Biochimie* **1999**, *81* (1-2), 39–44.

(5) Liu, Y.; Liu, Y.; Yang, Z.; Utzat, C.; Wang, G.; Basu, A. K.; Zou, Y. *Biochemistry* **2005**, *44* (19), 7361–7368.

(6) Cleaver, J. E.; States, J. C. *Biochem J* **1997**, *328 (Pt 1)*, 1–12.

(7) Batty, D. P.; Wood, R. D. *Gene* **2000**, *241* (2), 193–204.

(8) van Steeg, H.; Mullenders, L. H.; Vijg, J. *Mutat Res* **2000**, *450* (1-2), 167–180.

(9) Riedl, T.; Hanaoka, F.; Egly, J. M. *The EMBO Journal* **2003**, *22* (19), 5293–5303.

(10) Park, C.-J.; Choi, B.-S. *FEBS J* **2006**, *273* (8), 1600–1608.

(11) Niedernhofer, L. J.; Garinis, G. A.; Raams, A.; Lalai, A. S.; Robinson, A. R.; Appeldoorn, E.; Odijk, H.; Oostendorp, R.; Ahmad, A.; Van Leeuwen, W.; Theil, A. F.; Vermeulen, W.; van der Horst, G. T. J.; Meinecke, P.; Kleijer, W. J.; Vijg, J.; Jaspers, N. G. J.; Hoeijmakers, J. H. J. *Nature* **2006**, *444* (7122), 1038–1043.

(12) Liu, Y.; Wang, Y.; Rusinol, A. E.; Sinensky, M. S.; Liu, J.; Shell, S. M.; Zou, Y. *The FASEB Journal* **2008**, *22* (2), 603–611.

(13) Fang, E. F.; Scheibye-Knudsen, M.; Brace, L. E.; Kassahun, H.; SenGupta, T.; Nilsen, H.; Mitchell, J. R.; Croteau, D. L.; Bohr, V. A. *Cell* **2014**, *157* (4), 882–896.

(14) Sugasawa, K.; Ng, J. M.; Masutani, C.; Iwai, S.; van der Spek, P. J.; Eker, A. P.; Hanaoka, F.; Bootsma, D.; Hoeijmakers, J. H. *Mol Cell* **1998**, *2* (2), 223–232.

(15) Köberle, B.; Roginskaya, V.; Wood, R. D. *DNA Repair (Amst)* **2006**, *5* (5), 641–648.

(16) Yang, Z.; Roginskaya, M.; Colis, L. C.; Basu, A. K.; Shell, S. M.; Liu, Y.; Musich, P. R.; Harris, C. M.; Harris, T. M.; Zou, Y. *Biochemistry* **2006**, *45* (51), 15921–15930.

(17) Guzder, S. N.; Sommers, C. H.; Prakash, L.; Prakash, S. *Mol Cell Biol* **2006**, *26* (3), 1135–1141.

(18) Park, C. H.; Sancar, A. *Proc Natl Acad Sci USA* **1994**, *91* (11), 5017–5021.

(19) Ikegami, T.; Kuraoka, I.; Saijo, M.; Kodo, N.; Kyogoku, Y.; Morikawa, K.; Tanaka, K.; Shirakawa, M. *Nat. Struct. Biol.* **1998**, *5* (8), 701–706.

(20) Evans, E. *The EMBO Journal* **1997**, *16* (3), 625–638.

(21) Krasikova, Y. S.; Rechkunova, N. I.; Maltseva, E. A.; Petruseva, I. O.; Lavrik, O. I. *Nucleic Acids Res* **2010**, *38* (22), 8083–8094.

(22) Tapias, A.; Auriol, J.; Forget, D.; Enzlin, J. H.; Scharer, O. D.; Coin, F.; Coulombe, B.; Egly, J. M. *J Biol Chem* **2004**, *279* (18), 19074–19083.

(23) Gilljam, K. M.; Feyzi, E.; Aas, P. A.; Sousa, M. M. L.; Muller, R.; Vagbo, C. B.; Catterall, T.

C.; Liabakk, N. B.; Slupphaug, G.; Drablos, F.; Krokan, H. E.; Otterlei, M. *J Cell Biol* **2009**, *186* (5), 645–654.

(24) Missura, M.; Buterin, T.; Hindges, R.; Hübscher, U.; Kaspárková, J.; Brabec, V.; Naegeli, H. *The EMBO Journal* **2001**, *20* (13), 3554–3564.

(25) Camenisch, U.; Dip, R.; Schumacher, S. B.; Schuler, B.; Naegeli, H. *Nat. Struct. Mol. Biol.* **2006**, *13* (3), 278–284.

(26) Lambert, M. W.; Yang, L. *Biochem Biophys Res Commun* **2000**, *271* (3), 782–787.

(27) Tsodikov, O. V.; Ivanov, D.; Orelli, B.; Staresincic, L.; Shoshani, I.; Oberman, R.; Scharer, O. D.; Wagner, G.; Ellenberger, T. *The EMBO Journal* **2007**, *26* (22), 4768–4776.

(28) Ikegami, T.; Kuraoka, I.; Saijo, M.; Kodo, N.; Kyogoku, Y.; Morikawa, K.; Tanaka, K.; Shirakawa, M. *J. Biochem.* **1999**, *125* (3), 495–506.

(29) Kvaratskhelia, M.; Miller, J. T.; Budihas, S. R.; Pannell, L. K.; Le Grice, S. F. J. *Proc Natl Acad Sci USA* **2002**, *99* (25), 15988–15993.

(30) Shkriabai, N.; Patil, S. S.; Hess, S.; Budihas, S. R.; Craigie, R.; Burke, T. R.; Le Grice, S. F. J.; Kvaratskhelia, M. *Proc Natl Acad Sci USA* **2004**, *101* (18), 6894–6899.

(31) Liu, Y.; Kvaratskhelia, M.; Hess, S.; Qu, Y.; Zou, Y. *J Biol Chem* **2005**, *280* (38), 32775–32783.

(32) Yang, Z.-G.; Liu, Y.; Mao, L. Y.; Zhang, J.-T.; Zou, Y. *Biochemistry* **2002**, *41* (43), 13012–13020.

(33) Shell, S. M.; Li, Z.; Shkriabai, N.; Kvaratskhelia, M.; Brosey, C.; Serrano, M. A.; Chazin, W. J.; Musich, P. R.; Zou, Y. *J Biol Chem* **2009**, *284* (36), 24213–24222.

(34) Shell, S. M.; Hess, S.; Kvaratskhelia, M.; Zou, Y. *Biochemistry* **2005**, *44* (3), 971–978.

(35) Gomes, X. V.; Henricksen, L. A.; Wold, M. S. *Biochemistry* **1996**, *35* (17), 5586–5595.

(36) Heyduk, E.; Heyduk, T. *Biochemistry* **1994**, *33* (32), 9643–9650.

(37) Iakoucheva, L. M.; Kimzey, A. L.; Masselon, C. D.; Bruce, J. E.; Garner, E. C.; Brown, C. J.; Dunker, A. K.; Smith, R. D.; Ackerman, E. J. *Protein Sci* **2001**, *10* (3), 560–571.

(38) Zappacosta, F.; Pessi, A.; Bianchi, E.; Venturini, S.; Sollazzo, M.; Tramontano, A.; Marino, G.; Pucci, P. *Protein Sci* **1996**, *5* (5), 802–813.

(39) Buchko, G. W.; Tung, C. S.; McAteer, K.; Isern, N. G.; Spicer, L. D.; Kennedy, M. A.

Nucleic Acids Res **2001**, *29* (12), 2635–2643.

(40) Camenisch, U.; Dip, R.; Vitanescu, M.; Naegeli, H. *DNA Repair (Amst)* **2007**, *6* (12), 1819–1828.

(41) Ma, J.; Wang, S.; Zhao, F.; Xu, J. *Bioinformatics* **2013**, *29* (13), i257–i265.

(42) Peng, J.; Xu, J. *Proteins* **2011**, *79 Suppl 10* (S10), 161–171.

(43) Peng, J.; Xu, J. *Proteins* **2011**, *79* (6), 1930–1939.

(44) Sugitani, N.; Shell, S. M.; Soss, S. E.; Chazin, W. J. *J. Am. Chem. Soc.* **2014**, 140724070100004.

Supplemental Data

Figure S1

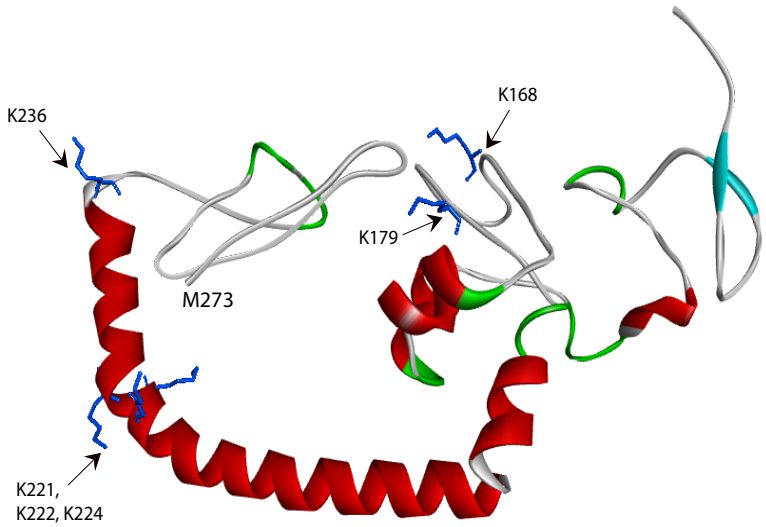


Figure S1. I-TASSER structural model of redefined DNA binding domain of XPA.
This representative model was generated by submitting the redefined XPA DNA binding domain amino acid sequence (aa98-273) for protein structure prediction to the I-TASSER software package. 5 models were generated and the model with best fit was selected. Lysine residues in the structure are presented in stick representation. K168 and K179 are found in the previously reported DNA-binding domain of XPA, while K221, K222, K224, and K236 are found within the extended structure. The two structures resemble two arms of a clamp. Biotin-modified lysine residues protected from modification in the presence of ds-ssDNA junctions are shown in blue.

Yang Zhang. I-TASSER server for protein 3D structure prediction. BMC Bioinformatics, 9:40 (2008).

Ambrish Roy, Alper Kucukural, Yang Zhang. I-TASSER: a unified platform for automated protein structure and function prediction. Nature Protocols, vol 5, 725-738 (2010).

Ambrish Roy, Jianyi Yang, Yang Zhang. COFACTOR: an accurate comparative algorithm for structure-based protein function annotation. Nucleic Acids Research, vol 40, W471-W477 (2012).

CHAPTER 6

REPLICATION FACTOR C1, THE LARGE SUBUNIT OF REPLICATION FACTOR C, IS PROTEOLYTICLY TRUNCATED IN HUTCHINSON-GILFORD PROGERIA SYNDROME

Abstract

Hutchinson-Gilford progeria syndrome (HGPS) is caused by a LMNA gene mutation which produces a mutant lamin A protein (progerin). However, how progerin causes the progeria remains unknown. Here we report that the large subunit (RFC1) of replication factor C is cleaved by an unknown serine protease in HGPS cells, leading to the production of a truncated RFC1 of ~75 kDa which appears to be defective in loading PCNA and pol δ onto DNA for replication. Our findings provide a mechanistic interpretation for the observed replicative arrest and progeria phenotypes of HPGS, and may lead to novel strategies in HGPS treatment.

Introduction

Hutchinson-Gilford progeria syndrome (HGPS) is a genetic disease that results in premature aging due to a *de novo* point mutation (1824C→T) in the *LMNA* gene (Eriksson et al. 2003; De Sandre-Giovannoli et al. 2003). The mutation results in sporadic activation of a cryptic splice donor site in exon 11 of the prelamin A pre-mRNA, leading to a 50 amino acids deletion within prelamin A. The deleted region includes an endoprotease cleavage site required for maturation of lamin A, an intermediate filament protein in the nuclear lamina underlying the inner nuclear membrane (Goldman et al. 2002). Lack of this cleavage

leads to production of a farnesylated and carboxylmethylated lamin A mutant protein (progerin) which causes nuclear envelope dysfunction, DNA double-strand breaks (DSBs), activation of DNA damage responses, early replicative arrest, and, ultimately, accelerated aging (Liu et al. 2006; Bridger & Kill 2004; Musich & Zou 2009; Capell & Collins 2006; Busch et al. 2009; Liu et al. 2008). In addition, progerin occurs in the cells of healthy individuals and is responsible for the HGPS-like nuclear defects in the cells of normal aging individuals (Cao et al. 2011; Scaffidi & Misteli 2006). The level of progerin also increases with the age of healthy individuals and be correlated with aging-related cardiovascular diseases (Ragnauth et al. 2010; Olive et al. 2010).

Materials and Methods

Cell Culture and Drug Treatment

Fibroblasts from two HGPS patients with the mutation at codon 608 (G608G) within exon 11 of the LMNA gene (AG11513A and AG11498B designated as HGPS-1 and HGPS-2, respectively) were obtained from Coriell Cell Repository. HGPS patient cells were also obtained from the Progeria Research Foundation with the point mutation 1824C→T (HGADFN143, HGADFN178, HGADFN271, designated as HGPS-3, 4, and 5, respectively). HeLa cells were purchased from American Type Culture Collection (ATCC, CCL-2). All cultures were maintained in Dulbecco's modified Eagle's medium (DMEM) (A549 and HeLa) or Eagle's minimal essential medium (EMEM) (HGPS) supplemented with 10% or 15% fetal bovine serum, secondary and primary cells respectively, and antibiotics (50 units/ml penicillin and 50 ug/ml streptomycin) at 37°C under an atmosphere containing

5% CO_2. For protease and proteosome inhibitor treatments, cells were cultured to 70% confluence, and treated with the indicated concentrations of pepstain, AEBSF, E64, and MG132 for 48 hours (pepstain, E64 and MG132 were a gift from Dr. David Johnson, ETSU). To induce DNA damage cells were treated with camptothecin (CPT) after growth to 70% confluence, followed by incubation with 10 μM CPT for 6 hours, no release. Cells were treated with 20 J/m^2 followed by a 6 hour release period before harvesting.

Transfection of siRNA

To knockdown RFC1 by RNAi, the cells were transfected with a pool of RFC1 siRNAs (Santa Cruz, sc-37631) or a scrambled sequence as a control using INTEFERin transfection reagent Polyplus-transfection SA (France) following the manufacturer's instructions. Further analyses were performed 72 hours after transfection.

Preparation of cell lysates, nuclear extracts and immunoblots

The following steps were performed on ice in presence of protease inhibitor cocktail (Thermo Scientific, Pierce, 78430). For whole cell lysates, cell pellets were resuspended in lysis buffer (50mM Tris-Cl, 150mM NaCl, 1%NP-40) and incubated on ice for a minimum of 10 minutes. SDS gel loading buffer (Molecular Cloning) was added to the samples and boiled for 5 minutes. For nuclear extraction, cells were resuspended in 50 μl of buffer A (20 mM HEPES, pH 7.8, 10 mM KCl, 0.15 mM EDTA, 0.15 mM EGTA) followed by lysis with 3 μl 0.5% NP-40. Then 15 ml of sucrose restoration buffer (50 mM HEPES, pH 7.0, 0.25 mM EDTA, 10 mM KCl, 70% sucrose) was added and the samples were centrifuged at 5000 rpm

for 5 minutes. The cytoplasmic fraction was removed and the nuclei were washed in 50 µl buffer B (10 mM HEPES, pH 8.0, 0.1 M NaCl, 0.1 mM EDTA, 25% glycerol), lysed by resuspension in 40 µl buffer C (10 mM HEPES, pH 8.0, 0.4 M NaCl, 25% glycerol, 0.1 mM EDTA) and incubated for 30 minutes on ice on a rotating rack. Samples were centrifuged for 10 minutes at 10,000 rpm and nuclear proteins were removed.

Chromatin fractionation Assay

Cellular localization of endogenous degraded RFC1 in HGPS cells was investigated by the chromatin fractionation technique based on the method of Mendez and Stillman. Briefly, HGPS cells were plated in a 100-mm dish. At 80% confluence, cells were washed twice with ice-cold PBS and the cell pellet was resuspended in 100 µl of Solution A (10 mM Hepes–KOH pH 7.9, 10 mM KCl, 1.5 mM $MgCl_2$, 0.34 M sucrose, 1X protease inhibitor, 1mM PMSF and 0.1% TritonX-100. After 5 minutes incubation on ice, samples were centrifuged at 1300 g for 4 minutes. The pellet was washed once in Solution A and designated as intact nuclei. Nuclei were resuspended in 100 µl of Solution B (3 mM EDTA , 0.2 mM EGTA and 1X protease inhibitor cocktail) and incubated on ice for 10 minutes. The pellet was washed once with Solution B, and designated as the chromatin fraction.

Western Blotting

Cells cultured in 100-mm dishes or 25 cm^2 flasks were grown to 70% confluence and then trypsinized. Cell number was counted using a hemacytometer. The cells were centrifuged at 500 ***g*** for 5 minutes and washed twice with PBS in the presence of protease

inhibitor cocktail. The cell pellet was lysed in 2x SDS gel loading buffer and volumes corresponding to $5*10^6$ cells were subjected to SDS-PAGE. SDS–PAGE was carried out in 8 or 4-12% gels (BioRad). Immunoblotting was performed as previously described (Liu et al., 2005) with primary antibodies directed against RFC1 (Bethyl Laboratories, A300-141A, Santa Cruz, sc-20993 and Fisher/Millipore, 09-094), Lamin A/C (Cell Signaling Technology, Beverly, MA), Histone 3 (Santa Cruz), β-Actin (Sigma-Aldrich), PCNA (Santa Cruz), Pol δ (Santa Cruz). Proteins were transferred to PVDF membranes (SantaCruz, USA) and immunoblotting was performed with the indicated antibodies followed by HRP-conjugated antibodies (SantaCruz). The complexes were revealed using SuperSignal West Pico Chemiluminescent Substrate (Thermo Scientific, Pierce).

Results and Discussion

Despite these findings, the molecular mechanism by which progerin causes HGPS or premature aging is largely unknown. Understanding the mechanisms is essential for identification of molecular targets for therapeutic treatment of HGPS disease. Since early replicative arrest is a hallmark phenotype of HGPS and DSBs can form from collapse of stalled replication forks during replication, we reasoned that the replication machinery in HGPS could be dysfunctional. Indeed, analysis of the nuclear extracts of HGPS cells demonstrated that the integrity of replication factor C (RFC) was compromised in HGPS cells (Fig. 1A). Specifically, the large subunit of RFC (RFC1, 140 kDa) was found to be degraded to a ~75 kDa C-terminal fragment (RFC1-75), suggesting that a proteolytic cleavage of RFC1 occurred in HGPS cells. The Western blot results were confirmed using three different RFC1 C-terminus specific antibodies. Also, this RFC1-75 band was absent in

blots probed with N-terminal specific antibodies (data not shown). In addition, HGPS cells were transfected with RFC1 siRNA to knock down the expression of this protein. The knockdown was efficient for both intact and truncated RFC1 (Fig. 1B), further confirming that the truncated fragment was from RFC1. Also, the level of RFC1-75 increased with cell passage in culture (Fig. 1A). This RFC1 cleavage is a common event of HGPS since the RFC1-75 fragment was seen in cells from five different HGPS patients (Fig. 1C). To determine if RFC1 degradation is a result of DNA damage, HeLa cells were treated with UV-C irradiation or camptothecin (CPT) to induce bulky DNA damage and DSBs, respectively. The treatment resulted in no cleavage of RFC1 even though significant amount DSBs were induced (Fig. 1D), suggesting that the cleavage is independent of DNA damage but an event unique to HGPS.

RFC1 is the large subunit of the RFC complex consisting of RFC1, RFC2, RFC3, RFC4 and RFC5. RFC1 contains the major DNA-binding domains of RFC and is directly involved in RFC-PCNA interaction. RFC, widely known as the DNA clamp loader, plays a crucial role in replication: RFC binds to the 3' end of the primed nascent DNA strand and loads PCNA and DNA polymerase δ or ε onto the replication forks. It is highly possible that the truncation of RFC1 may disrupt the assembly of replication machinery and stall the forks, ultimately leading to replication fork collapse and formation of DSBs. Indeed, recruitment of PCNA and pol δ to chromatin was inhibited as the cleaved RFC1 accumulated in HGPS cells (Fig. 1E), suggesting that RFC1-75 is functionally defective in recruiting PCNA and pol δ to replication forks. This is consistent with the fact that HGPS cells are characterized by early replicative arrest.

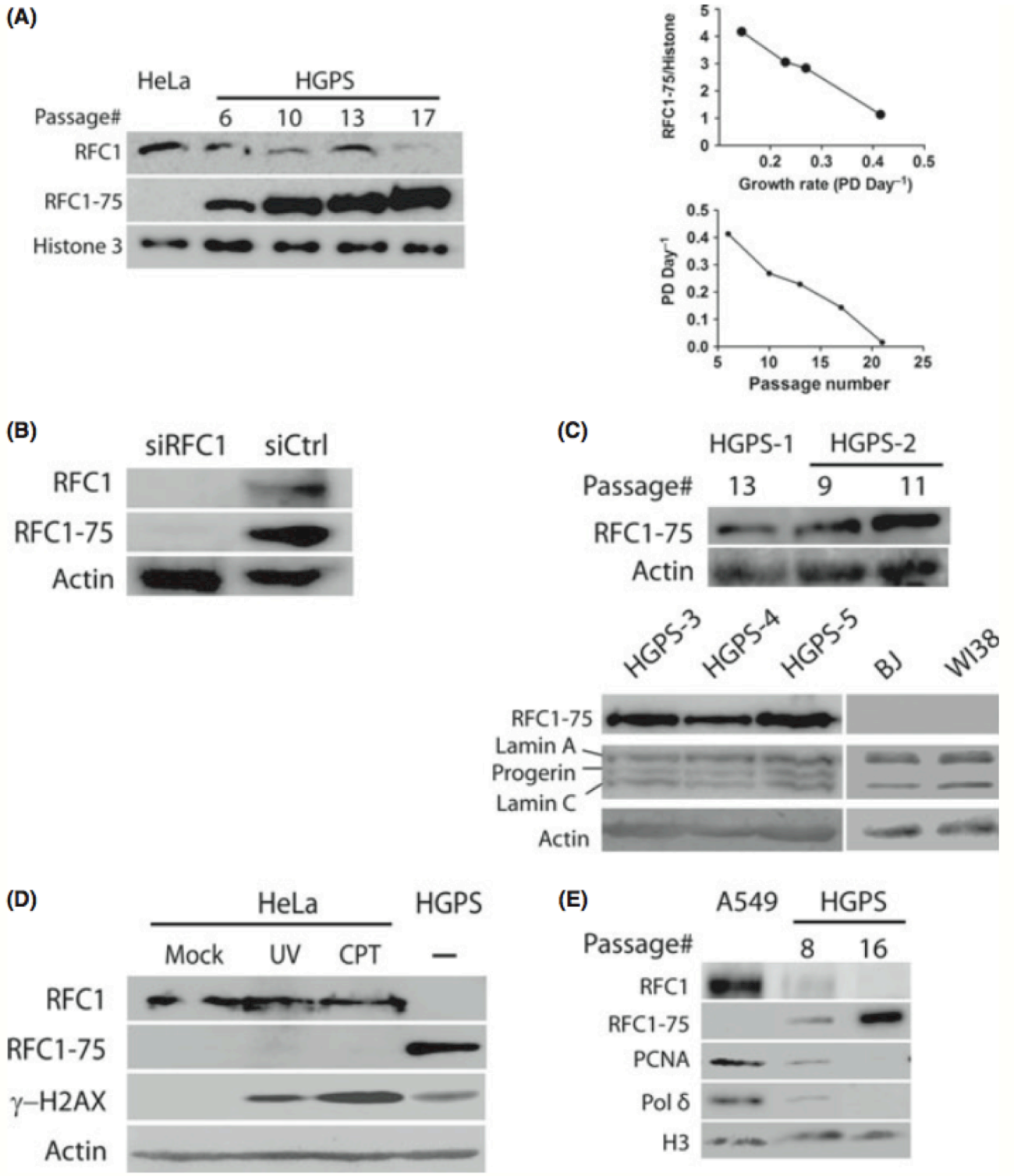


Fig. 1 Proteolytic cleavage of RFC1 in HGPS cells. (A) HGPS fibroblasts were grown in culture to various passage numbers. Nuclear extracts were analyzed by Western blotting with anti-RFC1 antibody. HeLa cells were used as a control cell line and histone 3 as a loading control. (B) RFC1 siRNAs were transfected into HGPS cells for 72 hrs. A scrambled siRNA was used as a control. (C) Cells from five different HGPS patients were lysed and analyzed by Western blotting. BJ and WI38 are human normal fibroblasts used as control cells. (D) HeLa cells were treated or mock-treated with 20 J/m^2 UV or 10 uM CPT. The cell lysates were analyzed using anti-RFC1 antibodies. The γ–H2AX bands indicate formation of DSBs induced by the treatments. (E) Chromatin association assay was performed for HGPS cells with different passage numbers. A549 cells served as a control.

The finding that RFC1 is unexpectedly truncated in HGPS could have a significant and broad implication in addressing the mechanisms of replicative senescence and aging due to the close relationship between replication and cell proliferation and to the relevance of progerin to aging. In addition, it has been noticed that HGPS patients appear to be cancer free. Although this could be due to their short average lifespan of 13 years, the fact that RFC1 is functionally defective in HGPS cells clearly is a possible explanation.

It is of great interest to identify the protease(s) that cleaved RFC1 in HGPS cells since the identification may lead to a therapeutic treatment of HGPS. Thus, inhibitors targeting different types of proteases were tested. Pepstain, a potent inhibitor of aspartyl proteases, and E64 which inhibits cysteine peptidases, showed no inhibition to RFC1 cleavage (Fig. 2A). In contrast, AEBSF, a potent serine protease inhibitor, efficiently inhibited the cleavage in a dose-dependent manner. We further examined the effect of the proteasome inhibitor MG132 which showed no substantial effect on RFC1 cleavage (Fig. 2B). Taken together, our results suggest that the protease responsible for the RFC1 cleavage is likely a serine protease. Obviously, future effort to identify the protease would be very helpful not only to understand the mechanisms of HGPS disease progression but also to provide new strategies for treatment of the HGPS disease.

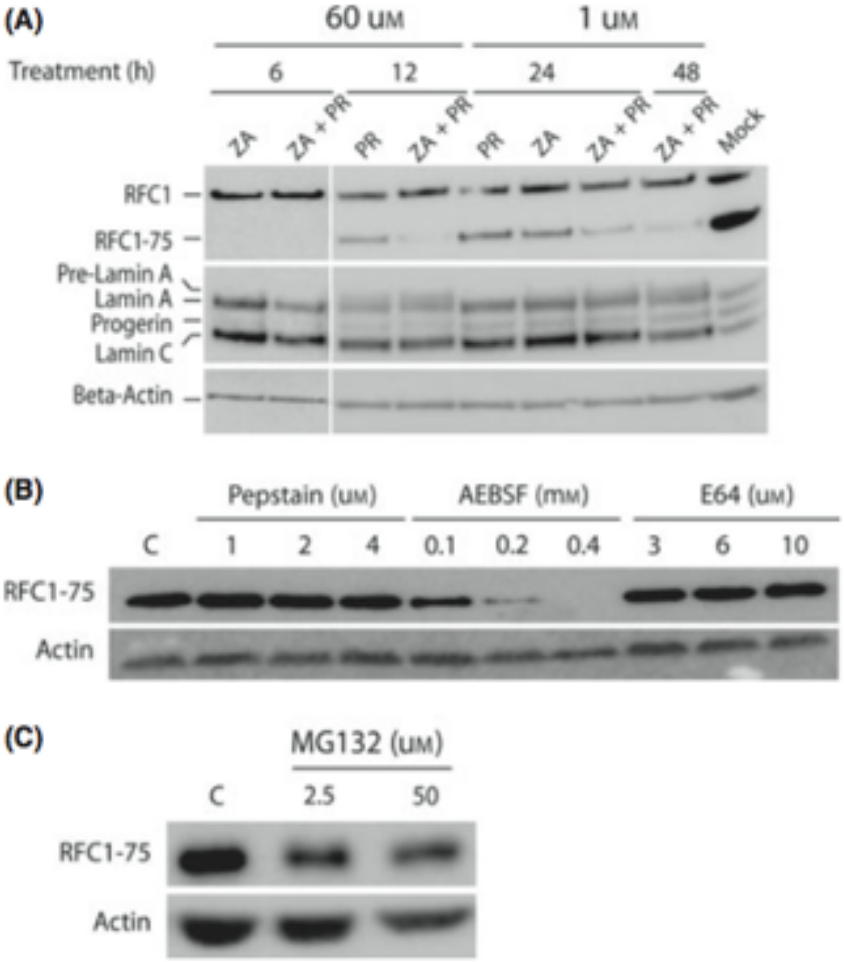


Fig. 2 Degradation of RFC1 in HGPS cells is prevented by protease inhibitors. (A) HGPS cells were treated with the indicated concentrations of protease inhibitor. Cell lysates were then analyzed by Western blotting. "C" indicates the HGPS cells that were mock treated. (B) HGPS cells were treated with the proteasome inhibitor MG132 and analyzed as in (A).

References

Bridger JM & Kill IR (2004) Aging of Hutchinson-Gilford progeria syndrome fibroblasts is characterised by hyperproliferation and increased apoptosis. *Exp. Gerontol.* 39, 717–724.

Busch A, Kiel T, Heupel W-M, Wehnert M & Hübner S (2009) Nuclear protein import is reduced in cells expressing nuclear envelopathy-causing lamin A mutants. *Exp Cell Res* 315, 2373–2385.

Cao K, Blair CD, Faddah DA, Kieckhaefer JE, Olive M, Erdos MR, Nabel EG & Collins FS (2011) Progerin and telomere dysfunction collaborate to trigger cellular senescence in normal human fibroblasts. *J Clin Invest*.

Capell BC & Collins FS (2006) Human laminopathies: nuclei gone genetically awry. *Nat. Rev. Genet.* 7, 940–952.

De Sandre-Giovannoli A, Bernard R, Cau P, Navarro C, Amiel J, Boccaccio I, Lyonnet S, Stewart CL, Munnich A, Le Merrer M & Lévy N (2003) Lamin a truncation in Hutchinson-Gilford progeria. *Science* 300, 2055.

Eriksson M, Brown WT, Gordon LB, Glynn MW, Singer J, Scott L, Erdos MR, Robbins CM, Moses TY, Berglund P, Dutra A, Pak E, Durkin S, Csoka AB, Boehnke M, Glover TW & Collins FS (2003) Recurrent de novo point mutations in lamin A cause Hutchinson-Gilford progeria syndrome. *Nature* 423, 293–298.

Goldman RD, Gruenbaum Y, Moir RD, Shumaker DK & Spann TP (2002) Nuclear lamins: building blocks of nuclear architecture. *Genes Dev* 16, 533–547.

Liu Y, Rusinol A, Sinensky M, Wang Y & Zou Y (2006) DNA damage responses in progeroid syndromes arise from defective maturation of prelamin A. *J Cell Sci* 119, 4644–4649.

Liu Y, Wang Y, Rusinol AE, Sinensky MS, Liu J, Shell SM & Zou Y (2008) Involvement of xeroderma pigmentosum group A (XPA) in progeria arising from defective maturation of prelamin A. *The FASEB Journal* 22, 603–611.

Musich PR & Zou Y (2009) Genomic instability and DNA damage responses in progeria arising from defective maturation of prelamin A. *Aging (Albany NY)* 1, 28–37.

Olive M, Harten I, Mitchell R, Beers JK, Djabali K, Cao K, Erdos MR, Blair C, Funke B, Smoot L, Gerhard-Herman M, Machan JT, Kutys R, Virmani R, Collins FS, Wight TN, Nabel EG & Gordon LB (2010) Cardiovascular pathology in Hutchinson-Gilford progeria: correlation with the vascular pathology of aging. *Arterioscler. Thromb. Vasc. Biol.* 30, 2301–2309.

Ragnauth CD, Warren DT, Liu Y, McNair R, Tajsic T, Figg N, Shroff R, Skepper J & Shanahan CM (2010) Prelamin A acts to accelerate smooth muscle cell senescence and is a novel

biomarker of human vascular aging. *Circulation* 121, 2200–2210.

Scaffidi P & Misteli T (2006) Lamin A-dependent nuclear defects in human aging. *Science* 312, 1059–1063.

CHAPTER 7

PROGERIN SEQUESTERS PCNA RESULTING IN REPLICATION FORK COLLAPSE IN HUTCHINSON-GILFORD PROGERIA SYNDROME

Abstract

Hutchinson–Gilford progeria syndrome (HGPS) is a rare genetic disorder that is characterized by the phenotypes of accelerated and premature aging. HGPS is caused by a point mutation in the *LMNA* gene resulting in production and accumulation of an uncleavable form of farnesylated prelamin A protein (termed progerin or LAΔ50) in cells. We previously reported that accumulation of progerin induces stalled replication forks leading to fork collapse, which then can be processed into double-strand breaks (DSBs) in progeroid cells as evidenced by an increase in the number of γH2AX foci. Interestingly, xeroderma pigmentosum group A (XPA), the DNA damage verification factor in nucleotide excision repair (NER), mislocalized with γH2AX foci in HGPS cells. This suggested a role for XPA at the stalled/collapsed forks induced by progerin. In the present study, we show that progerin accumulates and sequesters the replication protein proliferating cell nuclear antigen (PCNA) leaving an exposed DNA junction. Accumulation of exposed DNA junctions then are marked by XPA mislocalization. Furthermore, we demonstrate that progerin-induced apoptosis can be rescued by XPA binding at exposed junctions. Our results suggest that XPA may play a role in binding junctions to prevent apoptosis in HGPS cells.

Introduction

Hutchinson-Gilford progeria syndrome (HGPS) is a premature-aging disease caused by formation of a carboxyl-terminal-truncated form (progerin, LAΔ50) of the lamin A precursor, prelamin A (De Sandre-Giovannoli et al., 2003; Eriksson et al., 2003). Lamin A is an intermediate filament protein that is part of the nuclear lamina, which structurally supports the nucleus and organizes chromatin (Goldman et al., 2002). Zmpste24 is an endoproteinase essential for the proteolytic maturation of prelamin A (Corrigan et al., 2005); thus, loss of Zmpste24 proteinase activity arrests the processing of prelamin A at a stage similar to HGPS. Homozygous deficiency of Zmpste24 results in restrictive dermopathy (RD), which is a lethal perinatal progeroid disorder (Navarro et al., 2005). In RD a form of prelamin A accumulates, while progerin is accumulated in HGPS cells. RD and HGPS have been grouped with other genetic diseases arising from mutations in the lamin A gene (Gonzalo and Kreienkamp, 2015; Hutchison, 2011). The genetics diseases that arise from mutations in *LMNA* are collectively termed laminopathies.

The replication rate of HGPS cells in culture previously has been shown in young cells at a rate consistent with normal fibroblasts growth; however, as the cells age in culture they reach a level near senescence (Liu et al., 2006; Musich and Zou, 2011). The enzyme ribonucleotide reductase (RNR) is responsible for conversion of ribonucleotides to deoxyribonucleotides, an essential reaction for DNA synthesis and repair, in the nucleus. The enzyme contains two R1 and two R2 subunits (RRM1 & RRM2, respectively), which form a heterodimeric tetramer. The regulation of the complex depends on a balanced deoxyribonucleotide pool while pool alterations are known to lead to genomic instability and have been linked to multiple human diseases (Aye et al., 2014). Furthermore, RRM2

turnover is dependent on cell-cycle progression while the RRM1 subunit is maintained at a constant level. Recent work has demonstrated that RNR can be regulated by the checkpoint kinase ATR in both *E. coli* and yeast. In these studies ATR was found to mediate replication catastrophe and origin firing through the promotion of RRM2 accumulation in cycling cells (Buisson et al., 2015). Interestingly, progeroid cells genetically modified to contain extra alleles coding for RRM2 had a decrease in the severity of the progeroid disease (Lopez-Contreras et al., 2015). p53M2 is the subunit required instead of RRM2 for repair synthesis and also forms an active holoenzyme with RRM1. p53R2 is continuously expressed throughout the cell cycle; however, following DNA damage, p53R2 is transcriptionally induced in a p53-dependent manner. Even though HGPS cells are known to reach a senescence-like state in culture little is known about the relationship between cell proliferation and RNR activity in this system.

Past studies have shown that HGPS and RD cells exhibit double-strand break (DSB) accumulation, impairment of DNA repair, and activation of the p53-dependent stress signaling pathway (Liu et al., 2005; Manju et al., 2006; Scaffidi and Misteli, 2006; Varela et al., 2005), suggesting that genome instability caused by HGPS and RD might contribute to premature aging. The DSB accumulation in HGPS and Zmpste24-deficient cells as well as in senescing and aging mammalian cells appears to be due to DSBs that are not repaired (Sedelnikova et al., 2004). The underlying mechanism of the repair defect is still poorly understood; however, a previous study from our lab demonstrated that DSB repair proteins Rad51 and Rad50 were not localized to the DNA damage sites, whereas xeroderma pigmentosum group A (XPA), a DNA-damage verification protein in nucleotide excision repair, largely colocalized with the DSBs formed in HGPS and RD cells (Liu et al.,

2008). Furthermore, knockdown of XPA in HGPS cells significantly restored DSB repair. The mislocalization of XPA to DSBs may be responsible for the lack of repair of DNA damage in the progeria cells.

In a related study we identified a novel activity of XPA for recognition of double-strand/single-strand DNA (ds-ssDNA) junctions with 3'- and/or 5'-ssDNA overhangs (Yang et al., 2006). The binding affinity of XPA for these sites is between one and two orders of magnitude higher than its ability to bind to bulky DNA damage. In an in-depth analysis by a mass spectrometry footprinting assay we identified additional protein-DNA contacts that could relate to this higher-affinity binding (Hilton et al., 2014). The ds-ssDNA junction structures used in our past studies form as intermediates during some DNA metabolic pathways including replication and repair. When replication is stalled, the replication fork, if left unprotected, is structurally similar to the Y-shape DNA substrate and through further processing ssDNA overhangs are generated, comparable to the overhang substrates used in our study. The ssDNA overhangs then can be further processed into DSBs. The processing of DNA junction structures resulting from replication fork stalling and collapse into DSBs has been previously shown (Cox, 2002; Heller and Marians, 2006; Mirkin and Mirkin, 2007). Interestingly, one group (Bomgarden et al., 2006) reported that XPA, but not NER-mediated damage processing, is required for UV-induced S-phase ATR checkpoint activation, which is replication dependent (Jarrett et al., 2014).

Previous studies have shown that RFC1 is proteolyticly cleaved in a caspase-3-dependent manor during apoptosis initiation to inhibit DNA replication (Rhéaume et al., 1997). RFC1 cleavage is not thought to be caspase-3-dependent in HGPS cells; however, that study does provide evidence that RFC1 can be regulated by proteolytic cleavage to

inhibit replication. In the aged HGPS cells there is an inhibition of DNA replication as evidenced by the lack of PCNA and polymerase δ on the chromatin. This may be due in part to RFC1 being proteolyticly cleaved and absent from the chromatin. Remarkably, when the HGPS cells were treated with drugs used to prevent lamin A processing modifications RFC1 was not proteolyticly cleaved. Therefore, there is a direct link between the cleavage of RFC1 and the accumulation of progerin and XPA at junction DNA.

In the present study, using immunofluorescence microscopy, chromatin association and chromatin immunoprecipitation (ChIP) assays, we examined the localization of DNA replication proteins in relation to γ-H2AX, a marker for strand breaks, in HGPS patient fibroblasts. We found that XPA accumulation at DSB sites is age-dependent and progerin appears to directly interact with proliferating cell nuclear antigen (PCNA), which sequesters PCNA from stalled replication forks. The mislocalization of XPA to DSBs in addition to sequestration of PCNA by progerin prevents repair of DSBs in HGPS cells. Additionally, we found that the XPA-DNA interaction is regulated by ATR. Our data also suggests that the regulation by ATR in HGPS also may affect the expression of RNR in old-age HGPS. Interestingly, we also found that co-expression of progerin and XPA in XPA deficient cells recused the cells from apoptosis. These data suggest that progerin sequesters PCNA and RFC1 is proteolyticly cleaved resulting in a naked fork at the stalled replication fork and as a result XPA binds the DNA junction.

Materials and Methods

Cell Culture and Drug Treatment

Fibroblasts from two HGPS patients with the mutation at codon 608 (G608G) within exon 11 of the LMNA gene (AG11513A and AG11498B designated as HGPS-1 and HGPS-2, respectively) were obtained from Coriell Cell Repository. HGPS patient cells also were obtained from the Progeria Research Foundation with the point mutation 1824C→T (HGADFN178 designated as HGPS-4). Human RD fibroblasts were a gift from Dr. J. H. Miner (Washington University School of Medicine, St. Louis, MO, USA). The normal human fibroblasts, BJ cells, were purchased from American Type Culture Collection (ATCC, Rockville, MD, USA, no. CRL-2522). Normal fibroblasts also were obtained from the Progeria Research Foundation harvested from the father of HGADFN178 (designated as 090). HeLa cells were purchased from American Type Culture Collection (ATCC, CCL-2). All cultures were maintained in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% or 15% fetal bovine serum, secondary and primary cells respectively, and antibiotics (50 units/ml penicillin and 50 μg/ml streptomycin) at 37°C under an atmosphere containing 5% CO_2. To induce DNA damage cells were grown to 70% confluence, followed by incubation with 10 μM CPT for 6 hours, no release. Cells were treated with UV-C irradiation using a 254 nm lamp at a flounce of 0.83 J/m^2/ sec (20 J/m^2) followed by a 6-hour release period before harvesting. The ATR kinase inhibitor NU6027 was purchased from EMD Millipore (189299) and was used at a final concentration of 10 μM in cell culture medium. The ATM kinase inhibitor KU-55933 was purchased from Abcam (ab120637) and was used at a final concentration of 10 μM in cell culture medium.

Transfection of siRNA and Plasmids

To knockdown XPA and PCNA by RNAi, cells were each transfected with a pool of siRNAs (Santa Cruz, sc-36853 and sc-29441, respectively) or a scrambled sequence as a control using INTEFERin transfection reagent Polyplus-transfection SA (France) following the manufacturer's instructions. Further analyses were performed 72 hours after transfection. LAΔ50 constructs were a gift from Dr. Antonio Rusiñol. Plasmid transfection was carried out using jetPEI transfection reagent Polyplus-transfection SA (France) following the manufacturer's instructions. Further analyses were performed 48 hours after transfection.

Chromatin Fractionation Assay

Cellular localization of XPA in HGPS cells was investigated by the chromatin fractionation technique based on the method of Mendez and Stillman (Méndez and Stillman, 2000). Briefly, HGPS cells were plated in a 100-mm dish. At 80% confluence, cells were washed twice with ice-cold PBS and the cell pellet was resuspended in 100 µl of Solution A (10 mM Hepes–KOH pH 7.9, 10 mM KCl, 1.5 mM $MgCl_2$, 0.34 M sucrose, 1X protease inhibitor, 1mM PMSF and 0.1% TritonX-100). After 5 minutes incubation on ice, samples were centrifuged at 1300 g for 4 minutes. The pellet was washed once in Solution A and designated as intact nuclei. Nuclei were resuspended in 100 µl of Solution B (3 mM EDTA, 0.2 mM EGTA and 1X protease inhibitor cocktail) and incubated on ice for 10 minutes. The pellet was washed once with Solution B, and designated as the chromatin fraction.

Western Blotting

Cells cultured in 100-mm dishes or 25 cm^2 flasks were grown to 70% confluence and then trypsinized. Cell number was counted using a hemocytometer. The cells were centrifuged at 500 *g* for 5 minutes and washed twice with PBS in the presence of protease inhibitor cocktail. The cell pellet was lysed in 2x SDS gel loading buffer and volumes corresponding to $5*10^6$ cells were subjected to SDS-PAGE. Electrophoresis was carried out in 8 or 4-12% gels (BioRad). Immunoblotting was performed as previously described with primary antibodies directed against XPA (MA5-13835, ThermoFisher), lamin A/C (2032, Cell Signaling), GAPDH (MA5-15738, ThermoFisher), β-actin (A5316, Sigma-Aldrich), PCNA (sc-9857, Santa Cruz), and Pol δ (sc-17776, Santa Cruz). Proteins were transferred to PVDF membranes (Santa Cruz, USA) and immunoblotting was performed with the indicated antibodies followed by HRP-conjugated secondary antibodies (Santa Cruz). The complexes were revealed using SuperSignal West Pico Chemiluminescent Substrate (Thermo Scientific, Pierce).

Immunofluorescence microscopy

Cells grown on coverslips were fixed with cold methanol (-20°C) or extracted with 0.5% IGEPAL CA-630 followed by fixation with methanol. The fixed cells then were incubated with primary antibodies against γ-H2AX (rabbit, Bethyl; or mouse, Stressgen), XPA (mouse, ThermoFisher; or rabbit, Santa Cruz Biotechnology, Santa Cruz, CA, USA), MCM7 (mouse, Santa Cruz Biotechnology), and PCNA (rabbit, Santa Cruz Biotechnology). Secondary antibodies used include Alexa fluor 488-conjugated donkey anti-rabbit IgG and Alexa fluor 568-conjugated goat anti-mouse IgG (Molecular Probes, Eugene, OR, USA). Cells

were counterstained with DAPI to visualize nuclear DNA. Two blinded observers who randomly chose 100 cells for each experiment performed counting of foci.

Duolink – *in situ* Proximity Ligation Assays

For Duolink microscopic detection of protein interactions cells were grown on coverslips before the initiation of experimental treatments. Cells were fixed with 4% paraformaldehyde and permeablized with 0.2% Triton X-100 in PBS and blocked with 3% bovine serum albumin in PBS for 1 hr at room temperature. Proteins then were detected with primary antibodies and the Duolink assay was carried out according to the company's protocol (Sigma, DUO92101). Images were taken using the Life Technologies Evos microscope or an Olympus confocal microscope.

γ-H2AX association assay and coimmunoprecipitation

The γ-H2AX association assay used in this study was modified from the histone association assay described by Ricke *et al.* (Ricke and Bielinsky, 2005). Briefly, cells were treated with formaldehyde to cross-link interacting protein-DNA as well as protein-protein complexes. Nuclei were prepared by fractionation. The chromatin was sheared into 200-1500 bp fragments by sonication. The sheared chromatin was incubated with γ-H2AX antibody, followed by precipitation with protein G-sepharose beads. The immunoprecipitates were boiled for at least 30 min to reverse the crosslinks (Ricke and Bielinsky, 2005). Proteins that coprecipitate with γ-H2AX bound chromatin were detected

by Western blotting. The coimmunoprecipitation was performed using a Nuclear Complex Co-IP kit (Active Motif, Carlsbad, CA, USA), following the manufacturer's instructions.

Results

We first examined the status of PCNA-chromatin binding in BJ, RD and HGPS fibroblasts by probing γ-H2AX focus formation using immunofluorescence (IF) microscopy. BJ cells are considered normal human diploid fibroblasts and were used in this study as a control. The BJ cells were treated with UV-C at 20 J/m^2 to induce DNA damage. UV was used because the damage caused by UV irradiation causes replication-dependent strand breaks much like the breaks found in HGPS as a result of progerin accumulation. γ-H2AX is a marker used for single-strand and double-strand DNA breaks. In UV-treated BJ cells there is an accumulation of γ-H2AX foci that correlates with the formation of PCNA foci (Figure 1A). These foci also are found to colocalize meaning that PCNA and γ-H2AX are at the same sites in the nucleus of DNA damaged BJ cells. In these cells PCNA foci formation is consistent with active DNA synthesis, which in the treated cells is a result of DNA repair synthesis. In both the RD and HGPS cells γ-H2AX is observed to form distinct foci without DNA-damage treatment. Consistent with previous data, γ-H2AX foci are accumulated in RD and HGPS cells as a consequence of progerin-induced DSB formation (Liu et al., 2006; Liu et al., 2008). PCNA, however, does not form discrete foci in HGPS and RD cells. This indicates that PCNA is not bound at DSB sites and indicates there is little to no repair synthesis occurring in the progeroid cells. Based on previous data in our lab that showed XPA binds at DSB sites in HGPS we next examined the nuclear localization of PCNA and XPA (Liu et al., 2008). In the young, replicating HGPS cells PCNA forms discrete foci indicating that

replication is active and XPA is diffuse throughout the nucleus (Figure 1B). Being a damage-verification protein XPA is only expected to form foci when there is DNA damage. Additionally, XPA is known to accumulate at DSB sites in HGPS; however, it is unlikely that these young-age cells have accumulated many DSBs. When compared to the old-aged HGPS cells PCNA is no longer found in foci; rather, PCNA is spread throughout the nucleus much like XPA in the young cells. The same is true for XPA, where in the young cells there was diffuse staining and in the old cells there is formation of discrete XPA foci. The accumulation of XPA at foci in the nucleus in old-age HGPS cells is consistent with our previous study (Liu et al., 2008) as is loss of PCNA at the chromatin (Tang et al., 2012). In both the young and old cells there is no colocalization of PCNA and XPA suggesting that these proteins are not in the same location in the nucleus. Finally, we investigated if XPA binding seen in HGPS is located at DSBs that are the result of stalled replication forks. In old-aged HGPS cells we found that the γ-H2AX colocalized with MCM7 foci. MCM7 (Minichromosome Maintenance Complex Component 7) is the DNA helicase complex that is one factor making up the pre-replication complex and is thought to be involved in the initiation of replication (Chibazakura et al., 2011). γ-H2AX colocalization with MCM7 suggests that the DSBs in HGPS cells are the result of stalled replication fork collapse. Next, XPA and MCM7 foci localization was observed (Figure 1C). MCM7 foci colocalize with XPA, further supporting the finding that XPA binds at DSBs resulting from stalled replication

Figure 1

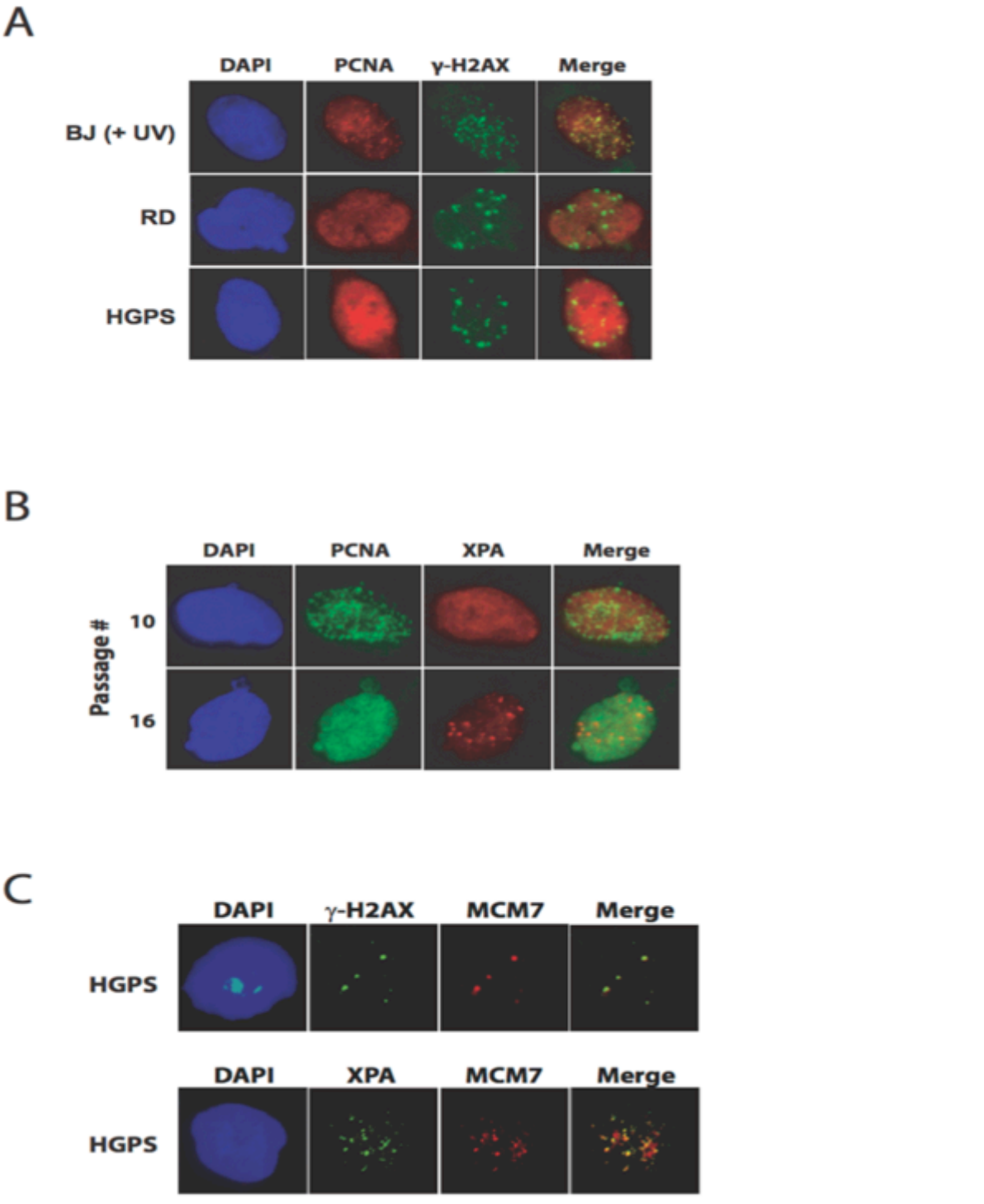


Figure 1. Replacement of PCNA by XPA at stalled replication forks leads to a lack of repair. A) Colocalization of γ-H2AX foci and the foci of PCNA in BJ (UV-treated), HGPS and RD fibroblasts. B) Nuclear focus localization of PCNA and XPA in early and late passage HGPS cells. C) Colocalization of γ-H2AX foci and the foci of MCM7 or XPA in late passage HGPS cells. Antigens were detected with corresponding antibodies by immunofluorescence microscopy. Uniform staining of the given protein throughout nucleus indicates homogenous distribution of the protein without foci formation.

fork collapse. Importantly, these data demonstrate that XPA is functioning outside of its known role as a bulky DNA-damage verification protein.

Next, we hypothesized that the slow replication rate of HGPS cells observed previously would affect RNR subunit turnover. As expected, the levels of RRM1 decrease as HGPS cells transition from rapidly growing (60% confluency) to continuous monolayer (100% to 150% confluency), a quiescent state (Figure 2A). This pattern applied to younger (p9) and older (p15) HGPS cells. In contrast, the level of RRM2 remained constant in younger cells of increasing confluency but increased 2.5x in the older cells as they settled into confluency. Interestingly, the level of p53RRM2 increased in young and old HGPS cells as they entered confluency, but significantly decreased in cells grown to quiescence. These data suggest that the endogenous DNA damage and the absence of PCNA and mislocalization of XPA observed in HGPS cells upset the normal turnover of the RNR subunits. Additionally, the increase in p53M2 suggests a transition from replication to repair synthesis.

We found that (6-4) PP (photoproducts) are repaired rapidly (data not shown), with no detectable differences between normal BJ fibroblasts and young or old HGPS fibroblasts all treated with UV-C at 20 J/m^2. Whereas, CPD adducts are repaired quickly in the transformed A549 lung carcinoma cells but more slowly in BJ cells and most slowly in the HGPS cells (Figure 2B). These differences indicate that in NER the repair of (6-4) PP is more crucial for short-term survival that CPD repair. These data demonstrate the UV-induced damage in HGPS cells is not repaired as quickly when compared to normal fibroblasts and A549 cells. Taken together, not only is replication affected in aging HGPS cells but the repair also is delayed.

Another question we addressed is what effect does a prolonged cell cycle length, prior to senescence, have on the DNA synthesis parameters of the cell cycle. To answer this question we used a Click-IT EdU kit (Figure 2C). The EdU was detected with the Alexa Flour 647 azide and the DNA stained with propidium iodide (PI) In young cells (PD16 → PD23) the bulk of the DNA synthesis occurs as genomic replication in the S-phase of the cell cycle. As the cells age there is a shift of DNA synthesis from S-phase to the G1-phase (PD23 → PD30 and 36) DNA repair. Thus, as DNA damage accumulates there is a shift from replication to DNA repair. This switching from replication to repair as the HGPS age in culture is in agreement with cell growth assays conducted with HGPS where the cells enter a senescent-like state in late passages.

To observe the PCNA bound to chromatin at sites of DNA DSBs we used a modified chromatin immunoprecipitation (ChIP) assay. Untreated HGPS cells were subjected to the ChIP followed by western blotting to detect the proteins bound at the sites of DSBs. HGPS cells also were treated with camptothecin (CPT), which is an inhibitor of the DNA enzyme topoisomerase I. When cells are treated with CPT there is an accumulation of ssDNA nicks, which then are processed into DSBs if not repaired, mimicking a stalled replication fork. The ChIP assay allows for IP of a particular target and if another protein is bound on the same small piece of chromatin that protein will be IPed as well. In the untreated HGPS cells after pulling-down with γ-H2AX the PCNA protein is not bound while MCM7 is bound (Figure 3A). This further demonstrates that PCNA does not bind DSB sites in HGPS cells. In the CPT-treated HGPS cells PCNA pulls-down with γ-H2AX. This suggests that PCNA is still capable of binding to exogenous DNA damage; however, PCNA is not capable of binding to

Figure 2

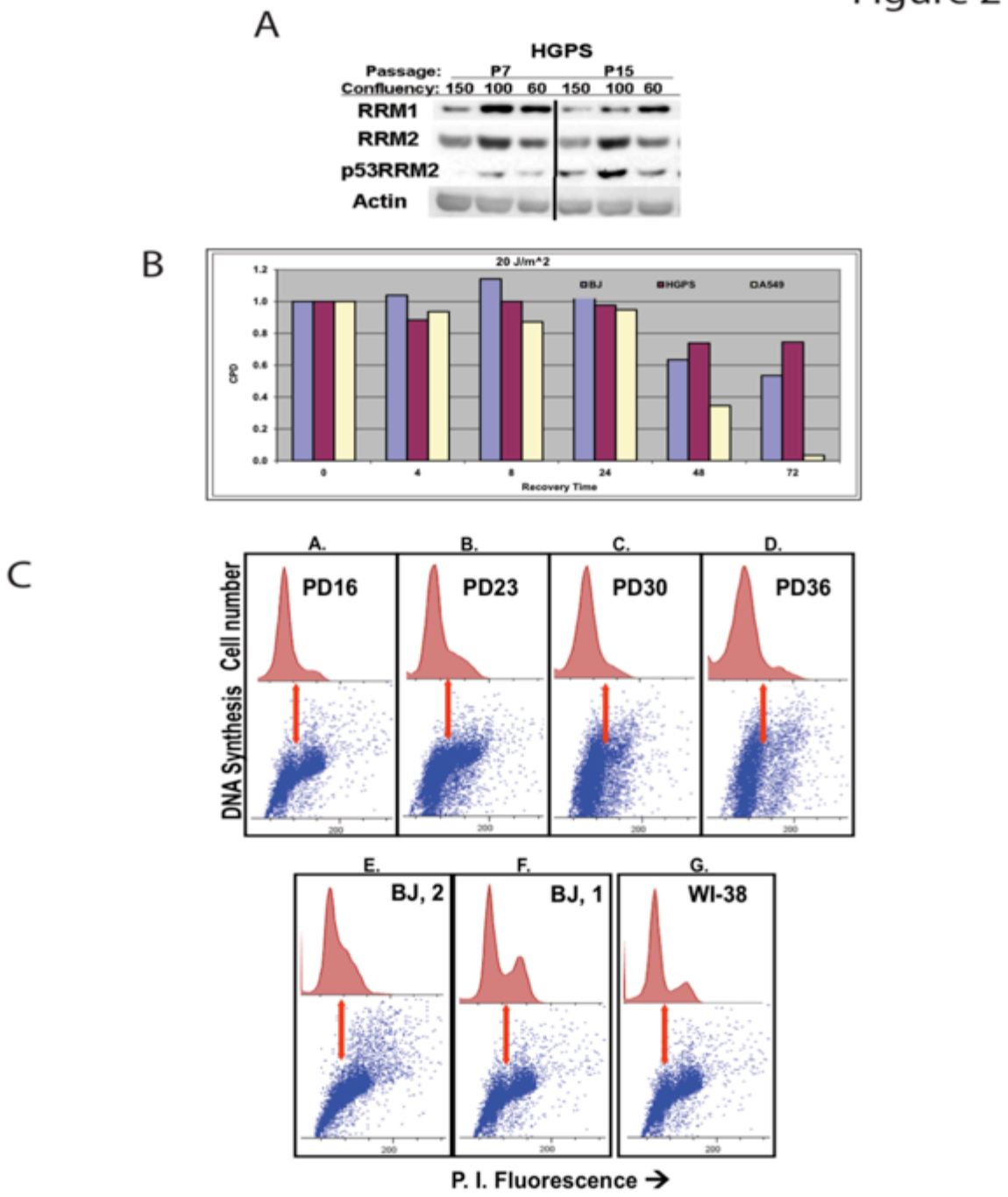


Figure 2. Replication abnormalities in HGPS. (A) Expression levels of RNR subunits were observed by western blotting. Cells were grown to the indicated confluence where 60% confluency is actively growing to 100% confluency is a continuous monolayer to 150% confluency is a quiescent state. (B) Cells were subjected to 20 J/m^2 UV-C followed by an ELISA assay to detect CPD repair as a function of time. (C) Newly synthesized DNA in culture-aged HGPS (A-D), BJ (E-F) and WI-38 (G) cells was labeled with 5-ethynyl-2'-deoxyuridine (EdU; Invitrogen). The labeling was for 1 hr with 20 μM EdU (A-D & F) or with 10 μM EdU for either 2 hr (E) or 4 hr (G). The cells were isolated, rinsed with PBS and processed as described in Invitrogen's "Click-IT EdU Manual". The EdU was detected with the Alexa Flour 647 azide (Invitrogen) and the DNA stained with propidium iodide (P.I.). The PD#s in panels A-to-D indicates the population doubling age of the HGPS cells.

endogenous DNA damage generated by progerin. When XPA is knocked down in HGPS cells using XPA specific siRNA, PCNA then can bind the endogenous DNA damage (Figure 3B). The replication polymerase, pol δ, also is found to bind the DSB site after XPA silencing. These data indicate that binding of XPA can prevent the binding of PCNA to the DSB sites in HGPS cells.

XPA binding to the DSBs in HGPS coincides with an absence of PCNA as shown here. Therefore, we speculated that if knockdown of XPA can restore PCNA binding, will XPA bind the DSBs generated by stalled replication forks when PCNA is knocked down? Silencing of PCNA was carried out in BJ cells by PCNA-specific siRNA. When PCNA is knocked down XPA forms discrete foci in untreated BJ cells, suggesting that XPA binding to these sites in normal cells as a novel function for XPA (Figure 4A). Next we used a proximity ligation assay (Duolink), in which two proteins must be within 30-40 nM proximity to generate a signal, to observe the interaction between PCNA and the nuclear lamina. In old-age HGPS cells there is little to no PCNA-lamin A binding. However, there is a huge increase in the amount of PCNA that is bound by progerin (Figure 4B). Western blot analysis demonstrates efficient knock down of PCNA. PCNA has been shown to interact with lamin A as part of replication regulation (Camozzi et al., 2014). PCNA-progerin interaction in old-age HGPS cells indicates that XPA binding to DSBs is a consequence of PCNA sequestration by progerin. Taken together, when the DNA replication machinery is absent, PCNA sequestration by progerin and RFC1 proteolytic cleavage, results in XPA binding of DSB sites in HGPS.

Figure 3

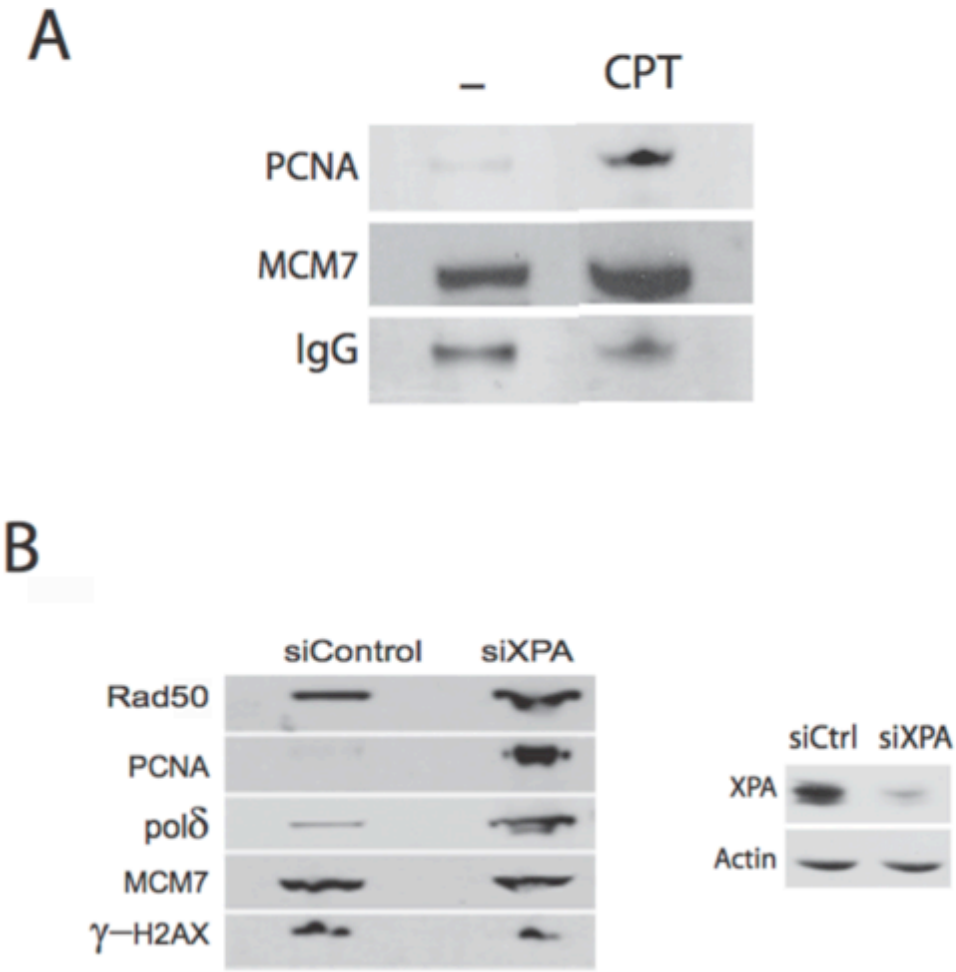


Figure 3. XPA-associated DNA double-strand breaks form at replication forks in HGPS. (A) HGPS cells treated with or without CPT were subjected to the modified ChIP assay with anti-γ-H2AX antibody for analysis by Western blotting. PCNA is largely absent at the replication forks with progerin-induced, but not CPT-induced, DSBs. (B) HGPS cells were transfected with XPA siRNA in the absence and presence of CPT treatment, followed by analysis using the modified ChIP assays with anti-γ-H2AX antibody. The XPA siRNA knockdown significantly restores the binding of PCNA and pol δ to the replication forks in the progeroid cells.

XPA accumulation at chromatin is age-dependent in HGPS cells as demonstrated using a chromatin association assay. In control BJ cells XPA is found in the cytoplasm and nucleoplasm while there is no XPA bound to the chromatin (Figure 5A, left). The chromatin-bound XPA was quantified and normalized to free XPA in the cells (Figure 5A, right). XPA binding to the chromatin is likely a result of PCNA sequestration by accumulation of progerin. The checkpoint kinase ATR has been previously shown to regulate XPA in response to damage (Shell et al., 2009). By using inhibitors specific to ATM (KU-55933) and ATR (NU6027) we observed that inhibition of ATR drastically reduced the amount of chromatin-bound XPA in HGPS cells (Figure 5B, left). The two inhibitors together reduced the chromatin-bound XPA even further, which may indicate additional regulation of XPA-chromatin binding. Whole cell lysates were generated using the same inhibitors and the downstream target of both ATM and ATR, p53 S15 phosphorylation, was detected (Figure 5B, right). These data confirmed that the inhibitors efficiently inhibited the kinase activity of both proteins.

We hypothesized that XPA functions to prevent apoptosis in HGPS cells. To investigate this hypothesis we used XPA deficient patient cells and expressed LAΔ50 (Figure 6A). Expression of LAΔ50 in $XPA^{-/-}$ cells more than doubled the number of cells that were killed via apoptosis when compared to $XPA^{-/-}$ cells expressing GFP control. Apoptosis was measured by caspase-3 activity using a fluorogenic substrate (Ac-DEVD-AFC). The relative fluorescence units were normalized to the protein concentration of the sample. By then expressing WT-XPA in both of these cells the amount of apoptosis is drastically reduced in the LAΔ50-expressing cells and returns to near that of the GFP-expressing cells. A high dose of CPT was used as a control for DNA damage and greatly increased the

Figure 4

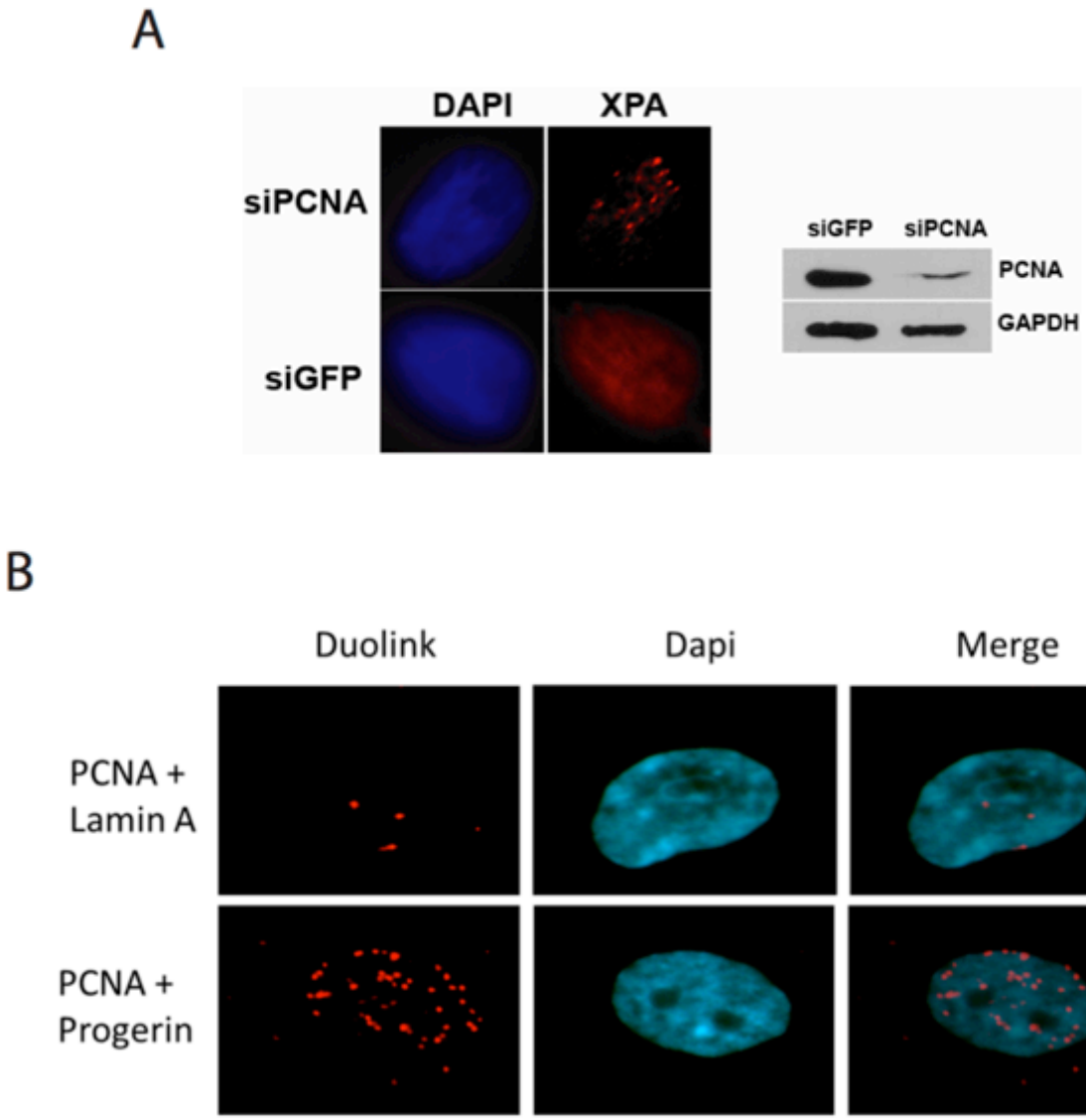


Figure 4. PCNA is sequestered by progerin in HGPS cells. (A) XPA nuclear focus formation in normal BJ cells with PCNA knockdown by PCNA-specific siRNA (left). Western blot representing the efficiency of PCNA knockdown (right). (B) Proximity ligation assay using PCNA and either lamin A- or progerin-specific antibodies. Following primary antibody incubation secondary antibodies conjugated to complimentary oligonucleotides that then are bound, ligated, and amplified to generate a red signal indicating interaction.

number of cells that enter apoptosis. These data suggest that XPA may play a role in preventing apoptosis in progeroid cells. Confirmation of the plasmid transfection can be seen in Figure 6B. XPA$^{-/-}$ cells expressing LAΔ50 are observed to accumulate both degraded lamins and an increase in PARP cleavage. This is indicative of cells that are initiating apoptosis. The apoptosis assay and western blot analysis demonstrates that XPA may play a role in preventing apoptosis in progeroid cells and this function of XPA is independent of its role as a DNA-damage verification protein in GG-NER.

Discussion

DSBs accumulate in HGPS cells and RD cells and as a result the cells have high levels of genome instability that might contribute to premature aging (Liu et al., 2005; Liu et al., 2006; Manju et al., 2006; Musich and Zou, 2009). It has been suggested that the accumulation of DSBs is due to a deficiency in DNA repair in progeria cells (Liu et al., 2008; Ukekawa et al., 2007). Previous studies from our lab found that DSB repair proteins Rad51 and Rad50 were absent at the progerin-induced DNA damage sites in progeria cells (Liu et al., 2008). The DSBs as a result of progerin were resistant to repair in the progeria cells; however, repair of CPT-induced DNA damage, although slower when compared to BJ cells, was still effective (Liu et al., 2008). Unexpectedly, the NER protein XPA was found to form nuclear foci that colocalize with γ-H2AX or DSBs. This mislocalization correlated with the inefficiency in DNA repair. XPA accumulation occurred exclusively in progeria cells but not in normal fibroblasts. RNAi knockdown of XPA in HGPS cells could restore DSB repair suggesting that the mislocalization of XPA to or near the laminopathy-induced DSB sites

Figure 5

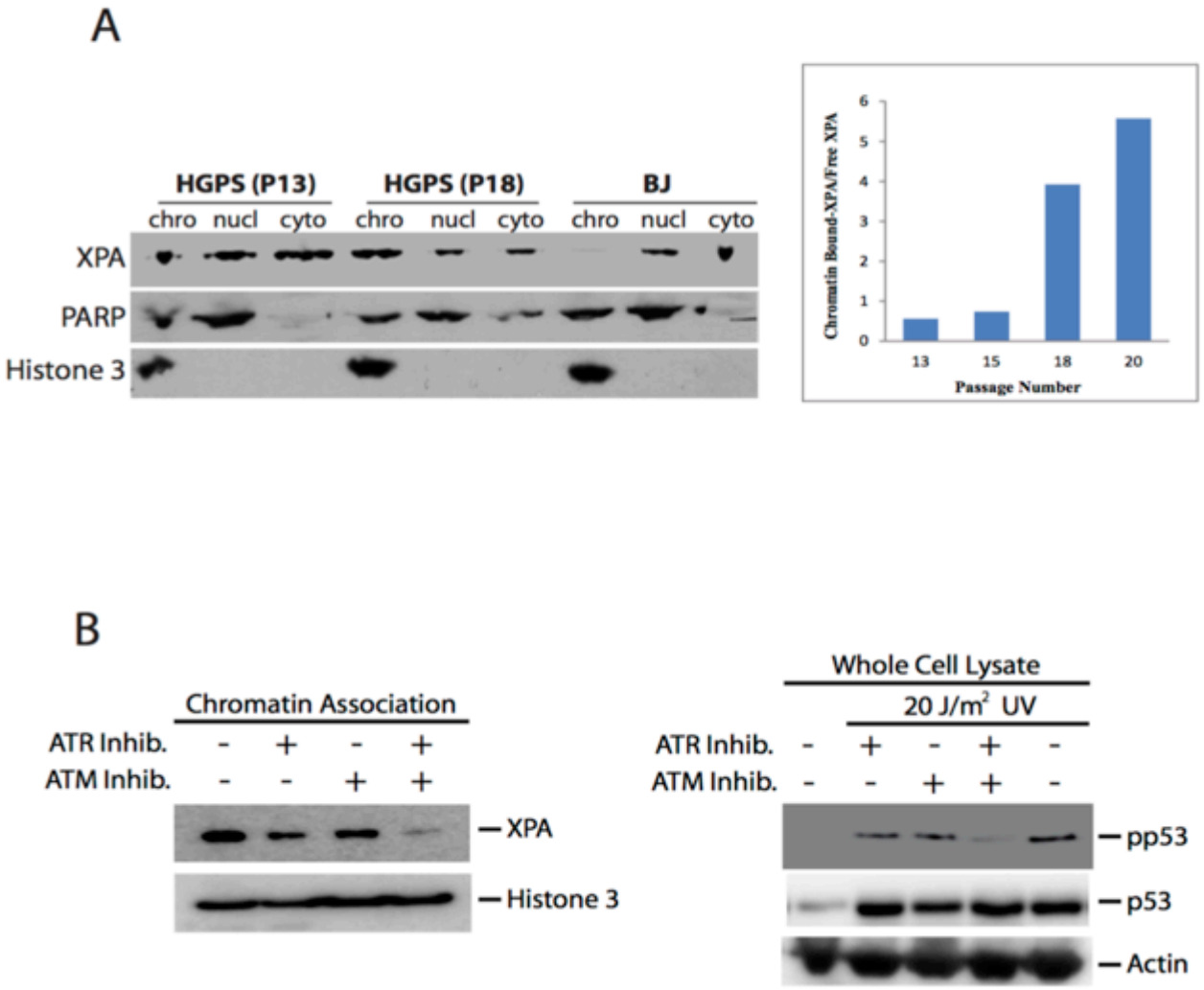


Figure 5. XPA chromatin accumulation in HGPS is age-dependent and regulated by ATR. (A) Chromatin association assay of BJ, young, and old HGPS fibroblasts. PARP was used as a control for the nuclear fraction and Histone 3 was used as a control for the "chromatin bound" protein (left panel). The amount of chromatin-bound XPA was quantified then normalized to free-XPA in the cell (right panel). (B) Chromatin association assay of old-age HGPS fibroblasts (left panel). ATR/ATM inhibitors were added 1 hr prior at 10 uM then the chromatin was collected or inhibitor treatment was followed by 20 J/m^2 UV-C treatment (right panel). Whole cell lysates were generated from BJ cells.

may play a role in blocking the accessibility of the damage sites to DSB-repair factors, thus inhibiting DNA repair. Another recent finding by our lab suggests that XPA can bind at ds-ssDNA junctions at higher affinity than at bulky DNA adducts and this binding is through an extended DNA-binding domain (Hilton et al., 2014; Yang et al., 2006). The study presented here suggests that XPA binding to stalled replication forks is the result of a sequestration of PCNA by direct interaction with progerin and XPA binding prevents apoptosis in HGPS cells.

The verification role of XPA in GGR-NER has been extensively studied (Nouspikel, 2009; Riedl et al., 2003; Sugasawa et al., 2001; Volker et al., 2001). XPA has not been found to play any role in DSB repair previously. Our past results indicated that the association of XPA with γ-H2AX is mediated by chromatin. Nuclear lamins directly interact with histones such as H2A; however, nuclear lamins also interact with DNA synthesis proteins such as PCNA (Mattout et al., 2007; Shumaker et al., 2008). The structures that result from stalled replication forks, like the ones seen in progeroid cells, are like ds-ssDNA junctions, which XPA is known to bind (Hilton et al., 2014; Sugitani et al., 2014; Yang et al., 2006). However, the effects of XPA binding at these sites in the cell are unclear. Here we demonstrate that in addition to XPA colocalization with γ-H2AX, or strand breaks, XPA colocalizes with MCM7, a subunit of the helicase complex involved in DNA synthesis (Wang et al., 2004). This colocalization links XPA to the stalled replication forks in HGPS.

Figure 6

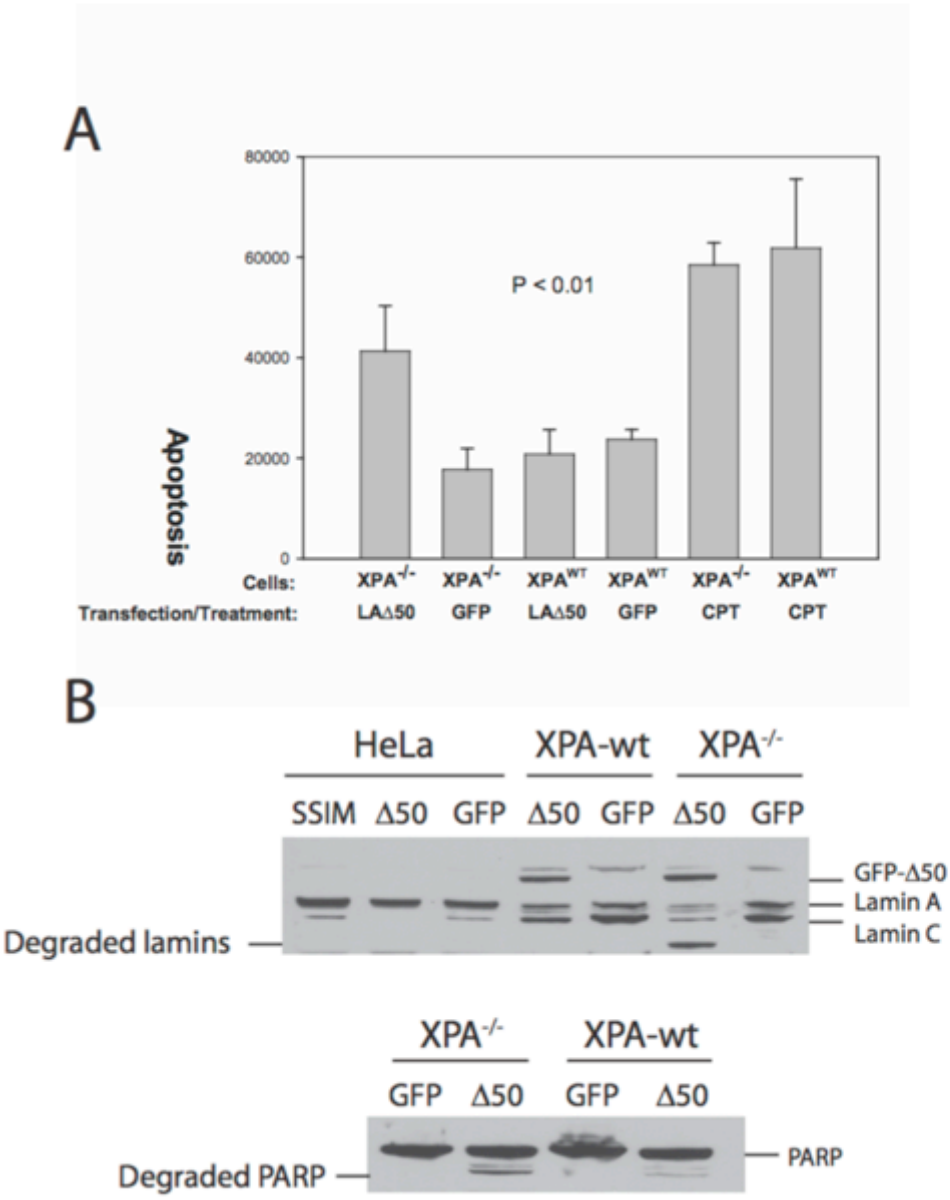


Figure 6. XPA prevents apoptosis in progeroid cells. (A) XPA^{WT} and XPA$^{-/-}$ cells were transfected with the plasmid expressing progerin (LA50) or empty parent vector (GFP). 24 hours after transfection, cells were harvested. The caspase-3 activity was measured by using a fluorogenic substrate (Ac-DEVD-AFC). The relative fluorescence units were normalized to the protein concentration of the sample.
(B) HeLa , XPA^{WT} and XPA$^{-/-}$ cells were transfected with the plasmid expressing progerin (LA50) or empty parent vector (GFP). 24 hours after transfection, cells were harvested. In the top panel, the blot was probed an antibody against laminA/C. In the bottom panel, the blot was probed with an antibody against PARP.

One potential outcome of XPA binding at the stalled forks is loss of DNA replicaiton proteins. It has been reported that disruption of lamin organization by injecting lamin A mutant protein into mammalian cells resulted in redistribution of PCNA and RFC, thus blocking the elongation phase of DNA replication (Moir et al., 2000; Spann et al., 1997). It has been proposed that the replication blockage, which may induce DSBs, is likely due to the loss of replicative PCNA trapped in lamin A aggregates (Moir et al., 2000; Spann et al., 1997). Therefore, a possible loss of PCNA at the collapsed replication forks in HGPS cells may leave the ds-ssDNA junctions unprotected. Indeed, by using immunofluorescence we observed that PCNA transitions from distinct foci formation to being spread throughout the nucleus in an age-dependent manner in HGPS. PCNA previously was observed to form foci in replicating cells (Leonhardt et al., 2000). Through proximity ligation assay we also found that in old age HGPS cells there is no longer interaction of PCNA and the lamina but there is interaction between PCNA and progerin. Furthermore, in normal fibroblast in which PCNA is knocked down XPA forms foci.

Likely, the DNA junctions near DSBs at the stalled replication forks are bound by XPA as a result of PCNA sequestration by progerin (Figure 7). In support of this conclusion, knocking down XPA in HGPS results in a re-association of DNA replication proteins at stalled replication forks. However, it is important to note that progerin sequestered PCNA is not likely to re-associate. It has been proposed that HGPS are characterized by hyperproliferation and an increase in apoptosis (Bridger and Kill, 2004). The question remains, if XPA is not present in HGPS cells will even more cells die of apoptosis. Indeed, in XPA-deficient patient cells expressing progerin there is a huge increase in the number of cells that die of apoptosis compared to control. Next, we used XPA-deficient patient cells

and found that expressing XPA in the deficient cells rescues the cells from apoptotic cell death. This suggests that XPA can serve to prevent apoptosis by binding naked junctions and preventing fork collapse in HGPS.

ATR is a major regulator of the DNA damage response and checkpoint signaling (Cimprich and Cortez, 2008). ATR can phosphorylate many downstream targets to initiate the checkpoint, halting replication, and has been observed to be involved in promoting structural stability of stalled forks by preventing dissociation of replisome components (Errico and Costanzo, 2012). ATR also can phosphorylate XPA in response to UV-induced DNA damage (Shell et al., 2009; Wu et al., 2006). Since ATR can regulate XPA following damage we asked if ATR might play a role in regulating XPA binding at stalled replication forks. By inhibiting ATR kinase activity using an ATR specific inhibitor we found that ATR kinase inhibition prevents XPA binding to the chromatin in a chromatin association assay. Inhibition of ATR in HGPS cells resulting in release of XPA from the stalled replication forks would likely lead to HGPS cell death via apoptosis. Interestingly, ATR recently has been shown to suppress DNA damage by coordinating RRM2 accumulation at replication origins (Buisson et al., 2015). In old-age HGPS cells there is an accumulation of p53RRM2 and a marked decrease in RRM1 expression. Furthermore, increasing RNR activity in progeroid mice alleviates the severity of the premature aging phenotype (Lopez-Contreras et al., 2015). RNR also has been observed to promote fork restart following replicative stress (Morafraile et al., 2015). RRM1 is turned over rapidly as a cell replicates its genome; however, it is interesting to speculate that ATR also may suppress RRM2 accumulation in HGPS, which could promote accumulation of the repair subunit p53RRM2 leading to the down regulation of RRM1 and fork restart.

The findings presented in this study represent another step in uncovering the underlying mechanism of DSB accumulation seen in the premature aging diseases HGPS and RD. Our results suggest a role for XPA in prevention of apoptosis in HGPS cells by binding naked replication forks that are the result of PCNA sequestration by progerin. Treatment of HGPS patients with pravastatin and zoledronic acid has been successful in extending the lifespan of mice with HGPS-like phenotypes (Varela et al., 2008) and HGPS patients (Gordon et al., 2014). Our lab found that treatment with these drugs can prevent the proteolytic cleavage of the RFC1 protein which is required in DNA replication and repair (McAlear et al., 1996; Tang et al., 2012). An interesting future study would be to use pravastatin and zoledronic acid in combination with farnesyltransferase inhibitors (FTIs) to observe the effects on not only DSB accumulation but also XPA and PCNA binding at stalled forks. A drug combination that can effectively inhibit DNA damage accumulation and improve nuclear morphology would be revolutionary in treatment of progeroid diseases. DNA damage accumulation has been linked to normal aging as well as progerin accumulation (Pacheco et al., 2014; Schumacher et al., 2008). Nuclear lamin defects also play a role in normal human aging (Scaffidi and Misteli, 2006). Further investigating the mechanisms behind the premature aging diseases could lead to development of novel treatments to alleviate some of the symptoms involved in aging.

Figure 7

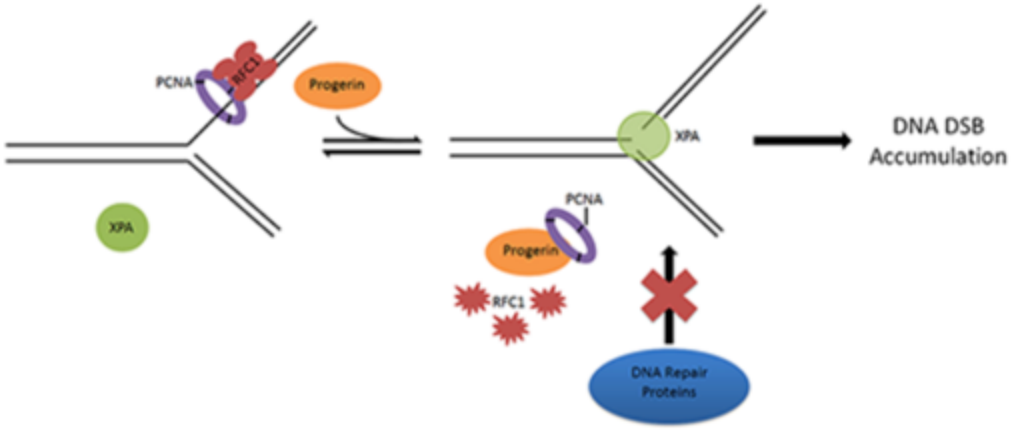


Figure 7. Proposed model for XPA binding at stalled replication forks as a result of PCNA sequestration by progerin and proteolytic cleavage of RFC1.

References

Aye, Y., Li, M., Long, M. J. C. and Weiss, R. S. (2014). Ribonucleotide reductase and cancer: biological mechanisms and targeted therapies. *Oncogene*.

Bomgarden, R. D., Lupardus, P. J., Soni, D. V., Yee, M.-C., Ford, J. M. and Cimprich, K. A. (2006). Opposing effects of the UV lesion repair protein XPA and UV bypass polymerase eta on ATR checkpoint signaling. *The EMBO Journal* **25**, 2605–2614.

Bridger, J. M. and Kill, I. R. (2004). Aging of Hutchinson-Gilford progeria syndrome fibroblasts is characterised by hyperproliferation and increased apoptosis. *Exp. Gerontol.* **39**, 717–724.

Buisson, R., Boisvert, J. L., Benes, C. H. and Zou, L. (2015). Distinct but Concerted Roles of ATR, DNA-PK, and Chk1 in Countering Replication Stress during S Phase. *Mol Cell* **59**, 1011–1024.

Camozzi, D., Capanni, C., Cenni, V., Mattioli, E., Columbaro, M., Squarzoni, S. and Lattanzi, G. (2014). Diverse lamin-dependent mechanisms interact to control chromatin dynamics. **5**, 427–440.

Chibazakura, T., Kamachi, K., Ohara, M., Tane, S., Yoshikawa, H. and Roberts, J. M. (2011). Cyclin A promotes S-phase entry via interaction with the replication licensing factor Mcm7. *Mol Cell Biol* **31**, 248–255.

Cimprich, K. A. and Cortez, D. (2008). ATR: an essential regulator of genome integrity. *Nat Rev Mol Cell Biol* **9**, 616–627.

Corrigan, D. P., Kuszczak, D., Rusinol, A. E., Thewke, D. P., Hrycyna, C. A., Michaelis, S. and Sinensky, M. S. (2005). Prelamin A endoproteolytic processing in vitro by recombinant Zmpste24. *Biochem J* **387**, 129–138.

Cox, M. M. (2002). The nonmutagenic repair of broken replication forks via recombination. *Mutat Res* **510**, 107–120.

De Sandre-Giovannoli, A., Bernard, R., Cau, P., Navarro, C., Amiel, J., Boccaccio, I., Lyonnet, S., Stewart, C. L., Munnich, A., Le Merrer, M., et al. (2003). Lamin a truncation in Hutchinson-Gilford progeria. *Science* **300**, 2055.

Eriksson, M., Brown, W. T., Gordon, L. B., Glynn, M. W., Singer, J., Scott, L., Erdos, M. R., Robbins, C. M., Moses, T. Y., Berglund, P., et al. (2003). Recurrent de novo point mutations in lamin A cause Hutchinson-Gilford progeria syndrome. *Nature* **423**, 293–298.

Errico, A. and Costanzo, V. (2012). Mechanisms of replication fork protection: a safeguard for genome stability. *Crit. Rev. Biochem. Mol. Biol.* **47**, 222–235.

Goldman, R. D., Gruenbaum, Y., Moir, R. D., Shumaker, D. K. and Spann, T. P. (2002). Nuclear lamins: building blocks of nuclear architecture. *Genes Dev* **16**, 533–547.

Gonzalo, S. and Kreienkamp, R. (2015). DNA repair defects and genome instability in Hutchinson-Gilford Progeria Syndrome. *Current Opinion in Cell Biology* **34**, 75–83.

Gordon, L. B., Rothman, F. G., López-Otín, C. and Misteli, T. (2014). Progeria: A Paradigmfor Translational Medicine. *Cell* **156**, 400–407.

Heller, R. C. and Marians, K. J. (2006). Replisome assembly and the direct restart of stalled replication forks. *Nat Rev Mol Cell Biol* **7**, 932–943.

Hilton, B., Shkriabai, N., Musich, P. R., Kvaratskhelia, M., Shell, S. and Zou, Y. (2014). A new structural insight into XPA-DNA interactions. *Biosci. Rep.* **34**, e00162–840.

Hutchison, C. J. (2011). The role of DNA damage in laminopathy progeroid syndromes. *Biochem. Soc. Trans.* **39**, 1715–1718.

Jarrett, S. G., Wolf Horrell, E. M., Christian, P. A., Vanover, J. C., Boulanger, M. C., Zou, Y. and D'Orazio, J. A. (2014). PKA-mediated phosphorylation of ATR promotes recruitment of XPA to UV-induced DNA damage. *Mol Cell* **54**, 999–1011.

Leonhardt, H., Rahn, H. P., Weinzierl, P., Sporbert, A., Cremer, T., Zink, D. and Cardoso, M. C. (2000). Dynamics of DNA replication factories in living cells. *J Cell Biol* **149**, 271–280.

Liu, B., Wang, J., Chan, K. M., Tjia, W. M., Deng, W., Guan, X., Huang, J.-D., Li, K. M., Chau, P. Y., Chen, D. J., et al. (2005). Genomic instability in laminopathy-based premature aging. *Nat. Med.* **11**, 780–785.

Liu, Y., Rusinol, A., Sinensky, M., Wang, Y. and Zou, Y. (2006). DNA damage responses in progeroid syndromes arise from defective maturation of prelamin A. *J Cell Sci* **119**, 4644–4649.

Liu, Y., Wang, Y., Rusinol, A. E., Sinensky, M. S., Liu, J., Shell, S. M. and Zou, Y. (2008). Involvement of xeroderma pigmentosum group A (XPA) in progeria arising from defective maturation of prelamin A. *The FASEB Journal* **22**, 603–611.

Lopez-Contreras, A. J., Specks, J., Barlow, J. H., Ambrogio, C., Desler, C., Vikingsson, S., Rodrigo-Perez, S., Green, H., Rasmussen, L. J., Murga, M., et al. (2015). Increased Rrm2 gene dosage reduces fragile site breakage and prolongs survival of ATR mutant mice. *Genes Dev* **29**, 690–695.

Manju, K., Muralikrishna, B. and Parnaik, V. K. (2006). Expression of disease-causing lamin A mutants impairs the formation of DNA repair foci. *J Cell Sci* **119**, 2704–2714.

Mattout, A., Goldberg, M., Tzur, Y., Margalit, A. and Gruenbaum, Y. (2007). Specific and

conserved sequences in D. melanogaster and C. elegans lamins and histone H2A mediate the attachment of lamins to chromosomes. *J Cell Sci* **120**, 77–85.

McAlear, M. A., Tuffo, K. M. and Holm, C. (1996). The large subunit of replication factor C (Rfc1p/Cdc44p) is required for DNA replication and DNA repair in Saccharomyces cerevisiae. *Genetics* **142**, 65–78.

Méndez, J. and Stillman, B. (2000). Chromatin association of human origin recognition complex, cdc6, and minichromosome maintenance proteins during the cell cycle: assembly of prereplication complexes in late mitosis. *Mol Cell Biol* **20**, 8602–8612.

Mirkin, E. V. and Mirkin, S. M. (2007). Replication fork stalling at natural impediments. *Microbiol. Mol. Biol. Rev.* **71**, 13–35.

Moir, R. D., Spann, T. P., Herrmann, H. and Goldman, R. D. (2000). Disruption of nuclear lamin organization blocks the elongation phase of DNA replication. *J Cell Biol* **149**, 1179–1192.

Morafraile, E. C., Diffley, J. F. X., Tercero, J. A. and Segurado, M. (2015). Checkpoint-dependent RNR induction promotes fork restart after replicative stress. *Sci Rep* **5**, 7886.

Musich, P. R. and Zou, Y. (2009). Genomic instability and DNA damage responses in progeria arising from defective maturation of prelamin A. *Aging (Albany NY)* **1**, 28–37.

Musich, P. R. and Zou, Y. (2011). DNA-damage accumulation and replicative arrest in Hutchinson-Gilford progeria syndrome. *Biochem. Soc. Trans.* **39**, 1764–1769.

Navarro, C. L., Cadiñanos, J., De Sandre-Giovannoli, A., Bernard, R., Courrier, S., Boccaccio, I., Boyer, A., Kleijer, W. J., Wagner, A., Giuliano, F., et al. (2005). Loss of ZMPSTE24 (FACE-1) causes autosomal recessive restrictive dermopathy and accumulation of Lamin A precursors. *Human Molecular Genetics* **14**, 1503–1513.

Nouspikel, T. (2009). DNA repair in mammalian cells : Nucleotide excision repair: variations on versatility. *Cell. Mol. Life Sci.* **66**, 994–1009.

Pacheco, L. M., Gomez, L. A., Dias, J., Ziebarth, N. M., Howard, G. A. and Schiller, P. C. (2014). Progerin expression disrupts critical adult stem cell functions involved in tissue repair. *Aging (Albany NY)* **6**, 1049–1063.

Rhéaume, E., Cohen, L. Y., Uhlmann, F., Lazure, C., Alam, A., Hurwitz, J., Sékaly, R. P. and Denis, F. (1997). The large subunit of replication factor C is a substrate for caspase-3 in vitro and is cleaved by a caspase-3-like protease during Fas-mediated apoptosis. *The EMBO Journal* **16**, 6346–6354.

Ricke, R. M. and Bielinsky, A.-K. (2005). Easy detection of chromatin binding proteins by the Histone Association Assay. *Biol Proced Online* **7**, 60–69.

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

Scaffidi, P. and Misteli, T. (2006). Lamin A-dependent nuclear defects in human aging. *Science* **312**, 1059–1063.

Schumacher, B., Garinis, G. A. and Hoeijmakers, J. H. J. (2008). Age to survive: DNA damage and aging. *Trends Genet.* **24**, 77–85.

Sedelnikova, O. A., Horikawa, I., Zimonjic, D. B., Popescu, N. C., Bonner, W. M. and Barrett, J. C. (2004). Senescing human cells and ageing mice accumulate DNA lesions with unrepairable double-strand breaks. *Nat Cell Biol* **6**, 168–170.

Shell, S. M., Li, Z., Shkriabai, N., Kvaratskhelia, M., Brosey, C., Serrano, M. A., Chazin, W. J., Musich, P. R. and Zou, Y. (2009). Checkpoint kinase ATR promotes nucleotide excision repair of UV-induced DNA damage via physical interaction with xeroderma pigmentosum group A. *J Biol Chem* **284**, 24213–24222.

Shumaker, D. K., Solimando, L., Sengupta, K., Shimi, T., Adam, S. A., Grunwald, A., Strelkov, S. V., Aebi, U., Cardoso, M. C. and Goldman, R. D. (2008). The highly conserved nuclear lamin Ig-fold binds to PCNA: its role in DNA replication. *J Cell Biol* **181**, 269–280.

Spann, T. P., Moir, R. D., Goldman, A. E., Stick, R. and Goldman, R. D. (1997). Disruption of nuclear lamin organization alters the distribution of replication factors and inhibits DNA synthesis. *J Cell Biol* **136**, 1201–1212.

Sugasawa, K., Okamoto, T., Shimizu, Y., Masutani, C., Iwai, S. and Hanaoka, F. (2001). A multistep damage recognition mechanism for global genomic nucleotide excision repair. *Genes Dev* **15**, 507–521.

Sugitani, N., Shell, S. M., Soss, S. E. and Chazin, W. J. (2014). Redefining the DNA-Binding Domain of Human XPA. *J. Am. Chem. Soc.* 140724070100004.

Tang, H., Hilton, B., Musich, P. R., Fang, D. Z. and Zou, Y. (2012). Replication factor C1, the large subunit of replication factor C, is proteolytically truncated in Hutchinson–Gilford progeria syndrome. *Aging Cell* **11**, 363–365.

Ukekawa, R., Miki, K., Fujii, M., Hirano, H. and Ayusawa, D. (2007). Accumulation of multiple forms of lamin A with down-regulation of FACE-1 suppresses growth in senescent human cells. *Genes Cells* **12**, 397–406.

Varela, I., Cadiñanos, J., Pendas, A. M., Gutiérrez-Fernández, A., Folgueras, A. R., Sánchez, L. M., Zhou, Z., Rodríguez, F. J., Stewart, C. L., Vega, J. A., et al. (2005). Accelerated ageing in mice deficient in Zmpste24 protease is linked to p53 signalling activation. *Nature* **437**, 564–568.

Varela, I., Pereira, S., Ugalde, A. P., Navarro, C. L., Suárez, M. F., Cau, P., Cadiñanos, J., Osorio, F. G., Foray, N., Cobo, J., et al. (2008). Combined treatment with statins and aminobisphosphonates extends longevity in a mouse model of human premature aging. *Nat. Med.* **14**, 767–772.

Volker, M., Moné, M. J., Karmakar, P., van Hoffen, A., Schul, W., Vermeulen, W., Hoeijmakers, J. H., van Driel, R., van Zeeland, A. A. and Mullenders, L. H. (2001). Sequential assembly of the nucleotide excision repair factors in vivo. *Mol Cell* **8**, 213–224.

Wang, X., Ira, G., Tercero, J. A., Holmes, A. M., Diffley, J. F. X. and Haber, J. E. (2004). Role of DNA replication proteins in double-strand break-induced recombination in Saccharomyces cerevisiae. *Mol Cell Biol* **24**, 6891–6899.

Wu, X., Shell, S. M., Yang, Z. and Zou, Y. (2006). Phosphorylation of nucleotide excision repair factor xeroderma pigmentosum group A by ataxia telangiectasia mutated and Rad3-related-dependent checkpoint pathway promotes cell survival in response to UV irradiation. *Cancer Res* **66**, 2997–3005.

Yang, Z., Roginskaya, M., Colis, L. C., Basu, A. K., Shell, S. M., Liu, Y., Musich, P. R., Harris, C. M., Harris, T. M. and Zou, Y. (2006). Specific and efficient binding of xeroderma pigmentosum complementation group A to double-strand/single-strand DNA junctions with 3“- and/or 5-”ssDNA branches. *Biochemistry* **45**, 15921–15930.

CHAPTER 8

ATR PLAYS A DIRECT ANTIAPOPTOTIC ROLE AT MITOCHONDRIA WHICH IS REGULATED BY PROLYL ISOMERASE PIN1

Abstract

ATR, a PI3K-like protein kinase, plays a key role in regulating DNA damage responses. Its nuclear checkpoint kinase function is well documented but little is known about its function outside the nucleus. Here we report that ATR has an antiapoptotic activity at mitochondria in response to UV damage, and this activity is independent of its hallmark checkpoint/kinase activity and partner ATRIP. ATR contains a BH3-like domain that allows ATR-tBid interaction at mitochondria, suppressing cytochrome c release and apoptosis. This mitochondrial activity of ATR is downregulated by Pin1 that isomerizes ATR from *cis*-isomer to *trans*-isomer at the phosphorylated-Ser428-Pro429 motif. However, UV inactivates Pin1 via DAPK1, stabilizing the pro-survival *cis*-isomeric ATR. In contrast, nuclear ATR remains in the *trans*-isoform disregarding UV. This cytoplasmic response of ATR may provide a mechanism for the observed antiapoptotic role of ATR in suppressing carcinogenesis and its inhibition in sensitizing anticancer agents for killing of cancer cells.

Introduction

Phosphatidylinositol 3-kinase (PI3K)-like protein kinase ATR (*Ataxia telangiectasia* and Rad3-related) plays a critical role in maintaining genome integrity against DNA damage (Cimprich and Cortez, 2008; Sancar et al., 2004; Wu et al., 2006; Zeman and Cimprich, 2014; Zou et al., 2003), and is a central player in preventing the onset of cancer (Bartkova et al., 2005). As a checkpoint kinase, ATR phosphorylates hundreds of downstream proteins during DNA damage responses (DDR) (Matsuoka et al., 2007). ATR in complex with ATR-interacting protein (ATRIP) senses replicative stress-inducing DNA damage, activates checkpoints, arrests the cell cycle and facilitates repair to restore DNA integrity (Cortez et al., 2001). Recently ATR also was shown to mediate a mechanical stress checkpoint response (Kumar et al., 2014). ATR protects skin from UV-induced mutagenesis by promoting repair of UV damage (Jarrett et al., 2014). On the other hand, ATR is essential for mouse viability during development and ATR-knockout embryonic mice die of apoptosis (Brown and Baltimore, 2000; de Klein et al., 2000; O'Driscoll, 2009). Inhibition of ATR enhances apoptosis through a p53-independent mechanism, subsequently suppressing UV-induced carcinogenesis (Heffernan et al., 2009; Kawasumi et al., 2011). Hypomorphic suppression of ATR inhibits p53-deficient cancer growth in mice (Schoppy et al., 2012). Given these properties, ATR appears to be a promising target for anticancer chemotherapy. However, little is known about the functions of ATR in the cytoplasm.

Pin1 (peptidylprolyl *cis*/*trans* isomerase NIMA-interacting 1) is a critical regulator of many biological processes (Hunter, 1998; Liou et al., 2011; Z. Lu and Hunter, 2014). Dysfunction of Pin1 has been related to human diseases such as cancer, neurodegeneration, aging, heart disease, *etc*. (Ayala et al., 2003; Driver and K. P. Lu, 2010; P. J. Lu et al., 1999;

Nakamura et al., 2012; Sorrentino et al., 2014; Toko et al., 2013; Z.-G. Yang et al., 2002). Pin1 recognizes phosphorylated Ser/Thr-Pro bonds as catalytic sites for isomerization that regulates protein conformation, *cis* or *trans*, to control protein functions. Previous studies have suggested that Pin1-mediated conformational changes regulate a wide range of protein functions such as protein interactions, subcellular localization, DNA repair and protein stability (Steger et al., 2013; Zacchi et al., 2002).

In this study, we present evidence that ATR contains a BH3-like domain that allows ATR to function at mitochondria by directly interacting with Bid (BH3 interacting-domain death agonist) following UV damage. ATR binding to Bid may block tBid-dependent recruitment of Bax (proapoptotic Bcl2–associated X) to mitochondria, and, thus, suppresses release of cytochrome c and apoptosis. The mitochondria-specific ATR (mtATR) is negatively regulated by Pin1 via isomerization of the phospho-Ser428-Pro motif of ATR. This mtATR is structurally different from the nuclear ATR as the former is a *cis*-isomeric form while the latter is a *trans*-isomer. Surprisingly, mtATR was not found in association with ATRIP and contains no checkpoint kinase activity. Our results reveal a role of ATR at mitochondria as a pro-survival protein, expanding our understanding of the cellular functions of ATR and Pin1.

Materials and Methods

Cell Culture, UV Irradiation, Treatments and Antibodies

All cells except HCT-116 ATR+/+ and HCT-116 ATRflox/- cells (a gracious gift from Dr. David Cortez, Vanderbilt University) were purchased either from American Type

Culture Collection (Manassas, Virginia) or Coriell Cell Repositories (Camden, New Jersey). Cells were maintained in DMEM supplemented with heat-inactivated 10% FBS (Thermo Scientific, SH3039603) and 1% penicillin-streptomycin, and grown at 37°C, 5% CO2. UV irradiation was performed using a 254-nm lamp at a flounce of 0.83 J/m2/sec. All UV treatments unless otherwise noted are 40 J/m2 followed by a two-hour recovery. Camptothecin (C9911) and hydroxyurea (H8627) were purchased from Sigma Chemical Co. Stock solutions of camptothecin (CPT) and hydroxyurea (HU) were made in DMSO and H2O, respectively, before dilution into cell culture medium at final concentrations of 10 μM for 4 hrs and 2 mM for 20 hrs, respectively. The ATR kinase inhibitor NU6027 was purchased from EMD Millipore (189299) and was used at a final concentration of 10 μM in cell culture medium. For western blotting and immunofluorescence microscopy primary goat polyclonal antibodies against ATRIP (sc-33410) and mitochondrial heat- shock protein 70 (mHSP70) (sc-1060), mouse monoclonal antibodies against p53 (sc- 6243), ubiquitin (sc-166553) or PARP (sc-8007) were purchased from Santa Cruz Biotechnology Co. Mouse monoclonal antibody against mHSP70 (MA3-028) was purchased from Thermo Scientific. Antibodies against ATR were purchased from Bethyl Laboratories, Inc. (A300-137A and A300-138A) or from Cell Signaling Technology (2790). ATR pS428 antibody was purchased from Cell Signaling Technology (2853S). GAPDH antibody was from GeneTex (GTX100118). Antibodies specific for Pin1 and pPin1 (S71) were from GeneScript (A01090), Santa Cruz Biotechnology (sc-15340), or the K. P. Lu lab (Harvard University). tBid antibody was from Santa Cruz Biotechnology (sc-6538). Mouse monoclonal antibody against lamin A/C (4777), rabbit polyclonal antibodies against phospho-Chk1 (S345) (2348) and p53 (S15) (9284) were purchased from Cell Signaling Technology. Antibody

kits that recognize Bcl-2 pro-apoptotic family proteins were purchased from Cell Signaling Technology (9942). Anti-VDAC2 antibody was purchased from GeneTex (GTX104745). Anti-ATP synthase antibody was purchased from Abcam (ab110273). A FITC-conjugated primary mouse anti-actin antibody (F3022) was obtained from Sigma Chemical Co. The anti-actin, anti-mHSP70, anti-PARP, anti- GAPDH and anti-lamin A/C antibodies were used in western blots to confirm successful subcellular fractionations and to normalize protein loadings.

Cell lysis and Immunoblotting

Cells were harvested by scraping or trypsinization, and suspended in lysis buffer (50 mM Tris-HCl, pH 7.8, 150 mM NaCl, 1 mM EDTA, 1% Triton X-100, 1x protease/phosphatase inhibitor cocktail (Thermo, 1861280)). 2x SDS loading buffer was added to the lysates and the mixtures were heated at 95°C for 5 min. Gradient gels of 3-8% Tris-Acetate (TA) SDS-PAGE (Invitrogen, EA0378) were used for maximum resolution of the slower migrating cytoplasmic ATR (ATR-H) band. For display of total ATR proteins standard SDS-PAGE was used. Proteins were transferred from the gels onto PVDF membranes (Millipore, IPVH00010). Chemiluminescent signal was captured using the Fuji Film imager LAS-4000.

RNAi and Plasmid Transfections

Pools of siRNAs against p53 and ATR, each containing at least three siRNA duplexes, were purchased from Santa Cruz Biotechnology (sc-29435 and sc-29763, respectively).

Other siRNA duplexes targeting ATR and other genes were synthesized by Genepharm using the following sequences: ATR siRNA, sense strand 5'- CCUCCGUGAUGUUGCUUGATT-3′ and antisense strand 5′- UCAAGCAACAUCACGGAGGTT-3′; ATRIP siRNA, sense strand 5′-AGAGAAACUGUUCCAAUUATT-3′ and antisense strand 5'- UAAUUGGAACAGUUUCUCUTT-3'; Chk1 siRNA, sense strand 5'- ACAGUAUUUCGGUAUAAUATT-3' and antisense strand 5’-UAUUAUACCGAAAUACUGUTT-3'; Bax siRNA, sense strand 5′-GACGAACUGGACAGUAACATT-3′ and antisense strand 5’- UGUUACUGUCCAGUUCGUCTT-3’; DAPK1 sense strand 5’- GAGAAUCGAUGUCCAGGAUTT-3’ and antisense strand 5’-CUCUUAGCUACAGGUCCUATT-3’. The siRNA transfection reagent, Interferin, was purchased from Polyplus and the transfections were performed according to the manufactures instructions. Typically, cells were harvested at 48- or 72-hours post-transfection. A pcDNA3-FLAG-ATR plasmid for transient expression of human wild- type ATR was a gift from Dr. Stephen J. Elledge at Harvard Medical School. The transfection of plasmids into all cell lines was carried out with jetPEI transfection reagent (Polyplus, 101-10) by following the company's instructions. Site-directed mutagenesis and deletion generation were performed using the In-Fusion HD cloning kit (Clontech, 638909).

Cytoplasmic and Nuclear Protein Extraction

Subcellular fractionation was performed using cytoplasm lysis buffer (10 mM HEPES, pH 7.9, 10 mM KCl, 3 mM $CaCl_2$, 1.5 mM $MgCl_2$, 0.34 M Sucrose, 10% Glycerol, 0.1% Triton X-100, protease and phosphotase Inhibitors), and nuclei lysis buffer (50 mM Tris-HCl, pH 7.9, 140 mM NaCl, 3 mM $CaCl_2$, 1 mM EDTA, 1% NP-40, 10% Glycerol, protease and phosphotase Inhibitors). Briefly, 10 volumes of cell lysis buffer with 1x protease and

phosphatase inhibitors cocktail were added to 1 volume of packed cells. After resuspension and incubation on ice for 10 min, cytoplasm was separated from nuclei at 500xg for 7 min at 4°C. Isolated nuclei were washed twice with 500 µL of the nuclear wash buffer, collected by centrifugation. Collected nuclear pellets were suspended in ice-cold nuclear storage buffer, and 1/10 volume of the nuclear lysis reagent was added to lyse nuclei with rotation for 15 min at 4°C. The nuclear lysate was collected by centrifugation at 20,000xg for 15 min at 4°C. In all of the fractionation experiments, WB of β-actin and PARP were assessed to check the quality of fractionation and protein loading.

Alkali Extraction

Isolated mitochondria were treated with 0.1 M Na_2CO_3 (pH 11.5) for 30 min on ice before centrifugation at 16,000xg for 15 min at 4°C. The supernatant containing peripheral membrane proteins was collected. The pellet containing mitochondrial inner membrane proteins was washed once and suspended in 250 mM sucrose, 1 mM EDTA, 20 mM HEPES-NaOH (pH 7.4). Input refers to untreated control mitochondria. Equal proportions were analyzed by WB.

Purification of Recombinant ATR and Pin1 Proteins

Flag-tagged ATR recombinant protein was isolated from HEK 293T cells 48 hrs after transfection with pcDNA3-Flag-ATR plasmid. Clarified lysates were IPed with monoclonal mouse-anti-FLAG M2 antibody (Sigma) and captured with Protein-G plus agarose beads (Pierce). After rinsing with buffer, including a 1M NaCl wash to remove ATR-associated

proteins (*i.e.* ATRIP), the ATR was released into elution buffer (50 mM Tris-HCl, pH7.5, 150 mM NaCl, 10% glycerol, 1 mM DTT) containing 200 μg/ml Flag peptide. After dialysis with elution buffer containing 50% glycerol aliquots of ATR protein were stored at -80oC. RK plasmid containing a GST-tagged Pin1 construct, a generous gift from Dr. Lu, was transformed into *E.coli* (BL21(DE3)) and expression induced with IPTG at 30 degree for 5 hours. After bacterial lysis by sonication and clarification, expressed GST-Pin1 protein was bound to glutathione-agarose beads (Sigma), washed, and released with elution buffer (25 mM Tris-HCl, pH 7.5, 1 mM EDTA, 1 mM DTT) containing 10 mM reduced glutathione. The isolated GST-Pin1 was dialyzed into elution buffer containing 50% glycerol and aliquots stored at -80°C.

Mitochondria Isolation and Cytochrome c Release Assay

Mitochondria were isolated using kits from Thermo or from Qiagen; the later kit benefits from an additional density gradient purification step. Briefly, cells with or without UV treatment were suspended in lysis buffer to disrupt the plasma membrane. Compartmentalized organelles, such as endoplasmic reticulum, mitochondria and nuclei, remained intact and were collected by centrifugation at 1000xg for 10 min. The resulting pellet was resuspended in disruption buffer, repeatedly passed through a 21 gauge needle, and re-centrifuged at 1000xg for 10 min to pellet nuclei, cell debris, and unbroken cells. The supernatant containing mitochondria was re-centrifuged at 6000xg for 10 min to pellet mitochondria. Mitochondria also were isolated according to Frezza *et al.* (Frezza et al., 2007). Briefly, cells were suspended in buffer IBC, and then homogenized using a glass-Teflon Potter-Elvehjem homogenizer at 1600 rpm for 30 strokes followed by several

centrifugation steps, ultimately pelleting mitochondria. The cytochrome c assay was performed by subjecting isolated mitochondria to incubation with recombinant activated Bid (Bid cleaved by caspase-8, tBid) (R&D Systems, 883-M8) (Luo et al., 1998) to assess cytochrome c release. Briefly, isolated mitochondria were suspended in buffer B, tBid was added, then the reaction mixture was incubated at 30°C for 30 min. Following centrifugation, the supernatant was collected and loaded on a 15% SDS-PAGE gel, whilst the mitochondria pellet was lysed with 1x SDS loading buffer and analyzed on a 10% SDS-PAGE gel, followed by western blotting.

In vitro Isomerization and Immunoprecipitation Assays

The cytoplasm of A549 cells was isolated after UV treatment. ATR antibody (Bethyl Labs) was added for pull-down overnight. IPed ATR protein was subjected to a 0.6 M NaCl wash to remove associated proteins and suspended in kinase buffer (New England Biolabs). Cdk1-Cyclin B (New England Biolabs, P6020S) was added and incubated for 1 hr at 30°C to phosphorylate the isolated IPed ATR. After centrifugation and washing of the bead-bound ATR purified Pin1 was added to the phosphorylated ATR and incubated for 1 hr at 30 °C. SDS-loading buffer was added and ATR isomerization assayed by 3-8% TA SDS-PAGE and WB. Alternatively, to measure Pin1 binding to phosphorylated ATR the IPed ATR was incubated first with Cdk1 for 1 hr at 30°C, washed thrice, and purified Pin1 added for a 2 hr incubation at 4°C. The ATR-beads were washed three times and suspended in SDS loading buffer for WB analysis. To assess tBid binding to ATR-H following UV treatment of A549 cells, the cytoplasmic fraction or isolated mitochondria was collected. ATR antibody (Bethyl Labs) was added for pull-down overnight. IPed ATR-beads were washed three

washes in Co-IP wash buffer (50 mM Tris-HCl, pH 7.6, 140 mM NaCl, 1 mM EDTA, 10% glycerol, 0.2% Tween-20). The beads were suspended in 1x SDS loading buffer, boiled at 95°C for 5 min and analyzed by WB.

Duolink – *in situ* Proximity Ligation Assays

For Duolink microscopic detection of protein interactions cells were grown on coverslips before the initiation of experimental treatments. After UV-irradiation and recovery, cells were fixed with 4% paraformaldehyde and permeablized with 0.2% Triton X-100 in PBS and blocked with 3% bovine serum albumin in PBS for 1 hr at room temperature. Proteins then were detected with primary antibodies and the Duolink assay was carried out according to the company's protocol (Sigma, DUO92101). Images were taken using the Life Technologies Evos microscope or an Olympus confocal microscope.

Immunogold Labeling and TEM

A549 cells were treated with 40 J/m^2 UV with a 2 hr recovery. Cells scraped off the surface were collected by centrifugation and rinsed in PBS and fixed in 2% paraformaldehyde/2.5% glutaraldehyde in 0.2 M cacodylate. Fixed samples were dehydrated through alcohol and agar-enrobed before embedding in Lowicryl resin (Lowicryl K4M Embedding Kit, Electron Microscopy Sciences EMS). The resin was polymerized by multiday illumination with ultraviolet light before cutting ultrathin (60 nm) sections which were placed on gold glider grids (EMS, #G200-Au). Sections were blocked in 5% normal donkey serum to prevent nonspecific labeling, rinsed with 0.1 M

Sorenson's buffer before incubation with primary ATR antibody (Bethyl, A300-138A) for two hours at room temperature. Incubation with anti-rabbit secondary antibody conjugated to 20 nm gold particles (Cytodiagnostics) was carried out for two hours at room temperature. Sections were rinsed and imaged by a Philips Tecra 10 transmission electron microscope (FEI Company).

Mitochondrial Membrane Potential Assays

TMRE staining (ab113852) of active mitochondria (incorporation is based on natural mitochondrial polarity) was carried out by addition to culture media 20 min before harvesting. Both the adherent and non-adherent cells were harvested, washed with PBS, and analyzed in a BD Accuri C6 flow cytometer.

In vitro Kinase Activity Assay

The *in vitro* kinase assay was performed on endogenous ATR that was IPed from cytoplasmic or nuclear extracts of A549 cells after UV irradiation. The IPed ATR was washed three times with PBS containing 0.05% NP40, followed by a kinase buffer wash [50 mM HEPES (pH 7.5), 150 mM NaCl, 10 mM $MgCl_2$, 10 mM $MnCl_2$, 2 mM DTT, 10% glycerol, 1x protease and phosphatase inhibitors, and 0.5 mM ATP]. The IPed ATR was suspended in 20 µL of kinase buffer containing 10 µCi of [γ-32P] ATP and 0.5 µg of GST-p53 protein (Signal Chem, P05-30BG). The reaction was incubated at 30°C for 30 minutes and stopped by the addition of SDS loading buffer. Proteins were separated by SDS-PAGE and the radiolabeled proteins were visualized by gel scanning using a PhosphorImage scanner

(Fuji, Stamford, CT). IPed endogenous ATR and the amount of GST-p53 in each sample were confirmed by WB.

Results

Isomerization of ATR by Prolyl Isomerase Pin1 Is Inhibited by UV

To explore possible cytoplasmic activities of ATR, cells were treated with DNA damaging agents, followed by subcellular fractionation and Western blot analysis (WB). As shown in Figure 1A-top panel, DNA damage (particularly UV damage) induced an electrophoretic mobility shift of cytoplasmic but not nuclear ATR. The induced-cytoplasmic ATR (cytoATR) migrated slower in 3-8% gradient SDS-PAGE than the nuclear or non-induced cytoATR, and is designated as ATR-H relative to the latter (ATR-L). UV-induced ATR-H formation also was observed in keratinocytes (Figure 1A-bottom panel). ATR-H formation was dose- and post-UV recovery time-dependent (Figure 1B). The UV-induced ATR-H was observed in other human cell lines including primary fibroblasts (Figure S1A). The identity of ATR-H was confirmed via complementation of ATR deficient ($ATR^{flox/-}$) cells and mass spectrometry analysis of isolated ATR-H from UV-treated cells (Figures S2A and S2B, respectively). Also, cytoplasmic localization of ATR increased from ~26% to ~47% following UV treatment.

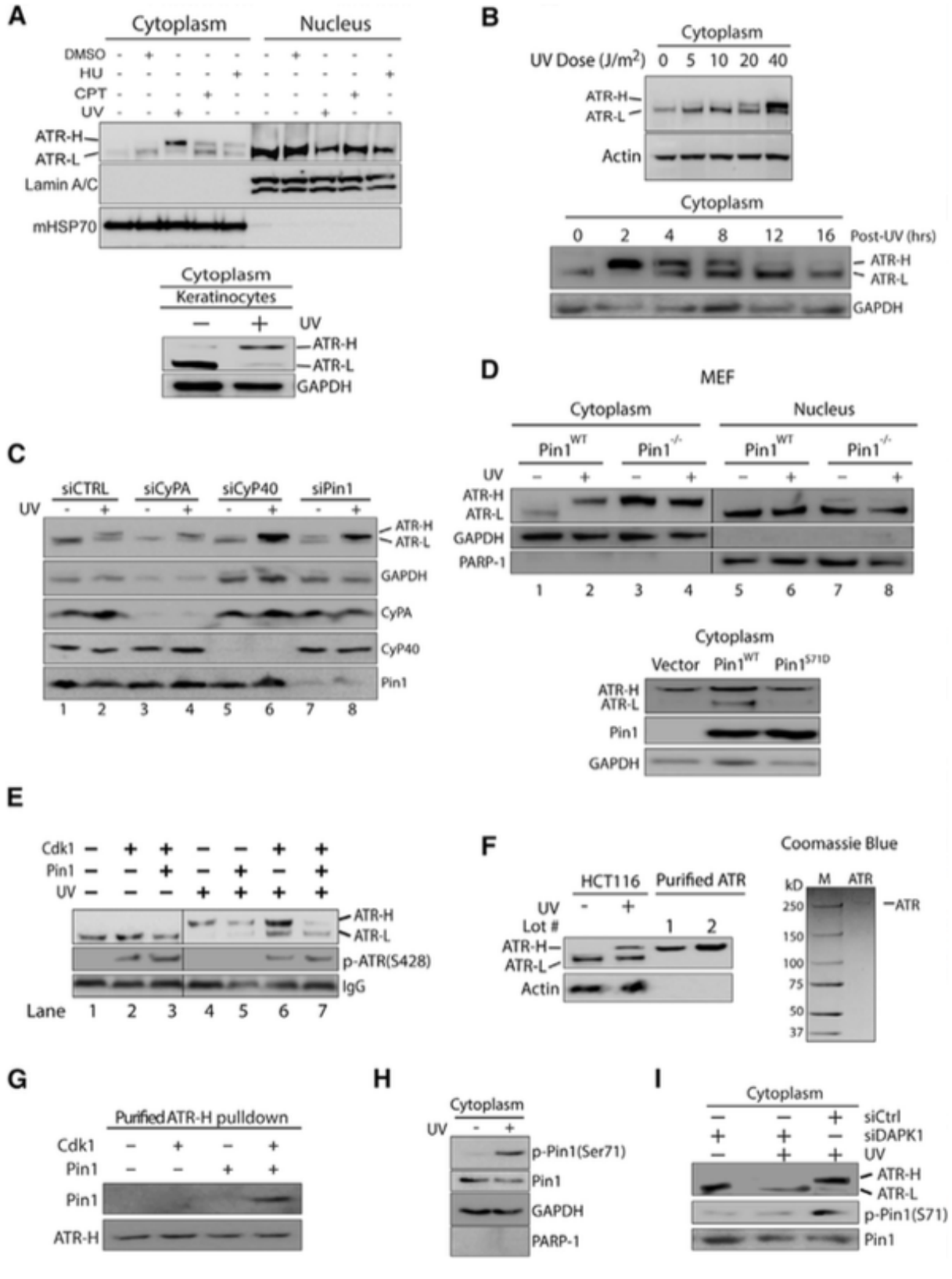
These observations implied that a post-translational modification to ATR protein may occur. Although ATR (301 kD) is known to undergo phosphorylation upon DNA damage (Liu et al., 2011; Nam et al., 2011), phosphorylation did not cause the ATR shift (Figure S2C). The same was true for modifications by glycosylation, ubiquitinylation,

sumoylation, palmitoylation, prenylation and N-myristoylation (Figure S2C or data not shown). Also, ATR-H is not a biosynthetic precursor to ATR-L as it occurred with translation inhibitors (data not shown). Alternatively, this mobility shift may reflect a backbone conformational change which persists in SDS-denaturing gel electrophoresis. Since isomerization of peptidyl-propyl bonds may confer a backbone conformational change, possible isomerization of ATR was examined. Prolyl isomerases cyclophilin-40 (CyP40), cyclophilin-A (CyPA) and Pin1 are known to be active in the cytoplasm. As shown in Figure 1C, Pin1 depletion resulted in formation of ATR-H even in the absence of UV (comparing ratios of ATR-H to ATR-L in lanes 1 vs. 7), while depletion of Cyp40 or CyPA had little effect (lanes 1 vs. 3 or 5). ATR-H formation in untreated cells transfected with Pin1 siRNA was also tested and confirmed in a keratinocyte cell line (Figure S1D). As expected, the siRNA effects on ATR-H production in the UV-induced cells were minimal as ATR-H already predominated (lane 2). These results suggest that Pin1was necessary for ATR-L formation and that lack of Pin1 led to ATR-H accumulation. Likewise, ATR-H was overwhelmingly found in the cytoplasm of $Pin1^{-/-}$ mouse embryonic fibroblasts (MEF) even without UV as compared with the predominant ATR-L in $Pin1^{+/+}$ MEF (Figure 1D, upper, lanes 3 vs. 1). Interestingly, for nuclear ATR lack of Pin1 resulted in only a very small amount of ATR-H with/without UV (lanes 7 and 8). Moreover, $Pin1^{-/-}$ MEF were transfected with Pin1-WT, and Pin1-S71D expression vectors (Figure 1D, lower). Since phosphorylation of Pin1 at Ser71 inhibits Pin1 isomerase activity (Lee et al., 2011), the phospho-mimic S71D mutation should make Pin1 catalytically inactive. As expected, little ATR-L was found in the cytoplasm of $Pin1^{-/-}$ cells expressing Pin1-S71D even without UV,

while there was a much higher level of cytoplasmic ATR-L in Pin1$^{-/-}$ cells expressing Pin1-WT. Together, these data support an essential role of Pin1 in maintaining ATR-L.

For further confirmation, endogenous cytoATR in control and UV-exposed cells was immunoprecipitated (IPed), washed with high-salt buffer, and subjected to an *in vitro* Pin1 isomerization assay. Since Pin1 recognizes the phosphorylated Ser/Thr-Pro motif of its substrates and Cdk1 was shown to phosphorylate Ser/Thr-Pro motif (Bernis et al., 2007; Holt et al., 2009), Cdk1 was used for *in vitro* phosphorylation of ATR though it is unknown if this occurs *in vivo*. In addition, to minimize isomerization of ATR-H back to ATR-L during purification of ATR from cells, no phosphatase inhibitors were used. As shown in Figure 1E, Cdk1 did phosphorylate ATR at Ser428 *in vitro* and the Cdk1-treated ATR-H was efficiently converted to ATR-L by Pin1 (lane 7). In addition, two lots of recombinant ATR were purified from HEK293 cells transfected with a pcDNA-ATR overexpression construct. As shown in Figure 1F, the purified ATR co-migrated with the ATR-H form, suggesting that ATR-L maintenance requires prolyl isomerase in cells and during protein purification.

Figure 1. Cytoplasmic ATR Is Prolyl-Isomerized by Pin1, which Is Inhibited upon UV Irradiation (A) Treatment with DNA-damaging agents reveals the formation of cytoplasmic ATR-H, which exhibited a slower electrophoretic mobility (top). Analysis of the HaCaT keratinocyte cells shows formation of ATR-H (lower). (B) ATR-H formation is UV dose and recovery time dependent. (C) siRNA knockdown of indicated peptidylprolyl isomerases shows that ATR-H formation increases in the cytoplasm of non-irradiated A549 cells only after Pin1 knockdown (lane 7). (D) Pin1 deficient (Pin1$^{-/-}$) MEF cytoplasm contains only ATR-H, while the proficient cells (Pin1wt) contain ATR-L under normal conditions and ATR-H following UV exposure (top). MEF Pin1$^{-/-}$ cells transfected with RK5 vector, Pin1WT or inactive Pin1S71D plasmids show significant accumulation of ATR-L only in the Pin1WT (bottom). (E) In vitro isomerization of purified ATR-H to ATR-L by purified recombinant Pin1. ATR-H phosphorylated by Cdk1 prior to Pin1 addition showed maximum conversion to ATR-L (lane 7). (F) Two lots (1 and 2) of recombinant ATR, purified from HEK293T cells (right), have the same electrophoretic mobility as endogenous ATR-H from UV-treated HCT116 cells. (G) Recombinant ATR-H phosphorylated by CDK1 directly interacts with recombinant Pin1. (H) Pin1 is phosphorylated at S71 in response to UV treatment. (I) Pin1 phosphorylation in response to UV is inhibited by knockdown of DAPK1 resulting in accumulation of ATR-L. See also Figures S1 and S2.


A
Cytoplasm
Nucleus
DMSO
HU
CPT
UV
ATR-H
ATR-L
Lamin A/C
mHSP70
Cytoplasm
Keratinocytes
UV
ATR-H
ATR-L
GAPDH
B
Cytoplasm
UV Dose (J/m2) 0 5 10 20 40
ATR-H
ATR-L
Actin
Cytoplasm
0 2 4 8 12 16 Post-UV (hrs)
ATR-H
ATR-L
GAPDH
C
siCTRL
siCyPA
siCyP40
siPin1
UV
ATR-H
ATR-L
GAPDH
CyPA
CyP40
Pin1
1 2 3 4 5 6 7 8
D
MEF
Cytoplasm
Nucleus
Pin1WT
Pin1-/-
UV
ATR-H
ATR-L
GAPDH
PARP-1
1 2 3 4 5 6 7 8
Cytoplasm
Vector
Pin1WT
Pin1S71D
ATR-H
ATR-L
Pin1
GAPDH
E
Cdk1
Pin1
UV
ATR-H
ATR-L
p-ATR(S428)
IgG
Lane 1 2 3 4 5 6 7
F
HCT116
Purified ATR
UV
Lot #
1 2
ATR-H
ATR-L
Actin
Coomassie Blue
kD
M
ATR
250
150
100
75
50
37
ATR
G
Purified ATR-H pulldown
Cdk1
Pin1
Pin1
ATR-H
H
Cytoplasm
UV
p-Pin1(Ser71)
Pin1
GAPDH
PARP-1
I
Cytoplasm
siCtrl
siDAPK1
UV
ATR-H
ATR-L
p-Pin1(S71)
Pin1

Also, purified Pin1 binds to the recombinant ATR-H after Cdk1 phosphorylation (Figure 1G). It should be noted that the enzymatic isomerization was carried out at 30°C while the IP at 4°C to facilitate maximum binding while preventing ATR isomerization. Interestingly, UV-irradiation led to phosphorylation of Pin1 at Ser71 in the cytoplasm of cells (Figure 1H), inactivating Pin1 and DAPK1 appears to be the kinase that phosphorylates Pin1 Ser71 upon UV (Figure 1I). Together these data suggest that: (a) UV may dephosphorylate ATR-H at its Pin1-binding site (Figure 1E, ATR-H/ATR-L ratios of lanes 7 vs. 5 and 1); (b) UV inactivation of Pin1 in the cytoplasm due to S71-phosphorylation by DAPK1 led to ATR-H formation (Figure 1H); and (c) active Pin1 and ATR-H phosphorylation at its Pin1-recognition site are required for ATR-L formation (Figure 1E).

ATR Contains a BH3-like Domain and Interacts with Bid at Mitochondria

To investigate the cellular functions of ATR-H, mitochondria were isolated from the cytoplasm of UV-treated cells. Comparison between the two subcellular compartments showed that ATR-H was exclusively associated with mitochondria (Figure 2A) and not free in the cytosol. The ATR-H association with mitochondria also was supported by immunofluorescence microscopic analysis of ATR colocalization with the mitochondria-specific protein MHSP70 (Figure 2B). Furthermore, transmission electron microscopy (TEM) coupled with immunogold labeling was employed to visualize the localization of ATR to the mitochondria in the cytoplasm (Figure 2C), showing that ATR accumulated at mitochondria after UV irradiation. More detailed localization of ATR-H on mitochondria was examined by an alkali extraction assay which separates outer and inner membrane

fractions of mitochondria (Bannwarth et al., 2012). ATR-H associated with the outer membrane rather than inner membrane (Figure 2D). This mitochondrial association appears due to a direct interaction between ATR and apoptotic protein Bid as demonstrated in the Duolink *in situ* proximity ligation assay (PLA) (Figure 2E). Bid plays an important role in facilitating proapoptotic activities (McDonnell et al., 1999; Zinkel et al., 2005). No interaction was found between ATR and the other tested apoptotic proteins. The ATR-H-Bid interaction was further confirmed by the co-immunoprecipitation (co-IP) of ATR-H with Bid (Figure 5D).

Protein sequence alignment of ATR with several Bcl2 family members reveals that ATR contains three potential BH3-like domains (Figure 3A), a hallmark of Bcl2 family proteins (Adams and Cory, 1998; Lessene et al., 2008). Domain-deletion expression constructs were made to examine these potential BH3-like domains in the ATR-Bid interaction. These constructs (ATR-WT, -Δ1 (aa175-187), -Δ2 (aa432-474) or -Δ3 (aa2345-2357)) had similar expression levels in the cytoplasm after transfection into $ATR^{flox/-}$ cells (Figure 3B). PLA revealed that ATR-Bid interaction (red foci) was normal in ATR-WT, -Δ1, or -Δ3 cells, but was abrogated in ATR-Δ2 cells (Figures 3C and 3D).

In support, dramatically much less ATR-H was found in association with mitochondria in ATR-Δ2 cells than in ATR-WT cells (Figure 3E). This indicates that the BH3-like domain (aa462-474) of ATR was required for ATR-Bid interaction. To assess Pin1 effects on the UV-induced ATR-Bid interaction, cells were depleted of Pin1, which significantly enhanced ATR-Bid foci formation even without UV (Figure 3F).

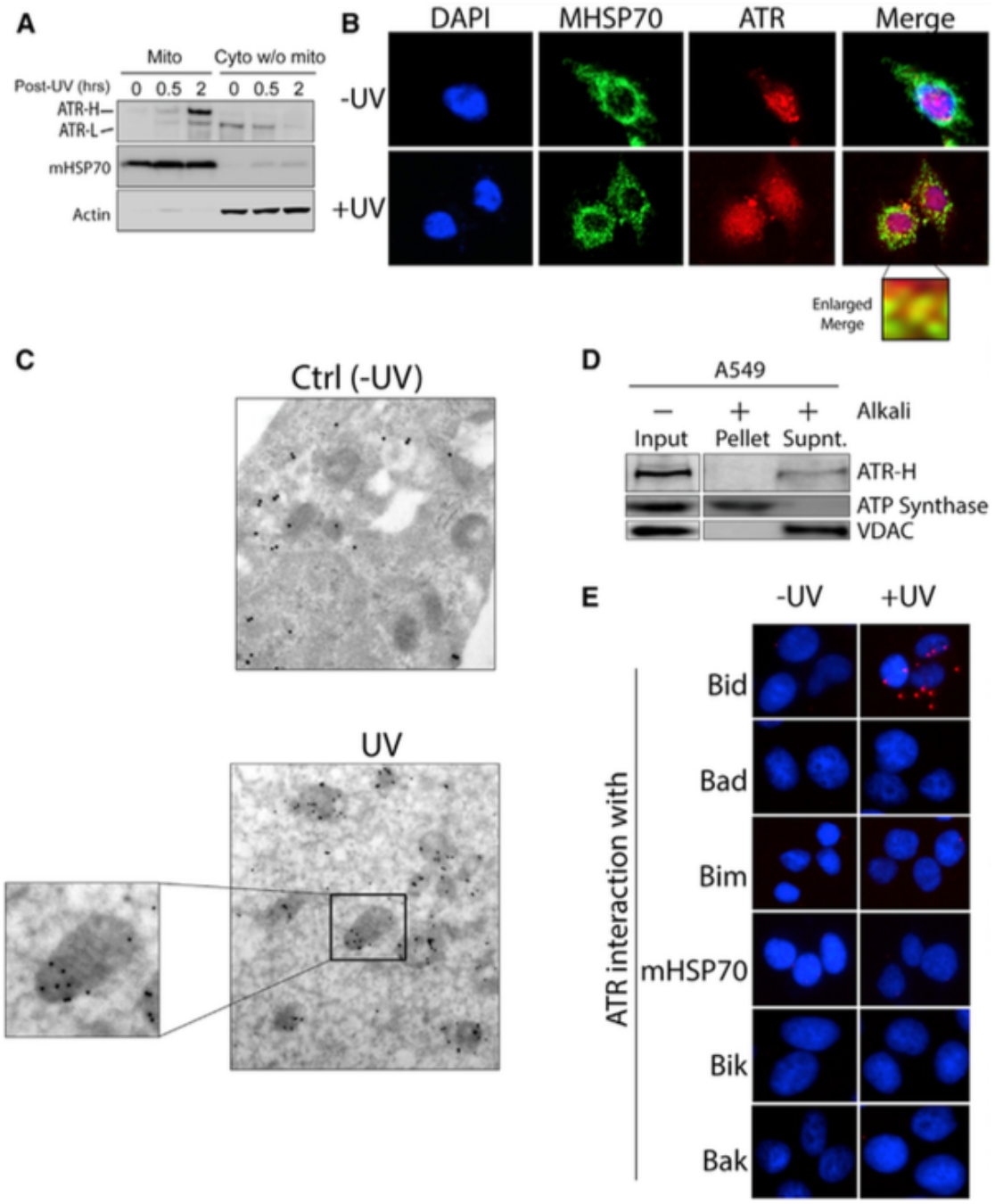


Figure 2. ATR Localizes to Mitochondria in Response to UV Irradiation (A) Cytosolic ATR accumulates at mitochondria (Mito) as ATR-H in a time-dependent manner. (B) ATR localizes to the mitochondria by immunofluorescence microscopy following UV irradiation. (C) Immuno-gold labeling followed by trans- mission electron microscopy (TEM) examination shows that ATR is dispersed in the cytoplasm (-UV) but localizes to the mitochondria following UV irradiation. Areas in rectangles were magnified to highlight mitochondrial localization of ATR. (D) Alkali extraction reveals that ATR-H interacts with the outer mitochondrial membrane. Intact mitochondria are represented by ‘‘Input.’’ ‘‘Pellet’’ represents the integral membrane proteins, and ‘‘Supnt’’ represents the soluble protein fraction. (E) Duolink PLA demonstrates that ATR interacts with proapoptotic protein Bid. Focus stacking reveals that the UV-induced ATR-Bid interaction predominately occurs outside of the nucleus.

Pin1 Isomerizes ATR at the Ser428-Pro Motif near the BH3-like Domain

To identify the Pin1-isomerization site in ATR we examined four Ser/Thr-Pro sites of ATR with reported/predicted S/T phosphorylation ($T^{22}P$, $S^{428}P$, $S^{1007}P$ and $T^{1989}P$). Alanine mutations were introduced at these Ser/Thr sites in ATR expression constructs which were transfected into $ATR^{flox/-}$ cells. ATR-S428A was predominately ATR-H even in the absence of UV, while all others were predominately ATR-L (Figure 4A upper). The residual amount of ATR-L in the S428A sample was likely due to the remaining endogenous wild-type ATR in $ATR^{flox/-}$ cells (Figure 3B, right panel). As expected, UV treatment generated a significant amount of ATR-H in all transfected cells (Figure 4A lower). Consistently, phosphorylation of cytoATR at Ser428-Pro was attenuated in cells treated with UV (Figure 4B upper). This also was true for the endogenous ATR (Figure 4B lower). These data suggest that besides the Ser71 phosphorylation of Pin1 to inhibit its isomerase activity, dephosphorylation of ATR-Ser428 may serve as a complementary mechanism to suppress Pin1 activity on ATR-H. Moreover, ATR-S428A expression abolished Pin1 activity on ATR, leading to a dramatic increase of ATR (ATR-H)-Bid interaction at mitochondria even in the absence of UV (Figure 4C).

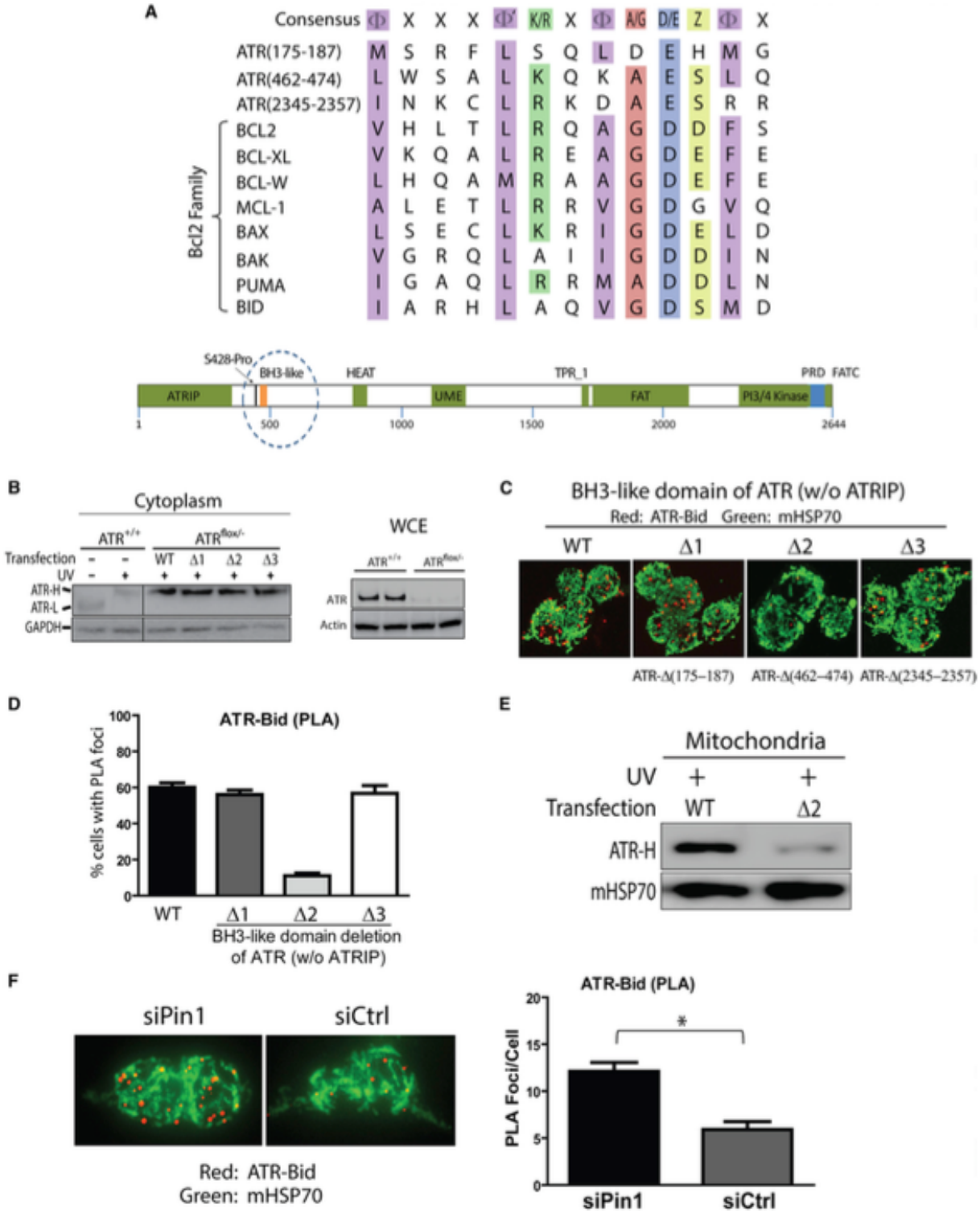

A
Consensus Φ X X X Φ K/R X Φ A/G D/E Z Φ X
ATR(175-187) M S R F L S Q L D E H M G
ATR(462-474) L W S A L K Q K A E S L Q
ATR(2345-2357) I N K C L R K D A E S R R
Bcl2 Family
BCL2 V H L T L R Q A G D D F S
BCL-XL V K Q A L R E A G D E F E
BCL-W L H Q A M R A A G D E F E
MCL-1 A L E T L R R V G D G V Q
BAX L S E C L K R I G D E L D
BAK V G R Q L A I I G D D I N
PUMA I G A Q L R R M A D D L N
BID I A R H L A Q V G D S M D
S428-Pro
BH3-like
HEAT
TPR_1
PRD FATC
ATRIP
UME
FAT
PI3/4 Kinase
1
500
1000
1500
2000
2644
B
Cytoplasm
WCE
ATR+/+
ATRflox/-
Transfection
UV
WT Δ1 Δ2 Δ3
ATR-H
ATR-L
GAPDH
ATR
Actin
C
BH3-like domain of ATR (w/o ATRIP)
Red: ATR-Bid Green: mHSP70
WT Δ1 Δ2 Δ3
ATR-Δ(175–187) ATR-Δ(462–474) ATR-Δ(2345–2357)
D
ATR-Bid (PLA)
% cells with PLA foci
100
80
60
40
20
0
WT Δ1 Δ2 Δ3
BH3-like domain deletion of ATR (w/o ATRIP)
E
Mitochondria
UV + +
Transfection WT Δ2
ATR-H
mHSP70
F
siPin1 siCtrl
Red: ATR-Bid
Green: mHSP70
ATR-Bid (PLA)
PLA Foci/Cell
20
15
10
5
0
*
siPin1 siCtrl

Figure 3. ATR Contains a BH3-like Domain that Is Required for Its Interaction with Mitochondria and Bid (A) Sequence alignment of BH3 domains in Bcl2 family proteins with ATR protein. In the consensus BH3 domain sequence, F represents a hydrophobic residue, Z a hydrophilic residue, D/E an acidic residue, A/G a small residue, K/R a basic residue, and X any residue. Three candidate BH3 domains were observed: ATR(175– 187), ATR(462–474), and ATR(2345–2357). Previously identified domains of ATR are shown in the lower bar diagram in relation to the ATR BH3-like domain and the Pin1 isomerization site (Ser428Pro). (B) Recombinant WT-ATR and ATR with deletion of one of the three BH3-like domains (D1, D2, D3, representing ATR lacking aa175–187, aa462–474, or aa2345– 2357, respectively) were expressed in the ATR-H form upon UV irradiation (left). $ATR^{flox/-}$ cells exhibit greatly reduced expression of endogenous ATR (right). (C) ATR-Bid interaction (PLA) requires BH3-like domain ATR(462–474). (D) Quantification of cells displaying ATR-Bid interaction in (C); three independent PLA experiments were performed. (E) The BH3-like domain (aa462–474) is necessary for ATR-H localization to mitochondria. (F) Pin1 isomerization of ATR reduces the ATR-Bid interaction (PLA), as shown by knockdown with indicated siRNA prior to UV treatment. The graph displays the average ATR-Bid interactions observed per cell from three independent experiments; data are represented as mean ± SD. See also Figure S6.

Together these data strongly support that Pin1 isomerizes ATR at p-Ser428-Pro429 which converts ATR from ATR-H to ATR-L in unstressed cells. An interesting question remaining is whether ATR-H is the *cis*- or the *trans*-isomer. Of the 20 canonical amino acids, peptide bonds formed between any pair of the 19 non-proline amino acids overwhelmingly adopt the *trans* conformation, while proline could be either the *cis*- or the *trans*-isomer with an equilibrium ratio depending on multiple protein features (Fischer and Schmid, 1990; Hinderaker and Raines, 2003). Thus, to identify the isoforms of ATR-H and ATR-L, a P429A point mutation was introduced in an ATR construct for expression in $ATR^{flox/-}$ cells. This mutation ensures a *trans*-isoform for S428-A429 bond. This mutation abolished UV-induced conversion of cytoplasmic ATR-L to ATR-H so that ATR-P429A remained in ATR-L form (Figure 4D last lane) and indicates that ATR-H is a *cis*-isomer while ATR-L is a *trans*-isomer. Also interestingly, although the N-terminally Flag-tagged ATR-WT expressed in cells could be IPed with anti-Flag antibody, the expressed Flag-

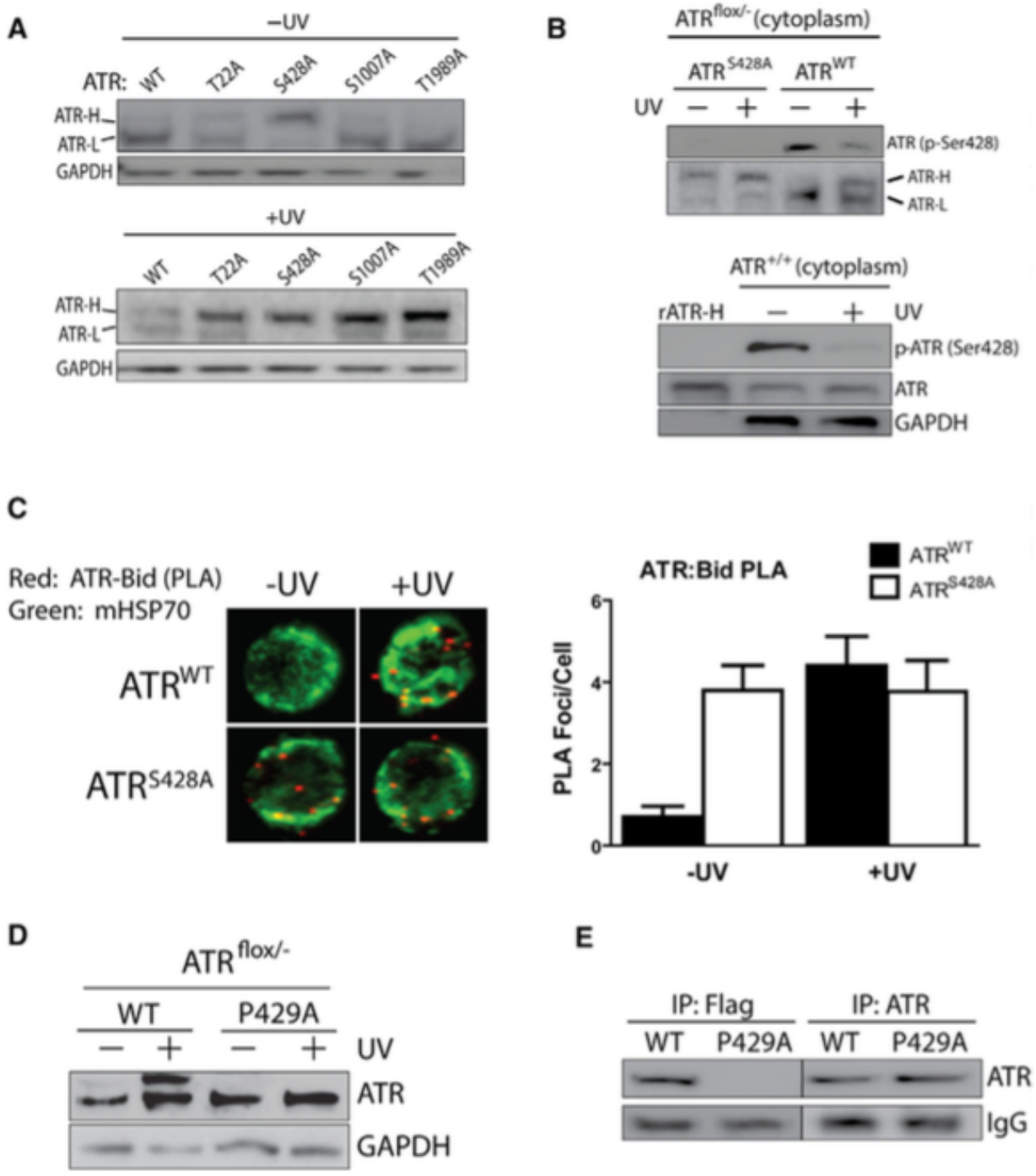


Figure 4. Pin1 Isomerizes ATR at the Phos-pho-Ser428-Pro Motif Near Its BH3-like Domain for ATR-Bid Interaction (A) Mutation of ATR at S428 resulted in ATR-H formation in non-UV-treated cells. (B) Phosphorylation of cytoATR at Ser428 in $ATR^{flox/-}$ cells (top). $ATR^{+/+}$ isolated from cytoplasm was analyzed to confirm phosphorylation status of Ser428 in UV-treated cells (bottom; note that this non-gradient SDS-PAGE does not separate ATR-H from ATR-L). rATR-H, purified recombinant ATR-H. (C) Phosphorylation of ATR at S428 is needed for ATR-Bid interaction. Data are represented as mean ± SD. (D) CytoATR mutated at P429 has the same electrophoretic mobility as the ATR-L following UV treatment. (E) $ATR^{flox/-}$ cells were transfected with constructs for expression of N-terminal Flag-tagged ATR-WT or mutated ATR-P429A followed by IP using anti- Flag antibodies or C-terminal-specific ATR anti- bodies.

tagged ATR-P429A mutant could not (Figure 4E). In contrast, both ATR-WT and its P429A mutant were efficiently IPed using ATR C-terminal antibodies (Figure 4E), suggesting that the Pin1-induced isomerization at P429 results in a conformational change in the N-terminal region of ATR where the BH3-like domain is located (Figure 3A). It is likely that the structural change makes the N-terminal Flag inaccessible to the antibody, and, thus, probably the BH3-like domain is sequestered as well.

ATR-Bid Binding Blocks Recruitment of Bax to Mitochondria and Prevents UV-Induced Apoptosis

Next, we investigated the effects of ATR-Bid interaction on UV-induced apoptosis. Mitochondria were isolated from UV-irradiated $ATR^{+/+}$ and $ATR^{flox/-}$ cells and post-UV accumulation of proapoptotic protein Bax on mitochondria was analyzed. In Figure 5A, UV-induced Bax accumulation at mitochondria significantly increased in a post-UV time-dependent manner in $ATR^{flox/-}$ cells while remaining unchanged in $ATR^{+/+}$ cells. It is known that Bax binding to tBid (truncated/apoptotically activated Bid) leads to activation of Bax at the mitochondria (Tait and Green, 2010). Once activated, Bax is capable of inducing mitochondrial outer membrane permeablization (MOMP) leading to release of cytochrome c and, thus, apoptosis (Billen et al., 2008; Westphal et al., 2014). Although tBid increasingly accumulated at mitochondria post UV in both $ATR^{+/+}$ and $ATR^{flox/-}$ cells, mitochondrial accumulation of Bax remained unchanged or even decreased in $ATR^{+/+}$ cells as compared to the significantly increased accumulation of Bax in $ATR^{flox/-}$ cells (Figure 5A). Bax is predominately cytosolic under unstressed conditions and upon activation is recruited from the cytoplasmic pool to mitochondria. Since the UV-induced mitochondrial Bax-tBid

association dramatically increased in $ATR^{flox/-}$ cells versus $ATR^{+/+}$ cells (Figure 5B), the negative effect of ATR on Bax mitochondrial accumulation is likely because ATR-H-tBid binding may block the Bax-tBid interaction.

Furthermore, an *in vitro* cytochrome c release assay (Luo et al., 1998) was performed. Mitochondria were isolated from UV-irradiated cells and supplied with purified tBid to induce cytochrome c release. As shown in Figure 5C, significantly more cytochrome c was released from $ATR^{flox/-}$ than $ATR^{+/+}$ cell mitochondria, consistent with the presence or absence of mitochondria-associated ATR-H. Also, more tBid was co-IPed with ATR-H than ATR-L from either a cytoplasmic fraction (including mitochondria) or a mitochondrial-only fraction (Figure 5D). Production of tBid by Bid cleavage under the experimental conditions was confirmed (Figure S3). To further verify ATR's direct involvement, isolated mitochondria from non-irradiated A549 cells were directly supplied with or without purified ATR-H and titrated with tBid in the cytochrome c release assay. WB analysis shows that the presence of ATR-H significantly reduced the cytochrome c release (Figure 5E). The same was true in an inverse titration study (Figure 5F). Cellular evidence for the direct involvement of ATR in the mitochondria-mediated apoptosis was demonstrated in Figure 5G where deletion of the identified ATR BH3-like domain (aa462-474) significantly reduced the survival relative to ATR-WT cells. In contrast, Pin1 depletion led to a substantial increase in survival of UV-treated cells (Figure 5H), fully consistent with the proapoptotic effect of Pin1 on the UV-induced cell death reported previously (Zacchi et al., 2002). It is worth noting that Pin1 can be either proapoptotic or antiapoptotic depending on its target proteins and the biological pathways it regulates (Sorrentino et al., 2014).

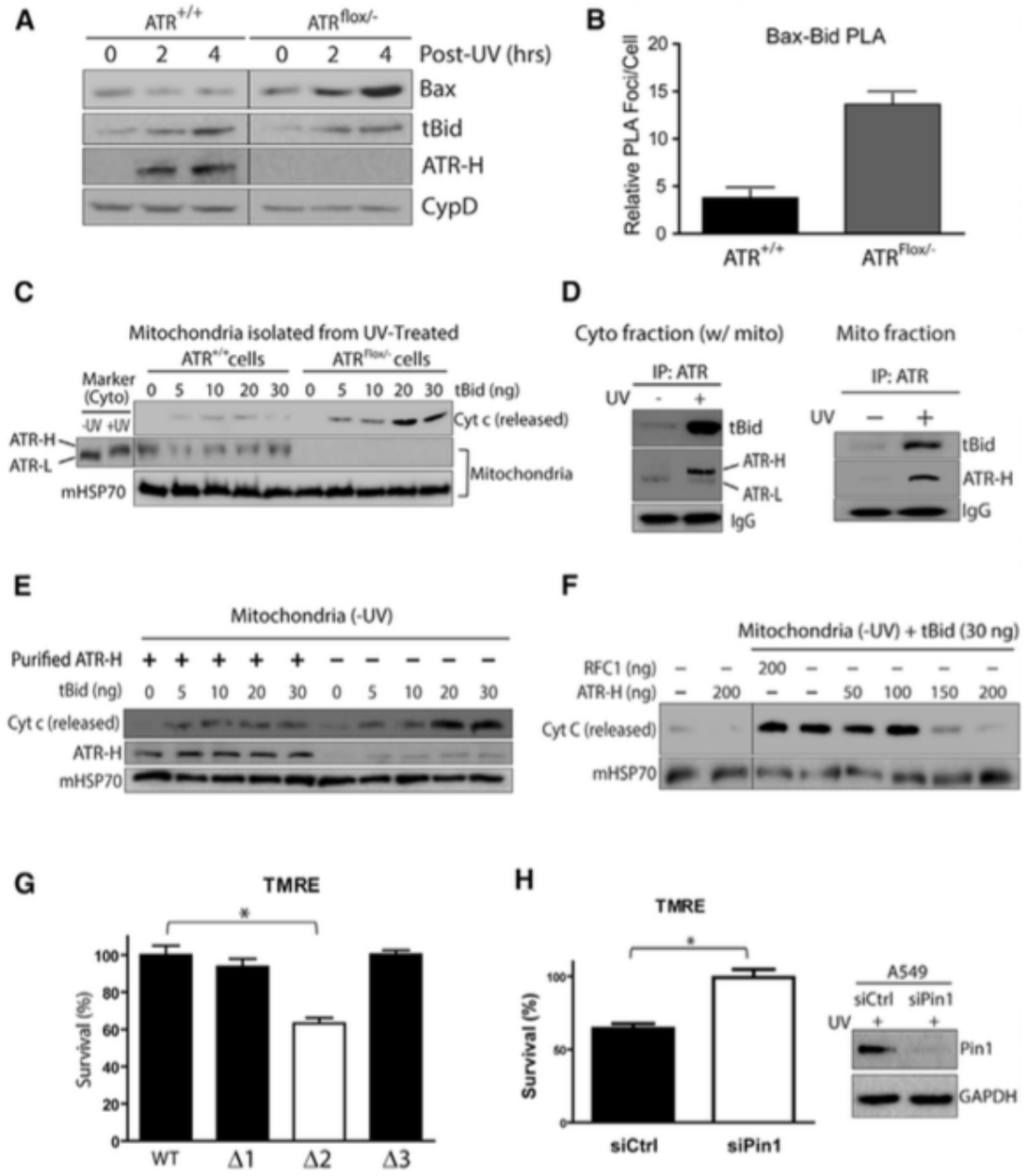


Figure 5. ATR Binding to Bid Inhibits Bax and Bak Recruitment to Mitochondria and Prevents UV-Induced Apoptosis (A) Lack of ATR in ATR$^{flox/-}$ allows recruitment of Bax and tBid to mitochondria following UV treatment. (B) ATR inhibits Bax-Bid complex formation. Confocal analysis of PLA foci was performed, and the average interactions between Bax-tBid observed per cell were quantified from three independent experiments. (C) The presence of ATR-H at mitochondria prevents cytochrome *c* release. Mitochondria isolated from UV-treated ATR$^{+/+}$ or ATR$^{flox/-}$ cells were supplemented with tBid to initiate cytochrome *c* release. Marker is the cytosolic fraction of ATR$^{+/+}$ cells ± UV. (D) tBid interacts with ATR-H upon UV exposure of cells. tBid co-IPed with ATR form the cytoplasmic and the mitochondrial fractions of UV-treated A549 cells. (E and F) Recombinant ATR-H can prevent cytochrome *c* release. Mitochondria isolated from untreated A549 cells were incubated with recombinant ATR-H and then with increasing amounts of tBid (E). The amount of recombinant ATR-H required to inhibit cytochrome *c* release was assessed (F). Purified RFC1 was used as a negative control for mitochondria binding. (G) The BH3-like domain D2 (aa462–474) of ATR is required for cell survival against UV. ATR$^{flox/-}$ cells were transfected with plasmid constructs as in Figure 3B before UV exposure and tetramethylrhodamine ethyl ester (TMRE) staining. (H) Silencing Pin1 in cells increases cell survival after UV exposure likely due to ATR-H accumulation. Data in (B), (G), and (H) are represented as mean ± SD. See also Figure S3.

ATR Mitochondrial Function is Independent of its Checkpoint Kinase Activity

Since ATR forms a complex with ATRIP in the nucleus, it is of interest to examine whether ATRIP is involved in the mitochondrial activity of ATR. Surprisingly, little to no ATRIP was found in the cytoplasm as compared to ATRIP in the nucleus (Figure 6A). Consistently, ATRIP, but not ATR, depletion had no effect on the UV-induced mitochondrial accumulation of Bax (Figure 6B). Given that ATRIP plays an essential role in ATR checkpoint activation, we examined the checkpoint kinase activity of cytoATR. ATR was IPed from cytoplasmic and nuclear fractions of ±UV-treated cells, and subjected to an *in vitro* kinase assay for ^{32}P-labeling of p53 by ATR. While p53 was efficiently phosphorylated by nuclear ATR, cytoATR failed to label p53 with ^{32}P (Figure 6C), likely due to the lack of ATRIP. Consistently, ATR kinase inhibition by NU6027 (ATR-KI), but not ATR depletion, failed to promote the UV-induced Bax accumulation at mitochondria in $ATR^{+/+}$ cells (Figure 6D). Further confirmation came from the UV-induced ATR-H formation in the cytoplasm of $ATR^{flox/-}$ cells expressing ATR-WT or ATR-KD (kinase-dead) constructs (Figure S4). It is worth mentioning that the proapoptotic effect of ATR depletion has been shown previously to be independent of p53 (Schoppy et al., 2012). Together our data suggest that the mitochondrial role of ATR may not depend on its checkpoint kinase activity.

Discussion

The present work reveals that ATR contains a BH3-like domain and plays a direct antiapoptotic role at mitochondria via interacting with tBid following DNA damage. As summarized in Figure 7, the mitochondrial activity of ATR is carried out by a peptidyl-prolyl isomeric *cis*-form of ATR (ATR-H) and is negatively regulated by the prolyl

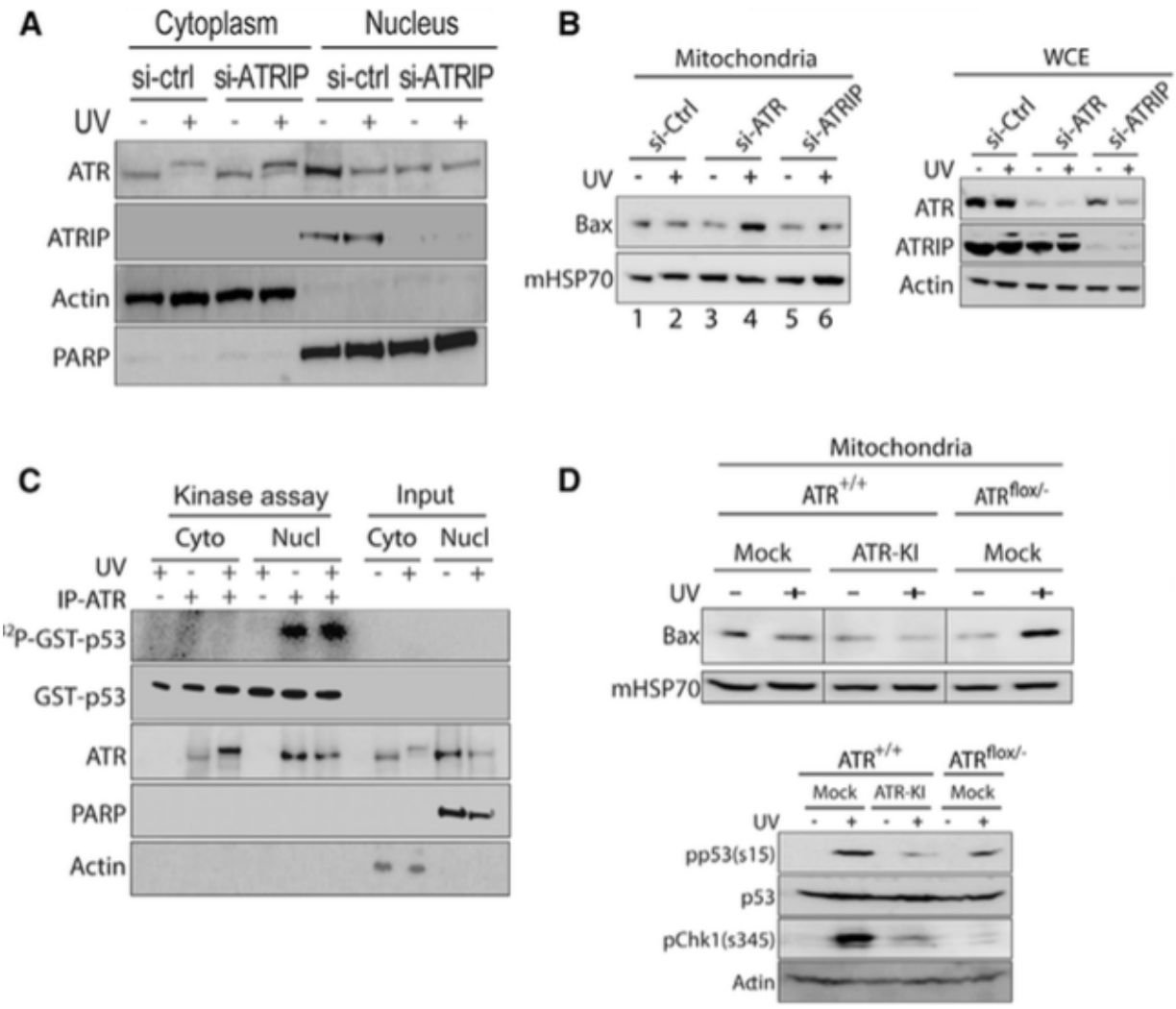


Figure 6. Mitochondrial-Specific ATR Functions Independently of Its Checkpoint Kinase Activity and ATRIP (A) A lack of ATRIP in the cytoplasmic fraction suggests that ATR kinase is not required for ATR-H formation. (B) ATRIP is not required for inhibition of Bax translocation by ATR-H. Bax association with isolated mitochondria (left) and the efficiency of the ATR and ATRIP knockdowns in whole cell extracts (WCE) are shown (right). (C) Confirmation of ATR-H's lack of kinase activity. In vitro phosphorylation of purified GST-p53 was carried out to assess the checkpoint kinase activity of IPed ATR (entire cytoplasmic ATR-H fraction versus one third of nuclear ATR-L fraction) in the presence of [γ^{32}P] ATP. (D) ATR kinase activity is not required to prevent Bax recruitment to mitochondria. Association of Bax with the indicated fractions under given conditions was analyzed. Phosphorylation of p53 on Ser15 (pp53(S15)) and Chk1 on Ser345 (pChk1(S345)) were monitored to confirm the loss of checkpoint kinase activity of ATR (right). See also Figure S4.

isomerase Pin1 which converts ATR-H to ATR-L (*trans*-ATR). Our data imply that this isomeric conversion results in a conformational change in the N-terminal region of ATR (Figure 4E), probably burying the BH3-like domain. In contrast, inhibition of the conversion may lead to generation of ATR-H with a structure exposing the BH3-like domain for mitochondria-specific functions. Since ATR is well known for its DDR role in the nucleus, including apoptosis suppression via the ATR-Chk1 pathway (Heffernan et al., 2009), this study reveals a Pin1-regulated function of ATR in the cytoplasm. Interestingly, although ATR and ATM (*ataxia telangiectasia* mutated) are closely related PI3K checkpoint kinases, unlike ATR, ATM deficiency promotes carcinogenesis. This study may provide a partial explanation for the differential effects.

ATR-H identification as the *cis*-isomeric ATR is supported by the previous biochemical data and evidence presented here: a) all amino acids except proline have a strong energetic preference for *trans* peptide bond conformation, while proline is naturally stabilized by a preceding *cis* peptide bond (Fischer et al., 1984; Fischer and Schmid, 1990; Hinderaker and Raines, 2003); b) Pin1 isomerizes ATR-H to ATR-L in both cell-based and *in vitro* assays (Figure 1); c) Pin1 catalyzes *cis* to *trans* isomerization of proteins such as Tau (P. J. Lu et al., 1999); d) importantly, a point mutation, P429A, in ATR which converts ATR residue 429 to an alanine of *trans*-isomeric form prevented formation of ATR-H even with UV treatment (Figure 4D), and e) additionally, ATR with either point mutation, S428A or P429A, was observed in the cytoplasm following UV treatment (Figure S6). Thus, as shown in Figure 7, Pin1 converts ATR-H to ATR-L (*trans*-ATR) in unstressed cells. Upon UV, however, cytoplasmic Pin1 is phosphorylated at S71 by DAPK1 which inactivates Pin1, preventing the cytoplasmic conversion of ATR-H to ATR-L. Pin1 previously has been shown

to be either pro- or anti-apoptotic depending on the targets, pathways and biological conditions it regulates (Sorrentino et al., 2014). In line with a previous report for Pin1's proapoptotic role against UV (Zacchi et al., 2002), here Pin1 plays a proapoptotic role. UV-inactivation of Pin1 enhances the mtATR-tBid association which suppresses the proapoptotic activity of Bax and thus initiation of apoptosis (Figure 7).

Localization of ATR to mitochondria may occur either through the binding of ATR-H to mitochondria-bound tBid or through cytoplasmic ATR-H-tBid interaction to facilitate mitochondrial localization of the ATR-H-tBid complex. Under either or both pretenses, the binding of ATR-H to tBid may lead to partial tBid sequestration, serving a role in preventing Bax/Bak activation. It has been shown that interaction of the Bid's BH3 domain with antiapoptotic Bcl-XL protein may involve BH3, BH1 and BH2 domains from Bcl-XL by a BH3-in-groove mechanism (Chou et al., 1999; Czabotar et al., 2013). Interestingly, sequence alignment analysis suggests that ATR does have several potential BH1- and BH2-like domains (Figure S5A). Even though multiple BH domains may participate in the interaction, loss of BH3 in ATR is likely to disrupt the complex formation (Figure 3). The structural details of how the ATR BH3-like domain interacts with Bid remain to be defined. However, recently a non-canonical interaction was reported for inhibition of proapoptotic proteins by antiapoptotic Bcl2 (Barclay et al., 2015), arguing that the possibility of a non-canonical mechanism should not be ruled out. Also interestingly, secondary structure prediction of ATR shows that ~72% of ATR sequence is likely alpha-helical (I-TASSER) (J. Yang et al., 2015), analogous to the average ~75% alpha-helix found in antiapoptotic Bcl2 family proteins. By contrast, an average globular protein contains only ~30% alpha-helix (Pace and Scholtz, 1998). Also, ATR is predicted to contain 45 HEAT repeats (domains

typically found in cytoplasmic proteins), each containing two alpha helices. The BH3-like domain described here overlaps the first helix within HEAT repeat #9 (Figure S5B). The Pin1 site is centered in a 46-residue region (Pondr-3 in Figure S5B) predicted to be multiply bent, non-alpha helical (http://www.pondr.com/index), making it accessible for modification. In ATR-H this organization may enhance the ATR-tBID interaction since HEAT repeats facilitate protein-protein interactions.

The proximate location of the Ser428-Pro motif to the BH3-like domain in ATR makes it possible that Pin1 could regulate the BH3-like domain accessibility *via* an isomerization-induced conformational change in ATR. It is possible that the conformational change may alter the accessibility of ATR's BH3-like domain for mitochondrial targeting, as implicated by the data in Figure 4E. Moreover, ATR-L formation depends on both dephosphorylation of Pin1-Ser71 and phosphorylation of ATR at Ser428. Our data indicate that ATR-H is under-phosphorylated at Ser428 after UV irradiation (Figures 4B and S1B), suggesting a second level of control besides Pin1-S71 phosphorylation. It is of interest to identify in future studies the kinases and phosphatases that modify ATR-Ser428 residue. Although Cdk1 can phosphorylate ATR at Ser428 *in vitro* (Figure 1E), there is no evidence that this is true in cells; but deserves further investigation.

Interestingly, cytoATR shows little checkpoint kinase activity, most likely due to the absence of ATRIP in the cytoplasm (Figure 6). In contrast, ATR in the nucleus is known to complex with ATRIP and render checkpoint activity. Nuclear ATR remains predominately in the ATR-L form (*trans*-ATR), even in the absence of Pin1 (Figure 1D). This may suggest: a) ATR-ATRIP complex formation and chromatin association, absent in the cytoplasm, might energetically favor ATR-L regardless of Pin1, and/or b) a back-up isomerization

system to Pin1 might exist in the nucleus to prevent ATR from forming ATR-H. More interesting questions would be: can ATR-H form a complex with ATRIP and, if so, does ATR-H contain checkpoint kinase activity after forming a complex with ATRIP?

DNA damage checkpoints and apoptosis are two major pathways of DDR. Upon moderate DNA damage, checkpoints are activated, leading to cell cycle arrest and DNA repair. It is believed that checkpoint activation is concurrent with suppression of apoptosis as eventually checkpoints will subside and normal cell cycling will resume after completion of DNA repair. The presented data demonstrate a mechanism by which the two pathways may work in a coordinate manner in DDR.

Our findings expand our understanding of the cellular functions of ATR and Pin1, and also provide a molecular basis for the observed pathological relevance of ATR with important translational implications. For example, since the cellular activities of ATR can be regulated by Pin1 through three post-translational modifications (Ser71 of Pin1, and Ser428 and Pro429 of ATR), manipulation of these modifications via pharmacological targeting may selectively inhibit one or both of the ATR activities for the purpose of developing new therapeutic treatments, especially for cancer (Lord and Ashworth, 2008; Weber and Ryan, 2014).

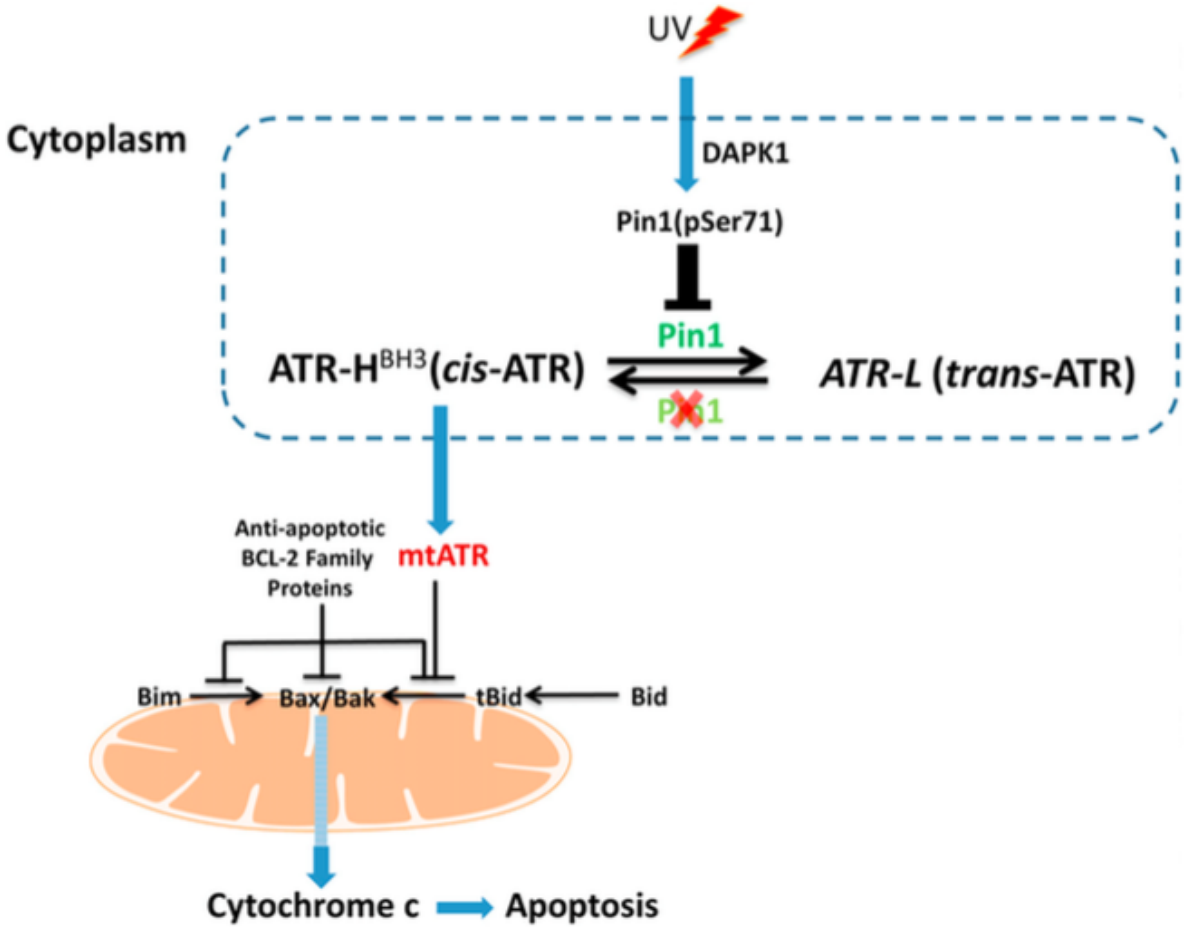


Figure 7. Proposed Mechanisms of Pin1 Regulation of ATR Isomeric Forms and the Function of ATR at Mitochondria in Response to UV Damage. ATR-H (*cis*-ATR) formation occurs in the cyto- plasm due to Pin1 inactivation by UV-induced phosphorylation at Ser71 by DAPK1 or by genetic deficiency; ATR-H can either localize to the mito- chondria, where it interacts with tBid inserted into the outer mitochondrial membrane via its now accessible BH3-like domain, and/or bind to cyto- solic tBid before mitochondria localization. The common result is that ATR-H functions as an an- tiapoptotic protein, preventing further recruitment of Bax to mitochondria and subsequent Bax/Bak activation and thus deterring cytochrome *c* release and apoptosis.

References

Adams, J.M., Cory, S., 1998. The Bcl-2 protein family: arbiters of cell survival. Science 281, 1322–1326.

Ayala, G., Wang, D., Wulf, G., Frolov, A., Li, R., Sowadski, J., Wheeler, T.M., Lu, K.P., Bao, L., 2003. The prolyl isomerase Pin1 is a novel prognostic marker in human prostate cancer. Cancer Res 63, 6244–6251.

Bannwarth, S., Figueroa, A., Fragaki, K., Destroismaisons, L., Lacas-Gervais, S., Lespinasse, F., Vandenbos, F., Pradelli, L.A., Ricci, J.-E., Rötig, A., Michiels, J.-F., Vande Velde, C., Paquis-Flucklinger, V., 2012. The human MSH5 (MutSHomolog 5) protein localizes to mitochondria and protects the mitochondrial genome from oxidative damage. Mitochondrion 12, 654–665. doi:10.1016/j.mito.2012.07.111

Barclay, L.A., Wales, T.E., Garner, T.P., Wachter, F., Lee, S., Guerra, R.M., Stewart, M.L., Braun, C.R., Bird, G.H., Gavathiotis, E., Engen, J.R., Walensky, L.D., 2015. Inhibition of Pro-Apoptotic BAX by a Noncanonical Interaction Mechanism. Mol Cell 57, 873–886. doi:10.1016/j.molcel.2015.01.014

Bartkova, J., Horejsí, Z., Koed, K., Krämer, A., Tort, F., Zieger, K., Guldberg, P., Sehested, M., Nesland, J.M., Lukas, C., Ørntoft, T., Lukas, J., Bartek, J., 2005. DNA damage response as a candidate anti-cancer barrier in early human tumorigenesis. Nature 434, 864–870. doi:10.1038/nature03482

Bernis, C., Vigneron, S., Burgess, A., Labbé, J.-C., Fesquet, D., Castro, A., Lorca, T., 2007. Pin1 stabilizes Emi1 during G2 phase by preventing its association with SCF(betatrcp). EMBO Rep 8, 91–98. doi:10.1038/sj.embor.7400853

Billen, L.P., Shamas-Din, A., Andrews, D.W., 2008. Bid: a Bax-like BH3 protein. Oncogene 27 Suppl 1, S93–104. doi:10.1038/onc.2009.47

Brown, E.J., Baltimore, D., 2000. ATR disruption leads to chromosomal fragmentation and early embryonic lethality. Genes Dev 14, 397–402.

Chou, J.J., Li, H., Salvesen, G.S., Yuan, J., Wagner, G., 1999. Solution structure of BID, an intracellular amplifier of apoptotic signaling. Cell 96, 615–624.

Cimprich, K.A., Cortez, D., 2008. ATR: an essential regulator of genome integrity. Nat Rev Mol Cell Biol 9, 616–627. doi:10.1038/nrm2450

Cortez, D., Guntuku, S., Qin, J., Elledge, S.J., 2001. ATR and ATRIP: partners in checkpoint signaling. Science 294, 1713–1716. doi:10.1126/science.1065521

Czabotar, P.E., Westphal, D., Dewson, G., Ma, S., Hockings, C., Fairlie, W.D., Lee, E.F., Yao, S., Robin, A.Y., Smith, B.J., Huang, D.C.S., Kluck, R.M., Adams, J.M., Colman, P.M., 2013. Bax

crystal structures reveal how BH3 domains activate Bax and nucleate its oligomerization to induce apoptosis. Cell 152, 519–531. doi:10.1016/j.cell.2012.12.031

de Klein, A., Muijtjens, M., van Os, R., Verhoeven, Y., Smit, B., Carr, A.M., Lehmann, A.R., Hoeijmakers, J.H., 2000. Targeted disruption of the cell-cycle checkpoint gene ATR leads to early embryonic lethality in mice. Curr Biol 10, 479–482.

Driver, J.A., Lu, K.P., 2010. Pin1: a new genetic link between Alzheimer's disease, cancer and aging. Curr Aging Sci 3, 158–165.

Fischer, G., Bang, H., Mech, C., 1984. [Determination of enzymatic catalysis for the cis-trans-isomerization of peptide binding in proline-containing peptides]. Biomed. Biochim. Acta 43, 1101–1111.

Fischer, G., Schmid, F.X., 1990. The mechanism of protein folding. Implications of in vitro refolding models for de novo protein folding and translocation in the cell. Biochemistry 29, 2205–2212.

Frezza, C., Cipolat, S., Scorrano, L., 2007. Organelle isolation: functional mitochondria from mouse liver, muscle and cultured fibroblasts. Nat Protoc 2, 287–295. doi:10.1038/nprot.2006.478

Heffernan, T.P., Kawasumi, M., Blasina, A., Anderes, K., Conney, A.H., Nghiem, P., 2009. ATR-Chk1 pathway inhibition promotes apoptosis after UV treatment in primary human keratinocytes: potential basis for the UV protective effects of caffeine. J. Invest. Dermatol. 129, 1805–1815. doi:10.1038/jid.2008.435

Hinderaker, M.P., Raines, R.T., 2003. An electronic effect on protein structure. Protein Sci 12, 1188–1194. doi:10.1110/ps.0241903

Holt, L.J., Tuch, B.B., Villén, J., Johnson, A.D., Gygi, S.P., Morgan, D.O., 2009. Global analysis of Cdk1 substrate phosphorylation sites provides insights into evolution. Science 325, 1682–1686. doi:10.1126/science.1172867

Hunter, T., 1998. Prolyl isomerases and nuclear function. Cell 92, 141–143.

Jarrett, S.G., Wolf Horrell, E.M., Christian, P.A., Vanover, J.C., Boulanger, M.C., Zou, Y., D'Orazio, J.A., 2014. PKA-mediated phosphorylation of ATR promotes recruitment of XPA to UV-induced DNA damage. Mol Cell 54, 999–1011. doi:10.1016/j.molcel.2014.05.030

Kawasumi, M., Lemos, B., Bradner, J.E., Thibodeau, R., Kim, Y.-S., Schmidt, M., Higgins, E., Koo, S.-W., Angle-Zahn, A., Chen, A., Levine, D., Nguyen, L., Heffernan, T.P., Longo, I., Mandinova, A., Lu, Y.-P., Conney, A.H., Nghiem, P., 2011. Protection from UV-induced skin carcinogenesis by genetic inhibition of the ataxia telangiectasia and Rad3-related (ATR) kinase. Proc Natl Acad Sci USA 108, 13716–13721.

doi:10.1073/pnas.1111378108

Kumar, A., Mazzanti, M., Mistrik, M., Kosar, M., Beznoussenko, G.V., Mironov, A.A., Garrè, M., Parazzoli, D., Shivashankar, G.V., Scita, G., Bartek, J., Foiani, M., 2014. ATR Mediates a Checkpoint at the Nuclear Envelope in Response to Mechanical Stress. Cell 158, 633–646. doi:10.1016/j.cell.2014.05.046

Lee, T.H., Chen, C.-H., Suizu, F., Huang, P., Schiene-Fischer, C., Daum, S., Zhang, Y.J., Goate, A., Chen, R.-H., Zhou, X.Z., Lu, K.P., 2011. Death-associated protein kinase 1 phosphorylates Pin1 and inhibits its prolyl isomerase activity and cellular function. Mol Cell 42, 147–159. doi:10.1016/j.molcel.2011.03.005

Lessene, G., Czabotar, P.E., Colman, P.M., 2008. BCL-2 family antagonists for cancer therapy. Nat Rev Drug Discov 7, 989–1000. doi:10.1038/nrd2658

Liou, Y.-C., Zhou, X.Z., Lu, K.P., 2011. Prolyl isomerase Pin1 as a molecular switch to determine the fate of phosphoproteins. Trends in Biochemical Sciences 36, 501–514. doi:10.1016/j.tibs.2011.07.001

Liu, S., Shiotani, B., Lahiri, M., Maréchal, A., Tse, A., Leung, C.C.Y., Glover, J.N.M., Yang, X.H., Zou, L., 2011. ATR autophosphorylation as a molecular switch for checkpoint activation. Mol Cell 43, 192–202. doi:10.1016/j.molcel.2011.06.019

Lord, C.J., Ashworth, A., 2008. Targeted therapy for cancer using PARP inhibitors. Current Opinion in Pharmacology 8, 363–369. doi:10.1016/j.coph.2008.06.016

Lu, P.J., Wulf, G., Zhou, X.Z., Davies, P., Lu, K.P., 1999. The prolyl isomerase Pin1 restores the function of Alzheimer-associated phosphorylated tau protein. Nature 399, 784–788. doi:10.1038/21650

Lu, Z., Hunter, T., 2014. Prolyl isomerase Pin1 in cancer. Cell Res 24, 1033–1049. doi:10.1038/cr.2014.109

Luo, X., Budihardjo, I., Zou, H., Slaughter, C., Wang, X., 1998. Bid, a Bcl2 interacting protein, mediates cytochrome c release from mitochondria in response to activation of cell surface death receptors. Cell 94, 481–490.

Matsuoka, S., Ballif, B.A., Smogorzewska, A., McDonald, E.R., Hurov, K.E., Luo, J., Bakalarski, C.E., Zhao, Z., Solimini, N., Lerenthal, Y., Shiloh, Y., Gygi, S.P., Elledge, S.J., 2007. ATM and ATR substrate analysis reveals extensive protein networks responsive to DNA damage. Science 316, 1160–1166. doi:10.1126/science.1140321

McDonnell, J.M., Fushman, D., Milliman, C.L., Korsmeyer, S.J., Cowburn, D., 1999. Solution structure of the proapoptotic molecule BID: a structural basis for apoptotic agonists and antagonists. Cell 96, 625–634.

Nakamura, K., Greenwood, A., Binder, L., Bigio, E.H., Denial, S., Nicholson, L., Zhou, X.Z., Lu, K.P., 2012. Proline isomer-specific antibodies reveal the early pathogenic tau conformation in Alzheimer's disease. Cell 149, 232–244. doi:10.1016/j.cell.2012.02.016

Nam, E.A., Zhao, R., Glick, G.G., Bansbach, C.E., Friedman, D.B., Cortez, D., 2011. Thr-1989 phosphorylation is a marker of active ataxia telangiectasia-mutated and Rad3-related (ATR) kinase. J Biol Chem 286, 28707–28714. doi:10.1074/jbc.M111.248914

O'Driscoll, M., 2009. Mouse models for ATR deficiency. DNA Repair (Amst) 8, 1333–1337. doi:10.1016/j.dnarep.2009.09.001

Pace, C.N., Scholtz, J.M., 1998. A helix propensity scale based on experimental studies of peptides and proteins. Biophysj 75, 422–427.

Sancar, A., Lindsey-Boltz, L.A., Unsal-Kaçmaz, K., Linn, S., 2004. Molecular mechanisms of mammalian DNA repair and the DNA damage checkpoints. Annu. Rev. Biochem. 73, 39–85. doi:10.1146/annurev.biochem.73.011303.073723

Schoppy, D.W., Ragland, R.L., Gilad, O., Shastri, N., Peters, A.A., Murga, M., Fernandez-Capetillo, O., Diehl, J.A., Brown, E.J., 2012. Oncogenic stress sensitizes murine cancers to hypomorphic suppression of ATR. J Clin Invest 122, 241–252. doi:10.1172/JCI58928

Sorrentino, G., Comel, A., Mantovani, F., Del Sal, G., 2014. Regulation of mitochondrial apoptosis by Pin1 in cancer and neurodegeneration. Mitochondrion 19 Pt A, 88–96. doi:10.1016/j.mito.2014.08.003

Steger, M., Murina, O., Hühn, D., Ferretti, L.P., Walser, R., Hänggi, K., Lafranchi, L., Neugebauer, C., Paliwal, S., Janscak, P., Gerrits, B., Del Sal, G., Zerbe, O., Sartori, A.A., 2013. Prolyl isomerase PIN1 regulates DNA double-strand break repair by counteracting DNA end resection. Mol Cell 50, 333–343. doi:10.1016/j.molcel.2013.03.023

Tait, S.W.G., Green, D.R., 2010. Mitochondria and cell death: outer membrane permeabilization and beyond. Nat Rev Mol Cell Biol 11, 621–632. doi:10.1038/nrm2952

Toko, H., Konstandin, M.H., Doroudgar, S., Ormachea, L., Joyo, E., Joyo, A.Y., Din, S., Gude, N.A., Collins, B., Völkers, M., Thuerauf, D.J., Glembotski, C.C., Chen, C.-H., Lu, K.P., Müller, O.J., Uchida, T., Sussman, M.A., 2013. Regulation of cardiac hypertrophic signaling by prolyl isomerase Pin1. Circulation Research 112, 1244–1252. doi:10.1161/CIRCRESAHA.113.301084

Weber, A.M., Ryan, A.J., 2014. ATM and ATR as therapeutic targets in cancer. Pharmacol. Ther. doi:10.1016/j.pharmthera.2014.12.001

Westphal, D., Kluck, R.M., Dewson, G., 2014. Building blocks of the apoptotic pore: how Bax

and Bak are activated and oligomerize during apoptosis. Cell Death Differ 21, 196–205. doi:10.1038/cdd.2013.139

Wu, X., Shell, S.M., Yang, Z., Zou, Y., 2006. Phosphorylation of nucleotide excision repair factor xeroderma pigmentosum group A by ataxia telangiectasia mutated and Rad3-related-dependent checkpoint pathway promotes cell survival in response to UV irradiation. Cancer Res 66, 2997–3005. doi:10.1158/0008-5472.CAN-05-3403

Yang, J., Yan, R., Roy, A., Xu, D., Poisson, J., Zhang, Y., 2015. The I-TASSER Suite: protein structure and function prediction. Nat Methods 12, 7–8. doi:10.1038/nmeth.3213

Yang, Z.-G., Liu, Y., Mao, L.Y., Zhang, J.-T., Zou, Y., 2002. Dimerization of human XPA and formation of XPA2-RPA protein complex. Biochemistry 41, 13012–13020.

Zacchi, P., Gostissa, M., Uchida, T., Salvagno, C., Avolio, F., Volinia, S., Ronai, Z., Blandino, G., Schneider, C., Del Sal, G., 2002. The prolyl isomerase Pin1 reveals a mechanism to control p53 functions after genotoxic insults. Nature 419, 853–857. doi:10.1038/nature01120

Zeman, M.K., Cimprich, K.A., 2014. Causes and consequences of replication stress. Nat Cell Biol 16, 2–9. doi:10.1038/ncb2897

Zinkel, S.S., Hurov, K.E., Ong, C., Abtahi, F.M., Gross, A., Korsmeyer, S.J., 2005. A role for proapoptotic BID in the DNA-damage response. Cell 122, 579–591. doi:10.1016/j.cell.2005.06.022

Zou, L., Liu, D., Elledge, S.J., 2003. Replication protein A-mediated recruitment and activation of Rad17 complexes. Proc Natl Acad Sci USA 100, 13827–13832. doi:10.1073/pnas.2336100100

Supplemental Data

Figure S1

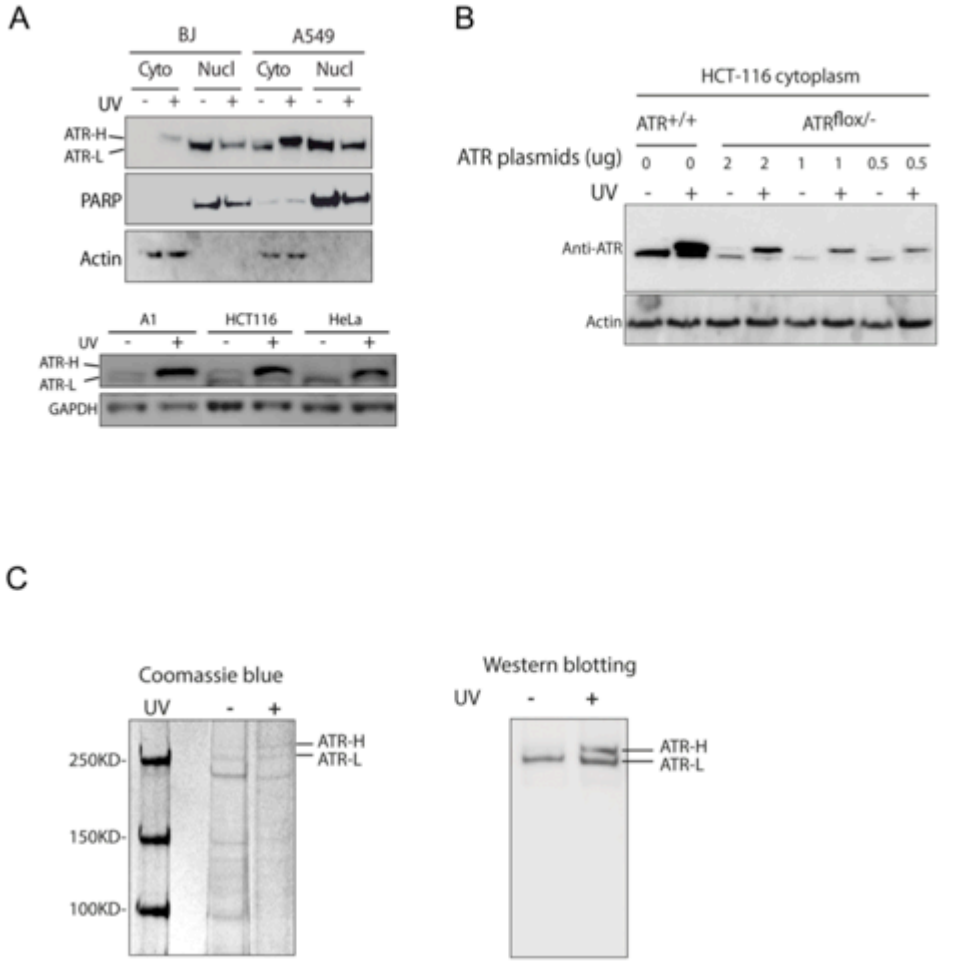


Figure S1. ATR-H forms in different types of cells and ATR-H is confirmed as a modified form of ATR protein, related to Figure 1. (A) UV-induced ATR-H formation occurs in the cytoplasm of multiple human cell lines including primary BJ and A1 (AG08498, Coriell) fibroblasts, and HCT116, A549 and HeLa cancer cell lines. PARP-1 and actin were probed as controls for fractionation (upper). GAPDH serves as a cytoplasmic loading control (lower). **(B)** HCT116 ATR$^{flox/-}$ cells were transiently transfected with the indicated amount of wild-type pcDNA-ATR construct. The transfection generated the same pattern of ATR-H formation in the cytoplasm as occurred in the ATR-proficient cells, HCT116 ATR$^{+/+}$, following UV irradiation and a 2-hr recovery. **(C)** ATR-H identity confirmation by MALDI analysis. Cytoplasmic lysates were collected from UV-treated A549 cells and ATR was immunoprecipitated, separated by 3-8% gradient SDS-PAGE, then the gel was stained with Coomassie blue (left). The two indicated ATR bands, ATR-H and ATR-L, were excised for protein identification by MALDI (Matrix-assisted laser desorption/ionization) mass spectrometry. ATR-H and ATR-L were both identified to be ATR according to the mascot search results. A portion of the immunoprecipitated sample also was analyzed by western blotting using anti-ATR antibody (right), confirming the identity of the ATR-H and ATR-L bands.

Figure S2

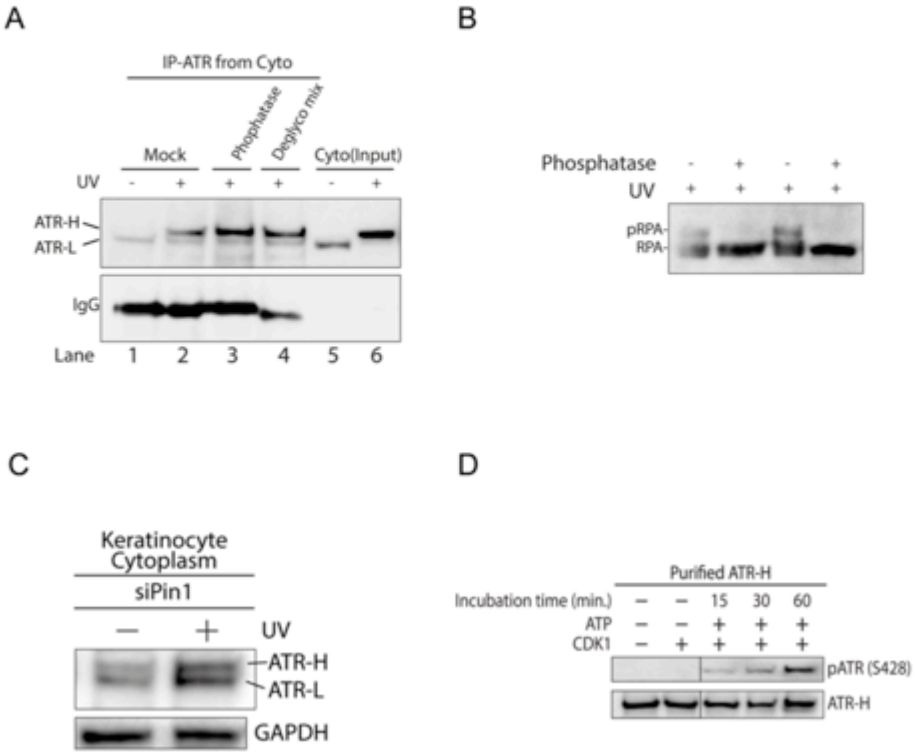


Figure S2. ATR-H formation is negatively regulated by Pin1, related to Figure 1. (A) The DNA damage-induced mobility change of cytoplasmic ATR is not due to phosphorylation or glycosylation. A549 cells were mock- or UV-treated, then cytoplasmic lysates were generated and ATR was immunoprecipitated. The lysates were then treated with enzyme buffer alone (Mock), lambda phosphatase (Phosphatase), or a deglycosylation enzyme mix (Deglyco mix). Detection of the faster-migrating IgG heavy chain indicates that the glycosylation treatment was complete (lane 4). **(B)** To confirm the efficiency of the phosphatase treatment (C, lane 3), hyper-phosphorylated RPA32 was induced by UV-irradiation of a duplicate A549 cell culture, followed by treatment of the lysates with phosphatase to completely remove phosphate groups. **(C)** Pin1 knockdown increases ATR-H formation in the cytoplasm of keratinocytes. siRNA knockdown of peptidylprolyl isomerase Pin1 increases the formation of ATR-H in the cytoplasm of non- irradiated keratinocytes. Cells were transfected with siPin1 for 72 hrs prior to mock treatment or UV exposure, then cells were analyzed by western blot following a 2-hr recovery. **(D)** ATR-H is phosphorylated by Cdk1-cyclin B *in vitro*. Purified ATR-H was treated with Cdk1-cyclin B in the presence or absence of ATP. The reaction with enzyme plus ATP was collected at 3 time points; most of the phosphorylation occurring after 1 hr incubation at 30°C. The PK buffer provided with the Cdk1-cyclin B (NEB) was used for incubation. Phosphorylated ATR was observed by probing with antibody to pATR (S428) and was compared to overall ATR loading.

Figure S3

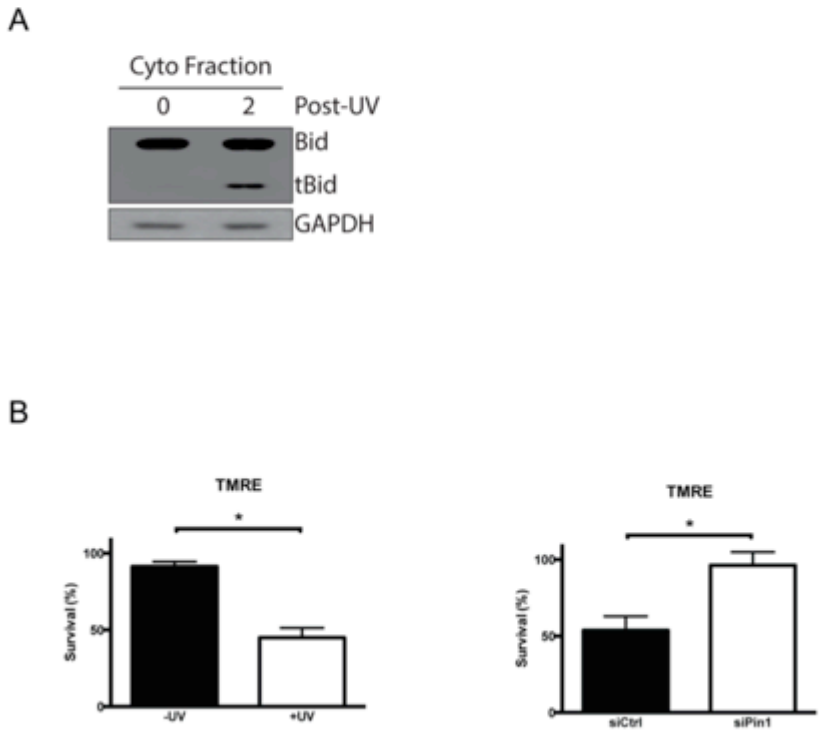


Figure S3. UV treatment induces Bid cleavage and Pin1 knockdown increases survival in UV-treated keratinocytes, related to Figure 5. (A) A549 cells were mock- treated or irradiated with 40 J/m^2 UV-C followed by a 2-hr recovery before cytoplasmic lysates were prepared for western blot analysis. An antibody specific to the C-terminus of Bid allowed detection of Bid and tBid. GAPDH was used as a cytoplasmic loading control. **(B)** Keratinocytes were mock or UV-treated and the percent survival was assessed by the TMRE assay (left). Knockdown of Pin1 increased cell survival in UV-treated keratinocytes (right). Cells were transfected with siPin1 for 72 hrs prior to mock treatment or UV exposure, then cells were analyzed by flow cytometry following a 16-hr recovery in a TMRE assay normalized to siPin1.

Figure S4

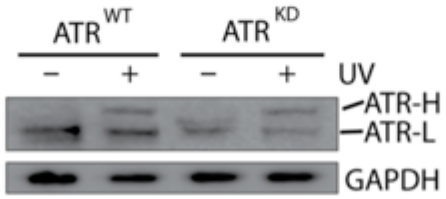


Figure S4. **UV-induced ATR-H formation occurs in cells expressing a kinase dead form of ATR, related to Figure 6.** HCT116 $ATR^{flox/-}$ cells were transiently transfected with wild-type ATR (ATR^{WT}) or kinase dead ATR (ATR^{KD}) constructs. The transfection generated the same pattern of ATR-H formation in the cytoplasm as occurred in the HCT116 $ATR^{+/+}$and $ATR^{flox/-}$ cells (observed in Figure S2A) following UV irradiation and a 2-hr recovery.

Figure S5

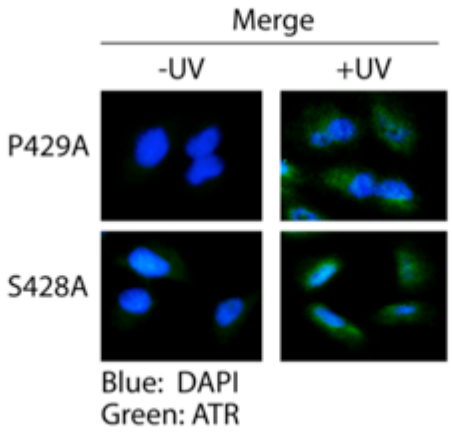


Figure S5. ATR mutants P429A and S428A localize to the cytoplasm following UV treatment, related to Figure 4. HCT116 $ATR^{flox/-}$ cells expressing S428A or P429A mutants of ATR resulted in cytoplasmic localization following UV irradiation and a two- hour recovery. Cells were transfected with mutant ATR plasmids for 48 hours prior to treatment. Nuclei were stained with DAPI and ATR was detected by ATR antibody (Bethyl, A300-138A) followed by incubation with an Alexa Fluor 488 conjugate.

Figure S6

A

Alignment of ATR sequence with Bcl2 BH1 Domain

ATR(556-2572)	T	F	L	H	D	P	L	V	E	W	S	K	P	V	K	G	H
BCL-2	E	L	F	R	D	G	V	-	N	W	G	R	I	V	A	F	F
BCL-XL	E	L	F	R	D	G	V	-	N	W	G	R	I	V	A	F	F
BCL-W	E	L	F	Q	G	G	P	-	N	W	G	R	L	V	A	F	F
MCL-1	H	V	F	S	D	G	V	T	N	W	G	R	I	V	T	L	I
BAX	D	M	F	S	D	G	N	F	N	W	G	R	V	V	A	L	F

Alignment of ATR sequence with Bcl2 BH2 Domain

ATR(2289-2304)	H	W	-	A	Y	I	A	G	F	D	D	M	V	E	I	L
ATR(2378-2393)	E	W	V	N	N	T	A	G	L	R	P	I	L	T	K	L
BCL-2	T	W	I	Q	D	N	G	G	W	D	A	F	V	E	L	Y
BCL-XL	P	W	I	Q	E	N	G	G	W	D	T	F	V	E	L	Y
BCL-W	D	W	I	H	S	S	G	G	W	A	E	F	T	A	L	Y
MCL-1	D	W	L	V	K	Q	R	G	W	D	G	F	V	E	F	F
BAX	G	W	I	Q	D	Q	G	G	W	D	G	L	L	S	Y	F

B

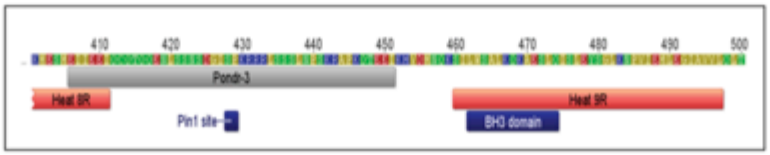


Figure S6. Potential BH-like domains and a proposed organization of the Pin1 site and BH3 domain in ATR, related to Figure 3. (A) Alignment of the ATR protein sequence with BH1 and BH2 domains in various Bcl-2 family proteins. The alignment was maximized by introducing insertions marked by dashes. Purple: identical amino acids; Light blue: conserved amino acids. **(B)** The Pin1 site in ATR occurs in a naturally disordered region while the BH3 domain occurs in a HEAT repeat. A 100-residue region of ATR is shown with the Pin1 site (pS428-P429) indicated. The location of the BH3-like domain is predicted by homology analysis (Fig. 3A) and documented by deletion mapping (Fig. 3C-3F). The predicted naturally disorder region (Pondr-3)[1] and two HEAT repeats (Heat 8R and 9R)[2] are indicated.

(1) Li, X., Romero, P., Rani, M., Dunker, A. K. and Obradovic, Z. (1999). "Predicting Protein Disorder for N-, C-, and Internal Regions." Genome Informatics 10: 30-40. http://www.pondr.com/index

(2) Perry, J. and N. Kleckner (2003). "The ATRs, ATMs, and TORs are giant HEAT repeat proteins." Cell 112(2): 151-155. http://dx.doi.org/10.1016/S0092-8674(03)00033-3

CHAPTER 9

SUMMARY

Maintenance of genome integrity is of the utmost importance to ensure proper cellular function and survival of the organism from one generation to the next. Accumulation of mutations that are not removed could lead to disease or cancer (Papamichos-Chronakis and Peterson 2013). To ensure that genome integrity is maintained cells have evolved a variety of pathways to cope with insults from both the environment and the byproducts of metabolic processes. One such pathway is NER. The NER pathway is able to eliminate a wide range of lesions in part due to the way in which the DNA-damage sensor proteins recognize the DNA opposite the damaged site. By recognizing destabilization of the DNA strand there is no need for a binding pocket that is specific to many different types of damage. This allows for the efficient removal of structurally unrelated lesions (Pakotiprapha et al. 2008; Shell et al. 2013 Sep 14). The $UvrA_2B$ complex and XPC-RAD23B serve to recognize destabilization when the DNA is adducted for the *E. coli* and human NER systems, respectively. The *E. coli* system has been a vital tool in understanding how damage is removed and has facilitated exploration of how both the *E. coli* and human pathways protect the genome. Many types of cancer bypass DNA damage pathways, which results in uncontrolled cell proliferation even in the presence of DNA damage. By better understanding these pathways a more targeted approach can be used to stop rogue cells from developing into tumors.

NER of Bulky Adducts in Prokaryotes

Although the UvrABC system has been studied extensively, there is still important information that can be learned through experimentally manipulating the UvrABC protein system, which is less complicated and purified proteins are readily available. Clarifying the exact mechanism of binding to various types of bulky lesions could lead to new information about protein-DNA interactions and potential avenues for drug therapy. The work presented in Chapter 2 demonstrates a novel 3'-next flanking sequence effect as a unique NER factor for bulky arylamine lesions in *E. coli.* Sequence context has previously been shown to play an important role in damage recognition (Y. Zou et al. 2003; Ruan et al. 2007). Arylamines and heterocyclic amines are well-known environmental mutagens/carcinogens, which have been implicated in the etiology of breast, liver and bladder cancers in humans (Luch 2005). Arylamine adducts are of importance because of their use in industry, usually as intermediates in the synthesis of dyes, rubber products, and other commercial materials (Radomski 1979). Heterocyclic amines are also of interest because of their production in the overcooking of red meat (Cooper and Grover 1990). Metabolic activation of these amines *in vivo* produces C8-substituted dG as the major DNA adducts. One example of this is activation of 4-aminobiphenyl produces ABP. Arylamines produce two major DNA adducts via *in vivo* activation: *N*-(2′- deoxyguanosin-8-yl)-2-aminofluorene (dG-AF) and *N*-(2′- deoxyguanosin-8-yl)-2-acetylaminofluorene (dG-AAF). These three adducts, FABP, FAF and FAAF, are used in this study.

The AF adduct in fully-paired duplex DNA has been shown to adopt an equilibrium of two prototype conformers: 'B-type', in which the carcinogen resides in the major groove of a relatively unperturbed double-helical DNA, and 'stacked (S)', in which the carcinogen is

base-displaced and the glycosidic linkage to the modified guanine is *syn*. An additional W-conformation is found in AAF, due to the addition of an acetyl group on the central nitrogen, in which the fluorene moiety is in the minor groove along with a *syn* glycosidic linkage. Furthermore, S/B-heterogeneity has been shown to be dependent on the flanking sequence, which modulates mutational and repair outcomes (Meneni et al. 2007; Jain et al. 2012). Swapping the polarity of the sequence used in this study *G**CN, where G* is the arylamine-adducted base and N is the base being changed, from A to T produces a dramatic increase in the destabilized stacked (S) conformation, resulting in a 2- to 3-fold lower NER efficiency. This data is consistent with lesion-induced DNA bending/distortion. Additionally, we found that FAAF was repaired three to four times more efficiently than the FABP and FAF lesions. This finding also is consistent with the extent of bending and helix destabilization, as well as the steric constraint in the duplex. The results presented in this study show that lesion-induced DNA bending/thermodynamic destabilization has a greater effect on the ability of the *E. coli* NER machinery to effectively eliminate arylamine-induced DNA damage than the S/B conformational heterogeneity, which was observed previously for AF and AAF in different sequence contexts than the ones tested here. This speaks further to the complexity with which NER recognition factors bind bulky lesions.

The next question that was addressed is how does clustering of arylamine adducts affect DNA-duplex stability and repair outcomes in *E. coli*. Clustered lesions are defined as two or more lesions occurring in a short stretch (<24 bp) of DNA (Ward 1988). It has been hypothesized that cluster lesions are less reparable than individual lesions and that mutation of one lesion is synergized by the presence of another lesion in the vicinity (Shikazono et al. 2006). Past work has demonstrated that the binding spectrum of 2-

acetylaminofluorene (AAF) is essentially random and all guanine residues exhibited equal reactivity toward AAF (Fuchs 1984). Taken together, this suggests that AAF clusters in the genome could cause a dramatic increase in mutagenesis. This was studied within the sequence context of the *Nar*I mutational hotspot (5′-$CG_1G_2CG_3$CC-3′) in which two FAAF-adducted guanines were placed in various combinations.

The findings of this study can be found in Chapter 3. Briefly, our results show that the diadduct duplexes are destabilized to a greater extent than FAAF monoadducts. A sequence and clustering effect was observed in the thermal and thermodynamic instability: *Nar*I-G_2G_3 > *Nar*I-G_1G_3 > *Nar*I-G_1G_2. Base stacking interactions are one of the most important factors that contribute to the thermal and thermodynamic stability of the DNA duplexes. In *Nar*I-G_1G_2 the base stacking was largely maintained while in the others tested there was significant disruption of base stacking. *Nar*I-G_1G_2 also was the most thermodynamically stable confirming the correlation between base stacking and thermodynamic stability.

Contrary to the proposed hypothesis the diadducts lead to enhanced incision rates in *E. coli*. However, this is not surprising because of the greater thermal and thermodynamic destabilization of the di-FAAF versus mono-FAAF duplexes (Jain et al. 2012). We observed different repair efficiencies depending on the location of di-FAAF adducts. The *Nar*I-G_1G_2 duplex showed a considerably lower efficiency than *Nar*I-G_2G_3 and *Nar*I-G_1G_3 [*Nar*I-G_2G_3 (100%) > *Nar*I-G_1G_3 (85%) > *Nar*I-G_1G_2 (44%)]. These NER results agree with the extent of the loss of base stacking and thermal and thermodynamic destabilization presented here. The *Nar*I-G_2G_3 duplex displayed the highest repair rate and also underwent the most severe duplex destabilization and loss of base stacking. In

contrast, the *Nar*I-G_1G_2 duplex showed the weakest reparability, which is consistent with the minimal base stacking and duplex destabilization seen in this adduct. *Nar*I-G_1G_3 showed an intermediary reparability and loss of base stacking in neighboring bases and destabilization of duplex. The greater reparability of the diadducts in *E. coli* was primarily due to their greater thermal and thermodynamic destabilization. Importantly, the incision efficiency occurred in a sequence-specific manner; *i.e.*, highly destacked and destabilized *Nar*I-G_2G_3 and *Nar*I-G_1G_3 are repaired more efficiently than the relatively stable *Nar*I-G_1G_2. These results indicate the importance of carcinogen-induced destacking and related thermal and thermodynamic destabilization in the repair of cluster arylamine adducts. The data presented here suggests that $UvrA_2$ does not contact the adduct directly; rather, the strand opposing the adduct by sensing the degree of destacking of the bases surrounding the lesion and destabilization of the DNA-duplex. The DNA-sensor XPC, in eukaryotes, may also contact lesions in a similar manner by sensing the destabilization of the DNA duplex at the damaged site.

NER of Bulky Adducts in Eukaryotes

To determine if the human NER machinery, primarily XPC, processes the cluster arylamine lesions in a similar way as UvrABC we conducted a study utilizing a human NER assay and surface plasmon resonance (SPR). This work is presented in this dissertation as Chapter 4. Various studies have been conducted to define the correlation of XPC-DNA damage equilibrium binding affinity with NER efficiency (Y.-C. Lee et al. 2014). However, little is known about the effects of XPC-DNA damage recognition kinetics on hNER. Previous work has shown that association of XPC is an important step in repair outcome;

however, our work illustrates that the XPC-DNA dissociation rate also plays a pivotal role in efficiently incising damaged DNA. In this study, we characterized for the first time the binding of XPC to mono- versus di-AAF-modified sequences by using the real-time monitoring SPR technique.

The current body of evidence suggests that XPC randomly binds DNA, damaged or not, and is retained at sites of duplex destabilization (Trego and Turchi 2006; Min and Pavletich 2007). However, other studies have found that the disruption of the incision complex is important for successful NER completion (Bergink et al. 2012). Importantly, in our study the retention time of XPC bound to damaged DNA shares an inverse relationship with NER efficiency. Less stable DNA correlates with higher binding affinity but lower rates of incision. This is particularly true when XPC remained bound to cluster adducts for a much longer period of time as compared to mono-adducts. Various studies have attempted to relate the protein binding interactions involved in DNA damage recognition with NER excision efficiency utilizing a wide array of damage to DNA (Sugasawa et al. 2009; Y.-C. Lee et al. 2014). This work proposes a novel mechanism beyond the conventional concept that the binding capability of DNA damage recognition proteins directly relates to the ability of the repair process to remove the damage from DNA. The presented SPR/hNER results suggest strongly that di-FAAF adducts fail to produce a productive complex for hNER even though the damage recognition binding is strong. In other words, robust XPC binding (K_D of $10^{-9} \sim 10^{-10}$ M) may be required for initiation of hNER. However, unusually strong XPC binding ($K_D < 10^{-11}$) and, more importantly, the extremely slow dissociation of XPC from di-FAAF adducts, and thus the long residence time ($t_{1/2}$) of XPC at the damage site, could be detrimental to recruitment of subsequent downstream proteins to complete the hNER

process. We demonstrate that in addition to the equilibrium binding affinity of XPC for DNA damage, kinetics and the off-rate of the interaction also play critical roles in determining the NER efficiency. This is particularly true for certain types of DNA adducts, such as the di-FAAF examined here, which have a long XPC residence time during DNA damage recognition. Since dissociation of XPC from the damage site after initial recognition is necessary for subsequent binding of other repair factors in the mechanism of NER, it is possible that such an extended residence of XPC at the damage site would likely make the dissociation the rate-limiting step of NER. This could lead to inefficient DNA repair even though the XPC-damaged DNA binding affinity is high. The current work advances our understanding of the intricacies of the NER mechanism by showing that the least stable DNA-adducted duplex is not always the first to be repaired by NER.

Our results suggest that XPC dissociation from the damage site could become a rate-limiting step in NER of certain types of DNA adducts, leading to repression of NER. In recognizing the types of adducts that are comparable to *Nar*1-G_2G_3 and *Nar*I-G_1G_3 cluster adducts, XPC can be stalled on these damage sites, preventing clearance of induced adducts. These data indicate that the slow dissociation rate of XPC from clustered adducts can be exploited along with the high proliferative rate of cancer cells allowing for a more efficiently targeted approach to cancer therapy. Thus, by treating with DNA damaging agents that induce cluster-like DNA lesions XPC will remain bound for a longer period of time, which will halt replication in the cell due to a lack of successful repair of the lesion. Development of a drug that can increase residence time of XPC at the lesion site is also of interest.

Another essential factor in NER is the *Xeroderma pigmentosum* group A (XPA) protein. The importance of XPA in NER is illuminated by evidence demonstrating that of all XP phenotypes, patients deficient in XPA are most severely affected, XPA-deficient mice show a complete deficiency in NER and XPA-deficient cells are most sensitive to UV irradiation (van Steeg et al. 2000; Riedl et al. 2003; C.-J. Park and Choi 2006). XPA has been established as the DNA-damage verification protein in human NER; however, this seemingly dispensable role does not seem to communicate the severe phenotype observed in XPA patients. This suggests that XPA may play additional roles outside of its DNA verification function in NER. XPA was found to bind to ds-ssDNA junctions with higher affinity than binding of XPA to bulky DNA lesions (Yang et al. 2006). DNA junctions also are present in a cell as repair intermediates in NER and other DNA metabolic pathways such as DNA replication.

To further characterize this interaction we determined XPA-junction binding via gel mobility shift assay, limited proteolysis and mass spectrometry. This work is presented in Chapter 6. We found that the previously identified DNA-binding domain (DBD) of XPA made contact with the junction structures; however, there were new contacts that had not been observed when analyzing XPA-adduct binding. These residues are just outside of the accepted binding domain and upon structure prediction analysis appear to be an extension of an alpha-helical region at the C-terminus of XPA. Just prior to publication of our study (Hilton et al. 2014) Sugitani *et al.* reported on an NMR study using Y-shape DNA as a substrate for XPA binding (Sugitani et al. 2014 Jul 24). Their work is in good agreement with our study and suggests that the DBD of XPA be extended from residue 219 to 239. Their NMR data shows that the extended DBD binds to junction DNA much stronger than a

construct with only the traditional shorter binding domain length. Their finding coupled with our finding that the extended XPA DBD binds at a higher affinity to junction DNA as compared with duplex DNA shows not only that XPA's DBD is more extensive than previously thought but also that XPA's main function in the cell could have been initially misinterpreted. We propose that the main function of XPA might actually be to bind and stabilize the ds-ssDNA junction intermediates formed in DNA metabolic reactions such as replication and NER. Binding of these structures by XPA stabilizes the stalled replication fork, preventing fork collapse and apoptosis in HGPS.

XPA Accumulates at DSB Sites in HGPS

In HGPS patient fibroblasts XPA surprisingly mislocalizes to double-strand break (DSB) sites that are the result of progerin accumulation(Y. Liu et al. 2008). In this work it was reported that the DSB repair proteins were not localized at DNA breaks but were diffuse throughout the nucleus of late passaged HGPS. However, as the cells aged in culture XPA went from diffuse in the nucleus to distinct foci that colocalized and immunoprecipitated with γ-H2AX, a marker for DNA strand breaks, indicating that XPA colocalized with DNA breaks. Indeed, knockdown of XPA was enough to restore the localization of DSB repair proteins to the break sites. Our hypothesis is that XPA blocks repair of progerin-induced strand breaks by binding at the stalled replication fork leading to an accumulation of DNA damage.

First, we asked how this new role for XPA binding to junction structures relates to the fate of the replication complex in HGPS. RFC1 is the catalytic subunit of the RFC complex and is responsible for loading PCNA, the sliding clamp that tethers DNA

polymerase to DNA, onto chromatin to ensure processive replication. Analysis of the nuclear extracts of HGPS cells demonstrated that the integrity of RFC1 was compromised in HGPS cells. Specifically, RFC1, full length is 140-kDa, was found to be degraded to a ~75-kDa C-terminal fragment (RFC1-75), suggesting that a proteolytic cleavage of RFC1 occurred in HGPS cells.

Previous studies have shown that RFC1 is proteolyticly cleaved in a caspase-3-dependent manor during apoptosis initiation to inhibit DNA replication (Rhéaume et al. 1997). Although this is not thought to be the mechanism underlying the observations in HGPS cells it does provide evidence that RFC1 can be regulated by proteolytic cleavage to inhibit replication. In the aged HGPS cells there is an inhibition of DNA replication as evidenced by lack of PCNA and polymerase δ on the chromatin. This may be due in part to the sequestration of PCNA by progerin followed by junction binding of XPA, discussed in the following section; conversely, it could be due to RFC1 being proteolyticly cleaved and absent from the chromatin. Pepstain, a potent inhibitor of aspartyl proteases, and E64, which inhibits cysteine peptidases, showed no inhibition of RFC1 cleavage; however, AEBSF, a potent serine protease inhibitor, efficiently inhibited the cleavage in a dose-dependent manner. Remarkably, when the HGPS cells were treated with drugs used to prevent lamin A processing modifications RFC1 was not proteolyticly cleaved. Therefore, there is a direct link between the cleavage of RFC1 and XPA binding at the naked junctions. Determining the protease responsible for RFC1 degradation could be coupled with inhibitors commonly used to prevent accumulation of progerin and provide a new therapeutic treatment for HGPS patients. Additionally, as mentioned earlier, progerin

accumulates in normal human cells also. With this in mind novel treatments for other aging-related diseases could potentially be developed.

In the next study, found in Chapter 6, we find that there is an age-dependent accumulation of XPA to DSB sites and this correlates with progerin accumulation, which sequesters PCNA from stalled replication forks. The data suggest that the PCNA is sequestered by progerin which also accumulates in an age-dependent manor (Y. Liu et al. 2006). Additionally, silencing of PCNA in normal fibroblasts leads to the formation of XPA foci in the nucleus. Because XPA can bind at junction structures we hypothesize that XPA may bind at stalled replication forks to stabilize and prevent fork collapse when PCNA is absent and RFC1 is proteolyticly cleaved. Interestingly, apoptosis was greatly reduced in XPA-deficient patient cells when wild-type XPA was transiently transfected. These data are in agreement with the theory that XPA-junction-binding may serve to prevent fork collapse, and ultimately cell clearance through apoptosis. XPA preventing fork collapse and apoptosis is more in-line with the observed severe phenotype that is observed in patients lacking functional XPA. To further investigate the mechanism behind XPA-DSB binding, ATR/ATM inhibitors were used. In HGPS cells treated with ATR/ATM kinase inhibitors the ATR kinase inhibition drastically prevented accumulation of XPA on chromatin. Therefore, the evidence suggests that ATR is involved in the regulation of XPA binding to DSB sites in HGPS. The inhibition of ATR leading to a decrease in XPA bound to chromatin demonstrates a role for the kinase activity of ATR in prevention of apoptosis. This study provides evidence for ATR as an antiapoptotic factor outside of its traditional role as a DNA damage checkpoint kinase. The next study, found in Chapter 8, provides even more striking evidence that ATR can regulate death through preventing initiation of apoptosis.

Importantly, both studies demonstrate that ATR has a role in the cell independent of its kinase activity.

ATR Functions as an Antiapoptotic Factor

ATR has been shown to play a role in suppression of apoptosis, which makes ATR an attractive target for chemotherapy (Heffernan et al. 2009; Myers et al. 2009; Pearl et al. 2015). For instance, as a DNA-damage checkpoint kinase ATR has been shown in Chapter 7 to be involved in prevention of apoptosis through XPA phosphorylation, which signals XPA to bind junctions preventing cell clearance in HGPS. Other work suggests that ATR and apoptosis are connected through phosphorylation of Chk1 and/or p53, activating downstream factors that regulate apoptosis (Pabla et al. 2008; Myers et al. 2009). Although the role of ATR in the damage checkpoints is well studied and has implications in preventing apoptosis we hypothesized that ATR plays a more direct role in prevention of apoptosis initiation. Indeed, we demonstrated that ATR is localized to mitochondria following UV-induced DNA damage. Moreover, ATR goes from being almost entirely nuclear to being both cytoplasmic and nuclear. Interestingly, we also found that ATR in the cytoplasm migrated slower in an SDS-PAGE gel than the nuclear ATR, suggesting a post-translational modification of ATR.

Through extensive work we determined that ATR was not being modified in response to UV-induced DNA damage by factors previously shown to modify ATR, such as phosphorylation, ubiquitination, etc. We found instead that the shift reflects a conformational change by prolyl isomerization of ATR regulated by Pin1. In the absence of Pin1 isomerization activity, cytoplasmic ATR was predominantly in the *cis*-conformation.

Further confirming this result, purification of recombinant ATR resulted in the higher, *cis*-conformation of ATR. When a protein is overexpressed, the regulatory mechanisms, or modifications, of the host cell are often bypassed. The *cis*-confirmation was further confirmed to be the native form of ATR in an *in vitro* isomerization assay, which demonstrated that phosphorylation of ATR is required for ATR interaction and isomerization by Pin1. Analysis of Pin1 following UV-treatment revealed that Pin1 is inactivated by phosphorylation of S71 by the DAPK1.

An alignment analysis of proteins that are known to interact with the mitochondria in the apoptotic response showed that ATR contains a BH3-like domain, which is known to be required for interaction between pro- and antiapoptotic factors. Alignment analysis also confirmed the presence of BH1 and BH2 domains, which also are involved in this interaction. Interestingly, the BH3-like domain is near the Pin1 isomerization site and part of a HEAT repeat domain. The HEAT repeats are usually found in cytoplasmic proteins and have been suggested to be involved in protein-protein interaction. An analysis of the protein structure stability also revealed that the Pin1 isomerization site is in a naturally disordered region. This is of interest because this site would likely be surface exposed and would be easily accessible for modification by phosphorylation or isomerization. Following UV damage, Pin1 is inactivated and ATR returns to the ATR-H (*cis*ATR) confirmation. In the ATR-H confirmation the BH3-like domain is likely to be surface exposed, whereas in the ATR-L confirmation the BH3-like domain is hidden. When this domain is exposed ATR can interact with tBid at the outer mitochondrial membrane. ATR uses the BH3-like domain to bind the proapoptotic protein tBid. tBid is known to be involved in Bax and Bak oligomerization which initiates the apoptotic response through cytochrome c release. We

proposed and demonstrated that ATR can prevent cytochrome c release by binding to tBid. These data confirm that ATR is directly involved in suppression of apoptosis that is DNA-damage-induced.

Since all of the known functions of ATR require kinase activity we explored whether or not the kinase activity was required for the newfound function of ATR at the mitochondria. Surprisingly, we found that the kinase activity was not required. Furthermore, the function of ATR at mitochondria is also independent of ATRIP, which has been shown to be required for the nuclear functions and kinase activity of ATR (Cortez et al. 2001). Our findings demonstrate for the first time that ATR is involved in pathways outside of the nucleus and function independent of its kinase activity. Our data suggest that ATR acts to prevent early apoptosis, allowing time for the nucleus to activate the appropriate factors to repair the damaged DNA. We think that ATR maintains cell homeostasis following introduction of DNA damage in a coordinated manner between the cytoplasm and nucleus. A clinical confirmation of our finding was published recently, where ATR signaling in the cytoplasm was reported to be turned off in aggressive cancers. This shows that targeting ATR in aggressive cancers could facilitate elimination of the cells by initiating apoptosis.

In summary, the work presented in this dissertation has focused on the exploration of the NER pathway in an effort to better understand the factors involved so that novel targets for the treatment of cancer can be identified. Comparison of *E. coli* and human NER pathways demonstrated that exploitation of XPC dissociation from DNA damage sites could be a novel target when used in combination with cisplatin-based drugs to more efficiency kill cancer cells. XPA also was shown to be an important target in the progression of aging,

specifically its ability to prevent clearance of cells that have accumulated DNA strand breaks. Finally, ATR was presented as an antiapoptoic factor that could be used as a target for chemotherapy by inhibiting ATR's antiapopotic functions in tumors.

REFERENCES

Abel EL, Hendrix SO, McNeeley SG, Johnson KC, Rosenberg CA, Mossavar-Rahmani Y, Vitolins M, Kruger M. 2007. Daily coffee consumption and prevalence of nonmelanoma skin cancer in Caucasian women. Eur. J. Cancer Prev. 16:446–452.

Abraham RT. 2001. Cell cycle checkpoint signaling through the ATM and ATR kinases. Genes Dev. 15:2177–2196.

Abuaf P, Hingerty BE, Broyde S, Grunberger D. 1995. Solution Conformation of the N-(Deoxyguanosin-8-yl)aminofluorene Adduct opposite Deoxyinosine and Deoxyguanosine in DNA by NMR and Computational Characterization. Chem. Res. Toxicol. 8:369–378.

Adams JM, Cory S. 1998. The Bcl-2 protein family: arbiters of cell survival. Science 281:1322–1326.

Al-Khalaf HH, Hendrayani S-F, Aboussekhra A. 2012. ATR controls the p21(WAF1/Cip1) protein up-regulation and apoptosis in response to low UV fluences. Mol. Carcinog. 51:930–938.

Aliper AM, Csoka AB, Buzdin A, Jetka T, Roumiantsev S, Moskalev A, Zhavoronkov A. 2015. Signaling pathway activation drift during aging: Hutchinson-Gilford Progeria Syndrome fibroblasts are comparable to normal middle-age and old-age cells. Aging 7:26–37.

Anderson KE, Hammons GJ, Kadlubar FF, Potter JD, Kaderlik KR, Ilett KF, Minchin RF, Teitel CH, Chou HC, Martin MV, et al. 1997. Metabolic activation of aromatic amines by human pancreas. Carcinogenesis 18:1085–1092.

Ayala G, Wang D, Wulf G, Frolov A, Li R, Sowadski J, Wheeler TM, Lu KP, Bao L. 2003. The prolyl isomerase Pin1 is a novel prognostic marker in human prostate cancer. Cancer Res. 63:6244–6251.

Aye Y, Li M, Long MJC, Weiss RS. 2014 Jun 9. Ribonucleotide reductase and cancer: biological mechanisms and targeted therapies. Oncogene.

Banavali NK, Mackerell AD. 2002. Free energy and structural pathways of base flipping in a DNA GCGC containing sequence. J. Mol. Biol. 319:141–160.

Bannwarth S, Figueroa A, Fragaki K, Destroismaisons L, Lacas-Gervais S, Lespinasse F, Vandenbos F, Pradelli LA, Ricci J-E, Rötig A, et al. 2012. The human MSH5 (MutSHomolog 5) protein localizes to mitochondria and protects the mitochondrial genome from oxidative damage. Mitochondrion 12:654–665.

Barclay LA, Wales TE, Garner TP, Wachter F, Lee S, Guerra RM, Stewart ML, Braun CR, Bird GH, Gavathiotis E, et al. 2015. Inhibition of Pro-Apoptotic BAX by a Noncanonical Interaction Mechanism. Mol. Cell 57:873–886.

Bartkova J, Horejsí Z, Koed K, Krämer A, Tort F, Zieger K, Guldberg P, Sehested M, Nesland JM, Lukas C, et al. 2005. DNA damage response as a candidate anti-cancer barrier in early human tumorigenesis. Nature 434:864–870.

Basu AK, Essigmann JM. 1988. Site-specifically modified oligodeoxynucleotides as probes for the structural and biological effects of DNA-damaging agents. Chem. Res. Toxicol. 1:1–18.

Batty DP, Wood RD. 2000. Damage recognition in nucleotide excision repair of DNA. Gene 241:193–204.

Beland FA, Kadlubar FF. 1990. Metabolic Activation and DNA Adducts of Aromatic Amines and Nitroaromatic Hydrocarbons. In: Cooper CS; Grover PL, editors. Chemical Carcinogenesis and Mutagenesis I. Vol. 94 / 1. Berlin, Heidelberg: Springer Berlin Heidelberg. (Handbook of Experimental Pharmacology). pp. 267–325; 59 p.

Bergink S, Toussaint W, Luijsterburg MS, Dinant C, Alekseev S, Hoeijmakers JHJ, Dantuma NP, Houtsmuller AB, Vermeulen W. 2012. Recognition of DNA damage by XPC coincides with disruption of the XPC-RAD23 complex. J. Cell Biol. 196:681–688.

Bermudez VP, Maniwa Y, Tappin I, Ozato K, Yokomori K, Hurwitz J. 2003. The alternative Ctf18-Dcc1-Ctf8-replication factor C complex required for sister chromatid cohesion loads proliferating cell nuclear antigen onto DNA. Proc. Natl. Acad. Sci. 100:10237–10242.

Bernis C, Vigneron S, Burgess A, Labbé J-C, Fesquet D, Castro A, Lorca T. 2007. Pin1 stabilizes Emi1 during G2 phase by preventing its association with SCF(betatrcp). EMBO Rep. 8:91–98.

Beveridge DL, DiCapua FM. 1989. Free energy via molecular simulation: applications to chemical and biomolecular systems. Annu. Re.v Biophys Biophys Chem. 18:431–492.

Billen LP, Shamas-Din A, Andrews DW. 2008. Bid: a Bax-like BH3 protein. Oncogene 27 Suppl 1:S93–104.

Blazy B, Culard F, Maurizot JC. 1987. Interaction between the cyclic AMP receptor protein and DNA. J. Mol. Biol. 195:175–183.

Boldt J, Mah MC, Wang YC, Smith BA, Beland FA, Maher VM, McCormick JJ. 1991. Kinds of mutations found when a shuttle vector containing adducts of 1,6-dinitropyrene replicates in human cells. Carcinogenesis 12:119–126.

Bomgarden RD, Lupardus PJ, Soni DV, Yee M-C, Ford JM, Cimprich KA. 2006. Opposing

effects of the UV lesion repair protein XPA and UV bypass polymerase eta on ATR checkpoint signaling. EMBO J. 25:2605–2614.

Bootsma D, Hoeijmakers JH. 1991. The genetic basis of xeroderma pigmentosum. Ann. Genet. 34:143–150.

Bradshaw JM, McFarland JM, Paavilainen VO, Bisconte A, Tam D, Phan VT, Romanov S, Finkle D, Shu J, Patel V, et al. 2015. Prolonged and tunable residence time using reversible covalent kinase inhibitors. Nat. Chem. Biol. 11:525–531.

Bridger JM, Kill IR. 2004. Aging of Hutchinson-Gilford progeria syndrome fibroblasts is characterised by hyperproliferation and increased apoptosis. Exp. Gerontol. 39:717–724.

Brooks BR, Brooks CL, MacKerell AD, Nilsson L, Petrella RJ, Roux B, Won Y, Archontis G, Bartels C, Boresch S, et al. 2009. CHARMM: the biomolecular simulation program. Brooks CL III; Case DA, editors. J. Comput. Chem. 30:1545–1614.

Broschard TH, Koffel-Schwartz N, Fuchs RPP. 1999. Sequence-dependent modulation of frameshift mutagenesis at NarI-derived mutation hot spots. J. Mol. Biol. 288:191–199.

Brown EJ, Baltimore D. 2000. ATR disruption leads to chromosomal fragmentation and early embryonic lethality. Genes Dev. 14:397–402.

Brown EJ, Baltimore D. 2003. Essential and dispensable roles of ATR in cell cycle arrest and genome maintenance. Genes Dev. 17:615–628.

Brown K, Hingerty BE, Guenther EA, Krishnan VV, Broyde S, Turteltaub KW, Cosman M. 2001. Solution structure of the 2-amino-1- methyl-6-phenylimidazo[4,5-b]pyridine C8-deoxyguanosine adduct in duplex DNA. Proc. Natl. Acad. Sci. 98:8507–8512.

Buchko GW, Tung CS, McAteer K, Isern NG, Spicer LD, Kennedy MA. 2001. DNA-XPA interactions: a (31)P NMR and molecular modeling study of dCCAATAACC association with the minimal DNA-binding domain (M98-F219) of the nucleotide excision repair protein XPA. Nucleic Acids Res. 29:2635–2643.

Buisson R, Boisvert JL, Benes CH, Zou L. 2015. Distinct but Concerted Roles of ATR, DNA-PK, and Chk1 in Countering Replication Stress during S Phase. Mol. Cell 59:1011–1024.

Burnouf D, Koehl P, Fuchs RP. 1989. Single adduct mutagenesis: strong effect of the position of a single acetylaminofluorene adduct within a mutation hot spot. Proc. Natl. Acad. Sci. 86:4147–4151.

Busch A, Kiel T, Heupel W-M, Wehnert M, Hübner S. 2009. Nuclear protein import is reduced in cells expressing nuclear envelopathy-causing lamin A mutants. Exp. Cell

Res. 315:2373–2385.

Bussi G, Donadio D, Parrinello M. 2007. Canonical sampling through velocity rescaling. J. of Chemical Physics 126:014101.

Bussi G, Zykova-Timan T, Parrinello M. 2009. Isothermal-isobaric molecular dynamics using stochastic velocity rescaling. J. of Chemical Physics 130:074101.

Cai Y, Kropachev K, Xu R, Tang Y, Kolbanovskii M, Kolbanovskii A, Amin S, Patel DJ, Broyde S, Geacintov NE. 2010. Distant neighbor base sequence context effects in human nucleotide excision repair of a benzo[a]pyrene-derived DNA lesion. J. Mol. Biol. 399:397–409.

Cai Y, Patel DJ, Broyde S, Geacintov NE. 2010. Base sequence context effects on nucleotide excision repair. J. Nucleic Acids 2010:1–9.

Camenisch U, Dip R, Schumacher SB, Schuler B, Naegeli H. 2006. Recognition of helical kinks by xeroderma pigmentosum group A protein triggers DNA excision repair. Nat. Struct. Mol. Biol. 13:278–284.

Camenisch U, Dip R, Vitanescu M, Naegeli H. 2007. Xeroderma pigmentosum complementation group A protein is driven to nucleotide excision repair sites by the electrostatic potential of distorted DNA. DNA Repair 6:1819–1828.

Camozzi D, Capanni C, Cenni V, Mattioli E, Columbaro M, Squarzoni S, Lattanzi G. 2014. Diverse lamin-dependent mechanisms interact to control chromatin dynamics. 5:427–440.

Cao K, Blair CD, Faddah DA, Kieckhaefer JE, Olive M, Erdos MR, Nabel EG, Collins FS. 2011 Jun 13. Progerin and telomere dysfunction collaborate to trigger cellular senescence in normal human fibroblasts. J. Clin. Invest.

Capell BC, Collins FS. 2006. Human laminopathies: nuclei gone genetically awry. Nat. Rev. Genet. 7:940–952.

Caron PR, Kushner SR, Grossman L. 1985. Involvement of helicase II (uvrD gene product) and DNA polymerase I in excision mediated by the uvrABC protein complex. Proc. Natl. Acad. Sci. 82:4925–4929.

Chakrabarti MC, Schwarz FP. 1999. Thermal stability of PNA/DNA and DNA/DNA duplexes by differential scanning calorimetry. Nucleic Acids Res. 27:4801–4806.

Chen S, Levin MK, Sakato M, Zhou Y, Hingorani MM. 2009. Mechanism of ATP-driven PCNA clamp loading by S. cerevisiae RFC. J. Mol. Biol. 388:431–442.

Chen X, Velmurugu Y, Zheng G, Park B, Shim Y, Kim Y, Liu L, Van Houten B, He C, Ansari A, et al. 2015. Kinetic gating mechanism of DNA damage recognition by Rad4/XPC. Nat.

Commun. 6:5849.

Chibazakura T, Kamachi K, Ohara M, Tane S, Yoshikawa H, Roberts JM. 2011. Cyclin A promotes S-phase entry via interaction with the replication licensing factor Mcm7. Mol. Cell Biol 31:248–255.

Cho B. 2010. Structure–Function Characteristics of Aromatic Amine-DNA Adducts. The Chemical biology of DNA damage.

Cho BP. 2004. Dynamic conformational heterogeneities of carcinogen-DNA adducts and their mutagenic relevance. J. of Environmental Science and Health.

Cho BP, Zhou L. 1999. Probing the conformational heterogeneity of the acetylaminofluorene-modified 2'-deoxyguanosine and DNA by 19F NMR spectroscopy. Biochemistry 38:7572–7583.

Chou JJ, Li H, Salvesen GS, Yuan J, Wagner G. 1999. Solution structure of BID, an intracellular amplifier of apoptotic signaling. Cell 96:615–624.

Cimprich KA, Cortez D. 2008. ATR: an essential regulator of genome integrity. Nat. Rev. Mol. Cell Biol. 9:616–627.

Cleaver JE, States JC. 1997. The DNA damage-recognition problem in human and other eukaryotic cells: the XPA damage binding protein. Biochem. J. 328 (Pt 1):1–12.

Connor F, Cary PD, Read CM, Preston NS, Driscoll PC, Denny P, Crane-Robinson C, Ashworth A. 1994. DNA binding and bending properties of the postmeiotically expressed Sry-related protein Sox-5. Nucleic Acids Res. 22:3339–3346.

Cooper CS, Grover PL, editors. 1990. Chemical Carcinogenesis and Mutagenesis I. Berlin, Heidelberg: Springer Berlin Heidelberg.

Cooper MA. 2003. Label-free screening of bio-molecular interactions. Anal. Bioanal. Chem. 377:834–842.

Corrigan DP, Kuszczak D, Rusinol AE, Thewke DP, Hrycyna CA, Michaelis S, Sinensky MS. 2005. Prelamin A endoproteolytic processing in vitro by recombinant Zmpste24. Biochem. J. 387:129–138.

Cortez D. 2003. Caffeine inhibits checkpoint responses without inhibiting the ataxia-telangiectasia-mutated (ATM) and ATM- and Rad3-related (ATR) protein kinases. J. of Biological Chemistry 278:37139–37145.

Cortez D, Guntuku S, Qin J, Elledge SJ. 2001. ATR and ATRIP: partners in checkpoint signaling. Science 294:1713–1716.

Cox MM. 2002. The nonmutagenic repair of broken replication forks via recombination. Mutat. Res. 510:107–120.

Czabotar PE, Westphal D, Dewson G, Ma S, Hockings C, Fairlie WD, Lee EF, Yao S, Robin AY, Smith BJ, et al. 2013. Bax crystal structures reveal how BH3 domains activate Bax and nucleate its oligomerization to induce apoptosis. Cell 152:519–531.

Darden T, York D, Pedersen L. 1993. Particle mesh Ewald: An N·log(N) method for Ewald sums in large systems. J. of Chemical Physics 98:10089–10092.

de Klein A, Muijtjens M, van Os R, Verhoeven Y, Smit B, Carr AM, Lehmann AR, Hoeijmakers JH. 2000. Targeted disruption of the cell-cycle checkpoint gene ATR leads to early embryonic lethality in mice. Curr. Biol. 10:479–482.

de Laat WL, Jaspers N. 1999. Molecular mechanism of nucleotide excision repair. Genes.

De Sandre-Giovannoli A, Bernard R, Cau P, Navarro C, Amiel J, Boccaccio I, Lyonnet S, Stewart CL, Munnich A, Le Merrer M, et al. 2003. Lamin a truncation in Hutchinson-Gilford progeria. Science 300:2055.

Dechat T, Adam SA, Taimen P, Shimi T, Goldman RD. 2010. Nuclear lamins. Cold Spring Harb. Perspect Biol. 2:a000547.

Delagoutte E, Bertrand-Burggraf E, Lambert IB, Fuchs RPP. 1996. Binding and Incision Activities of UvrABC Excinuclease on Slipped DNA Intermediates that Generate Frameshift Mutations. J. Mol. Biol. 257:970–976.

Dionne I, Brown NJ, Woodgate R, Bell SD. 2008. On the mechanism of loading the PCNA sliding clamp by RFC. Mol. Microbiol. 68:216–222.

Doll R, Peto R. 1981. The causes of cancer: quantitative estimates of avoidable risks of cancer in the United States today. J. Natl. Cancer Inst. 66:1191–1308.

Driver JA, Lu KP. 2010. Pin1: a new genetic link between Alzheimer's disease, cancer and aging. Curr. Aging Sci. 3:158–165.

Eckel LM, Krugh TR. 1994. Structural Characterization of Two Interchangeable Conformations of an N-2-Aminofluorene-Modified DNA Oligomer by NMR and Energy Minimization. Biochemistry 33:13611–13624.

Ellison V, Stillman B. 1998. Reconstitution of recombinant human replication factor C (RFC) and identification of an RFC subcomplex possessing DNA-dependent ATPase activity. J. Biol. Chem. 273:5979–5987.

Elmquist CE, Wang F, Stover JS, Stone MP, Rizzo CJ. 2007. Conformational differences of the C8-deoxyguanosine adduct of 2-amino-3-methylimidazo[4,5-f]quinoline (IQ) within the NarI recognition sequence. Chem. Res. Toxicol. 20:445–454.

Eriksson M, Brown WT, Gordon LB, Glynn MW, Singer J, Scott L, Erdos MR, Robbins CM, Moses TY, Berglund P, et al. 2003. Recurrent de novo point mutations in lamin A

cause Hutchinson-Gilford progeria syndrome. Nature 423:293–298.

Errico A, Costanzo V. 2012. Mechanisms of replication fork protection: a safeguard for genome stability. Crit. Rev. Biochem. Mol. Biol. 47:222–235.

Evans E. 1997. Open complex formation around a lesion during nucleotide excision repair provides a structure for cleavage by human XPG protein. EMBO J. 16:625–638.

Fang EF, Scheibye-Knudsen M, Brace LE, Kassahun H, SenGupta T, Nilsen H, Mitchell JR, Croteau DL, Bohr VA. 2014. Defective Mitophagy in XPA via PARP-1 Hyperactivation and NAD+/SIRT1 Reduction. Cell 157:882–896.

Feig M, Onufriev A, Lee MS, Im W, Case DA, Brooks CL. 2004. Performance comparison of generalized born and Poisson methods in the calculation of electrostatic solvation energies for protein structures. J. Comput. Chem. 25:265–284.

Fischer G, Bang H, Mech C. 1984. [Determination of enzymatic catalysis for the cis-trans-isomerization of peptide binding in proline-containing peptides]. Biomed. Biochim. Acta 43:1101–1111.

Fischer G, Schmid FX. 1990. The mechanism of protein folding. Implications of in vitro refolding models for de novo protein folding and translocation in the cell. Biochemistry 29:2205–2212.

Follis AV, Llambi F, Merritt P, Chipuk JE, Green DR, Kriwacki RW. 2015. Pin1-Induced Proline Isomerization in Cytosolic p53 Mediates BAX Activation and Apoptosis. Mol. Cell 59:677–684.

Frezza C, Cipolat S, Scorrano L. 2007. Organelle isolation: functional mitochondria from mouse liver, muscle and cultured fibroblasts. Nat. Protoc. 2:287–295.

Friedberg EC. 2005. Suffering in silence: the tolerance of DNA damage. Nat Rev Mol. Cell Biol. 6:943–953.

Friedberg EC, Walker GC, Siede W, Wood RD. 2005. DNA Repair and Mutagenesis. American Society for Microbiology.

Fuchs RP, Seeberg E. 1984. pBR322 plasmid DNA modified with 2-acetylaminofluorene derivatives: transforming activity and in vitro strand cleavage by the Escherichia coli uvrABC endonuclease. EMBO J. 3:757–760.

Fuchs RPP. 1984. DNA binding spectrum of the carcinogen N-acetoxy-N-2-acetylaminofluorene significantly differs from the mutation spectrum. J. Mol. Biol. 177:173–180.

Gao L, Zhang L, Cho BP, Chiarelli MP. 2008. Sequence verification of oligonucleotides containing multiple arylamine modifications by enzymatic digestion and liquid

chromatography mass spectrometry (LC/MS). J. of the American Society for Mass Spectrometry 19:1147–1155.

Geacintov NE, Broyde S, Buterin T, Naegeli H, Wu M, Yan S, Patel DJ. 2002. Thermodynamic and structural factors in the removal of bulky DNA adducts by the nucleotide excision repair machinery. Biopolymers 65:202–210.

Geacintov NE, Cosman M, Hingerty BE, Amin S, Broyde S, Patel DJ. 1997. NMR solution structures of stereoisometric covalent polycyclic aromatic carcinogen-DNA adduct: principles, patterns, and diversity. Chem. Res. Toxicol. 10:111–146.

Gillet LCJ, Scharer OD. 2006. Molecular mechanisms of mammalian global genome nucleotide excision repair. Chem. Rev. 106:253–276.

Gilljam KM, Feyzi E, Aas PA, Sousa MML, Muller R, Vagbo CB, Catterall TC, Liabakk NB, Slupphaug G, Drablos F, et al. 2009. Identification of a novel, widespread, and functionally important PCNA-binding motif. J. Cell Biol. 186:645–654.

Goldman RD, Gruenbaum Y, Moir RD, Shumaker DK, Spann TP. 2002. Nuclear lamins: building blocks of nuclear architecture. Genes Dev. 16:533–547.

Gomes XV, Henricksen LA, Wold MS. 1996. Proteolytic mapping of human replication protein A: evidence for multiple structural domains and a conformational change upon interaction with single-stranded DNA. Biochemistry 35:5586–5595.

Gonzalo S, Kreienkamp R. 2015. DNA repair defects and genome instability in Hutchinson-Gilford Progeria Syndrome. Current Opinion in Cell Biology 34:75–83.

Gordon LB, Rothman FG, López-Otín C, Misteli T. 2014. Progeria: A Paradigmfor Translational Medicine. Cell 156:400–407.

Gourdin AM, van Cuijk L, Tresini M, Luijsterburg MS, Nigg AL, Giglia-Mari G, Houtsmuller AB, Vermeulen W, Marteijn JA. 2014. Differential binding kinetics of replication protein A during replication and the pre- and post-incision steps of nucleotide excision repair. DNA Repair 24:46–56.

Greenblatt MS, Bennett WP, Hollstein M, Harris CC. 1994. Mutations in the p53 tumor suppressor gene: clues to cancer etiology and molecular pathogenesis. Cancer Res. 54:4855–4878.

Guengerich FP. 2006. Interactions of Carcinogen-Bound DNA with Individual DNA Polymerases. Chem. Rev. 106:420–452.

Guzder SN, Sommers CH, Prakash L, Prakash S. 2006. Complex formation with damage recognition protein Rad14 is essential for Saccharomyces cerevisiae Rad1-Rad10 nuclease to perform its function in nucleotide excision repair in vivo. Mol. Cell Biol. 26:1135–1141.

Harper JW, Elledge SJ. 2007. The DNA damage response: ten years after. Mol. Cell 28:739–745.

Hart K, Foloppe N, Baker CM, Denning EJ, Nilsson L, MacKerell AD Jr. 2012. Optimization of the CHARMM Additive Force Field for DNA: Improved Treatment of the BI/BII Conformational Equilibrium. J. of Chemical Theory and Computation 8:348–362.

Hashiguchi K, Matsumoto Y, Yasui A. 2007. Recruitment of DNA repair synthesis machinery to sites of DNA damage/repair in living human cells. Nucleic Acids Res. 35:2913–2923.

Heffernan TP, Kawasumi M, Blasina A, Anderes K, Conney AH, Nghiem P. 2009. ATR-Chk1 pathway inhibition promotes apoptosis after UV treatment in primary human keratinocytes: potential basis for the UV protective effects of caffeine. J. Invest. Dermatol. 129:1805–1815.

Heflich RH, Neft RE. 1994. Genetic toxicity of 2-acetylaminofluorene, 2-aminofluorene and some of their metabolites and model metabolites. Mutation Research/Reviews in Genetic Toxicology 318:73–174.

Heller RC, Marians KJ. 2006. Replisome assembly and the direct restart of stalled replication forks. Nat Rev Mol. Cell Biol. 7:932–943.

Hess B. 2008. P-LINCS: A Parallel Linear Constraint Solver for Molecular Simulation. J. of Chemical Theory and Computation 4:116–122.

Hess B, Bekker H, Berendsen HJC, Fraaije JGEM. 1997. LINCS: A linear constraint solver for molecular simulations. J. Comput. Chem. 18:1463–1472.

Hess B, Kutzner C, van der Spoel D, Lindahl E. 2008. GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation. J. of Chemical Theory and Computation 4:435–447.

Heyduk E, Heyduk T. 1994. Mapping protein domains involved in macromolecular interactions: a novel protein footprinting approach. Biochemistry 33:9643–9650.

Hilton B, Shkriabai N, Musich PR, Kvaratskhelia M, Shell S, Zou Y. 2014. A new structural insight into XPA-DNA interactions. Biosci. Rep. 34:e00162–840.

Hinderaker MP, Raines RT. 2003. An electronic effect on protein structure. Protein Sci. 12:1188–1194.

Hoeijmakers JH. 2001. Genome maintenance mechanisms for preventing cancer. Nature 411:366–374.

Holt LJ, Tuch BB, Villén J, Johnson AD, Gygi SP, Morgan DO. 2009. Global analysis of Cdk1 substrate phosphorylation sites provides insights into evolution. Science 325:1682–

1686.

Humphrey W, Dalke A, Schulten K. 1996. VMD: Visual molecular dynamics. J. of Mol. Graphics 14:33–38.

Hunter T. 1998. Prolyl isomerases and nuclear function. Cell 92:141–143.

Huntoon CJ, Flatten KS, Wahner Hendrickson AE, Huehls AM, Sutor SL, Kaufmann SH, Karnitz LM. 2013. ATR inhibition broadly sensitizes ovarian cancer cells to chemotherapy independent of BRCA status. Cancer Res. 73:3683–3691.

Hutchison CJ. 2011. The role of DNA damage in laminopathy progeroid syndromes. Biochem. Soc. Trans. 39:1715–1718.

Iakoucheva LM, Kimzey AL, Masselon CD, Bruce JE, Garner EC, Brown CJ, Dunker AK, Smith RD, Ackerman EJ. 2001. Identification of intrinsic order and disorder in the DNA repair protein XPA. Protein Sci. 10:560–571.

Ikegami T, Kuraoka I, Saijo M, Kodo N, Kyogoku Y, Morikawa K, Tanaka K, Shirakawa M. 1998. Solution structure of the DNA- and RPA-binding domain of the human repair factor XPA. Nat. Struct. Biol. 5:701–706.

Ikegami T, Kuraoka I, Saijo M, Kodo N, Kyogoku Y, Morikawa K, Tanaka K, Shirakawa M. 1999. Resonance assignments, solution structure, and backbone dynamics of the DNA- and RPA-binding domain of human repair factor XPA. J. Biochem. 125:495–506.

Jaciuk M, Nowak E, Skowronek K, Tańska A, Nowotny M. 2011. Structure of UvrA nucleotide excision repair protein in complex with modified DNA. Nature 18:191–197.

Jain N, Li Y, Zhang L, Meneni SR, Cho BP. 2007. Probing the sequence effects on NarI-induced -2 frameshift mutagenesis by dynamic 19F NMR, UV, and CD spectroscopy. Biochemistry 46:13310–13321.

Jain N, Meneni S, Jain V, Cho BP. 2009. Influence of flanking sequence context on the conformational flexibility of aminofluorene-modified dG adduct in dA mismatch DNA duplexes. Nucleic Acids Res. 37:1628–1637.

Jain V, Hilton B, Lin B, Jain A, Mackerell AD, Zou Y, Cho BP. 2013. Structural and thermodynamic insight into Escherichia coli UvrABC-mediated incision of cluster diacetylaminofluorene adducts on the NarI sequence. Chem. Res. Toxicol. 26:1251–1262.

Jain V, Hilton B, Lin B, Patnaik S, Liang F, Darian E, Zou Y, Mackerell AD, Cho BP. 2013. Unusual sequence effects on nucleotide excision repair of arylamine lesions: DNA bending/distortion as a primary recognition factor. Nucleic Acids Res. 41:869–880.

Jain V, Hilton B, Patnaik S, Zou Y, Chiarelli MP, Cho BP. 2012. Conformational and thermodynamic properties modulate the nucleotide excision repair of 2-aminofluorene and 2-acetylaminofluorene dG adducts in the NarI sequence. Nucleic Acids Res. 40:3939–3951.

Jain V, Vaidyanathan VG, Patnaik S, Gopal S, Cho BP. 2014. Conformational insights into the lesion and sequence effects for arylamine-induced translesion DNA synthesis: 19F NMR, surface plasmon resonance, and primer kinetic studies. Biochemistry 53:4059–4071.

Jarrett SG, Wolf Horrell EM, Christian PA, Vanover JC, Boulanger MC, Zou Y, D'Orazio JA. 2014. PKA-mediated phosphorylation of ATR promotes recruitment of XPA to UV-induced DNA damage. Mol. Cell 54:999–1011.

Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML. 1983. Comparison of simple potential functions for simulating liquid water. J. of Chemical Physics 79:926–935.

Juricek L, Bui L-C, Busi F, Pierre S, Guyot E, Lamouri A, Dupret J-M, Barouki R, Coumoul X, Rodrigues-Lima F. 2014 Sep 17. Activation of the aryl hydrocarbon receptor by carcinogenic aromatic amines and modulatory effects of their N-acetylated metabolites. Arch. Toxicol. 1–10.

Kalam MA, Basu AK. 2005. Mutagenesis of 8-oxoguanine adjacent to an abasic site in simian kidney cells: tandem mutations and enhancement of G-->T transversions. Chem. Res. Toxicol. 18:1187–1192.

Kawasumi M, Lemos B, Bradner JE, Thibodeau R, Kim Y-S, Schmidt M, Higgins E, Koo S-W, Angle-Zahn A, Chen A, et al. 2011. Protection from UV-induced skin carcinogenesis by genetic inhibition of the ataxia telangiectasia and Rad3-related (ATR) kinase. Proc. Natl. Acad. Sci. 108:13716–13721.

Kim BM, You M-H, Chen C-H, Lee S, Hong Y, Hong Y, Kimchi A, Zhou XZ, Lee T-H. 2014. Death-associated protein kinase 1 has a critical role in aberrant tau protein regulation and function. Cell Death Dis. 5:e1237.

Kinds of mutations formed when a shuttle vector containing adducts of (+/-)-7 beta, 8 alpha-dihydroxy-9 alpha, 10 alpha-epoxy-7,8,9, 10-tetrahydrobenzo[a]pyrene replicates in human cells. 1987. Kinds of mutations formed when a shuttle vector containing adducts of (+/-)-7 beta, 8 alpha-dihydroxy-9 alpha, 10 alpha-epoxy-7,8,9, 10-tetrahydrobenzo[a]pyrene replicates in human cells. 84:3787–3791.

Kisker C, Kuper J, Van Houten B. 2013. Prokaryotic nucleotide excision repair. Cold Spring Harb. Perspect. Biol. 5:a012591.

Koehl P, Valladier P, Lefèvre J-F, Fuchs RPP. 1989. Strong structural effect of the position of a single acetylaminofluorene adduct within a mutation hot spot. Nucleic Acids Res.

17:9531–9541.

Kolodner RD, Putnam CD, Myung K. 2002. Maintenance of genome stability in Saccharomyces cerevisiae. Science 297:552–557.

Kortt AA, Gruen LC, Oddie GW. 1997. Influence of mass transfer and surface ligand heterogeneity on quantitative BIAcore binding data. Analysis of the interaction of NC10 Fab with an anti-idiotype Fab'. J. Mol. Recognit. 10:148–158.

Köberle B, Roginskaya V, Wood RD. 2006. XPA protein as a limiting factor for nucleotide excision repair and UV sensitivity in human cells. DNA Repair 5:641–648.

Krasikova YS, Rechkunova NI, Maltseva EA, Petruseva IO, Lavrik OI. 2010. Localization of xeroderma pigmentosum group A protein and replication protein A on damaged DNA in nucleotide excision repair. Nucleic Acids Res. 38:8083–8094.

Kumar A, Mazzanti M, Mistrik M, Kosar M, Beznoussenko GV, Mironov AA, Garrè M, Parazzoli D, Shivashankar GV, Scita G, et al. 2014. ATR Mediates a Checkpoint at the Nuclear Envelope in Response to Mechanical Stress. Cell 158:633–646.

Kvaratskhelia M, Miller JT, Budihas SR, Pannell LK, Le Grice SFJ. 2002. Identification of specific HIV-1 reverse transcriptase contacts to the viral RNA:tRNA complex by mass spectrometry and a primary amine selective reagent. Proc. Natl. Acad. Sci. 99:15988–15993.

Lambert MW, Yang L. 2000. Xeroderma pigmentosum complementation group A protein acts as a processivity factor. Biochem. Biophys. Res. Commun. 271:782–787.

Lankaš F, Sponer J, Langowski J, Cheatham TE III. 2003. DNA Basepair Step Deformability Inferred from Molecular Dynamics Simulations. Biophys. J. 85:2872–2883.

Lavery R, Moakher M, Maddocks JH, Petkeviciute D, Zakrzewska K. 2009. Conformational analysis of nucleic acids revisited: Curves. Nucleic Acids Res. 37:5917–5929.

Lee SH, Kim DK, Drissi R. 1995. Human xeroderma pigmentosum group A protein interacts with human replication protein A and inhibits DNA replication. J. of Biological Chemistry 270:21800–21805.

Lee TH, Chen C-H, Suizu F, Huang P, Schiene-Fischer C, Daum S, Zhang YJ, Goate A, Chen R-H, Zhou XZ, et al. 2011. Death-associated protein kinase 1 phosphorylates Pin1 and inhibits its prolyl isomerase activity and cellular function. Mol. Cell 42:147–159.

Lee Y-C, Cai Y, Mu H, Broyde S, Amin S, Chen X, Min J-H, Geacintov NE. 2014. The relationships between XPC binding to conformationally diverse DNA adducts and their excision by the human NER system: is there a correlation? DNA Repair 19:55–63.

Lehmann AR. 2003. DNA repair-deficient diseases, xeroderma pigmentosum, Cockayne syndrome and trichothiodystrophy. Biochimie 85:1101–1111.

Leonhardt H, Rahn HP, Weinzierl P, Sporbert A, Cremer T, Zink D, Cardoso MC. 2000. Dynamics of DNA replication factories in living cells. J. Cell Biol. 149:271–280.

Lessene G, Czabotar PE, Colman PM. 2008. BCL-2 family antagonists for cancer therapy. Nat. Rev. Drug Discov. 7:989–1000.

Li C-L, Golebiowski FM, Onishi Y, Samara NL, Sugasawa K, Yang W. 2015. Tripartite DNA Lesion Recognition and Verification by XPC, TFIIH, and XPA in Nucleotide Excision Repair. Mol Cell 59:1025–1034.

Li L, Zou L. 2005. Sensing, signaling, and responding to DNA damage: organization of the checkpoint pathways in mammalian cells. J. Cell. Biochem. 94:298–306.

Liang F, Cho BP. 2010. Enthalpy-entropy contribution to carcinogen-induced DNA conformational heterogeneity. Biochemistry 49:259–266.

Liang F, Meneni S, Cho BP. 2006. Induced Circular Dichroism Characteristics as Conformational Probes for Carcinogenic Aminofluorene−DNA Adducts. Chem. Res. Toxicol. 19:1040–1043.

Lieber MR, Karanjawala ZE. 2004. Ageing, repetitive genomes and DNA damage. Nat Rev Mol. Cell Biol .5:69–75.

Liou Y-C, Zhou XZ, Lu KP. 2011. Prolyl isomerase Pin1 as a molecular switch to determine the fate of phosphoproteins. Trends in Biochemical Sciences 36:501–514.

Liu B, Wang J, Chan KM, Tjia WM, Deng W, Guan X, Huang J-D, Li KM, Chau PY, Chen DJ, et al. 2005. Genomic instability in laminopathy-based premature aging. Nat. Med. 11:780–785.

Liu S, Shiotani B, Lahiri M, Maréchal A, Tse A, Leung CCY, Glover JNM, Yang XH, Zou L. 2011. ATR autophosphorylation as a molecular switch for checkpoint activation. Mol. Cell 43:192–202.

Liu Yang, Reeves D, Kropachev K, Cai Y, Ding S, Kolbanovskiy M, Kolbanovskiy A, Bolton JL, Broyde S, Van Houten B, et al. 2011. Probing for DNA damage with β-hairpins: Similarities in incision efficiencies of bulky DNA adducts by prokaryotic and human nucleotide excision repair systems in vitro. DNA Repair 10:684–696.

Liu Yiyong, Kvaratskhelia M, Hess S, Qu Y, Zou Y. 2005. Modulation of replication protein A function by its hyperphosphorylation-induced conformational change involving DNA binding domain B. J. Biol. Chem. 280:32775–32783.

Liu Yiyong, Rusinol A, Sinensky M, Wang Y, Zou Y. 2006. DNA damage responses in

progeroid syndromes arise from defective maturation of prelamin A. J. Cell Sci. 119:4644–4649.

Liu Yiyong, Wang Y, Rusinol AE, Sinensky MS, Liu J, Shell SM, Zou Y. 2008. Involvement of xeroderma pigmentosum group A (XPA) in progeria arising from defective maturation of prelamin A. FASEB J. 22:603–611.

Liu Yu, Liu Yiyong, Yang Z, Utzat C, Wang G, Basu AK, Zou Y. 2005. Cooperative interaction of human XPA stabilizes and enhances specific binding of XPA to DNA damage. Biochemistry 44:7361–7368.

Lombard DB, Chua KF, Mostoslavsky R, Franco S, Gostissa M, Alt FW. 2005. DNA repair, genome stability, and aging. Cell 120:497–512.

Lopez-Contreras AJ, Specks J, Barlow JH, Ambrogio C, Desler C, Vikingsson S, Rodrigo-Perez S, Green H, Rasmussen LJ, Murga M, et al. 2015. Increased Rrm2 gene dosage reduces fragile site breakage and prolongs survival of ATR mutant mice. Genes Dev. 29:690–695.

Lord CJ, Ashworth A. 2008. Targeted therapy for cancer using PARP inhibitors. Current Opinion in Pharmacology 8:363–369.

Lu KP, Hanes SD, Hunter T. 1996. A human peptidyl-prolyl isomerase essential for regulation of mitosis. Nature 380:544–547.

Lu PJ, Wulf G, Zhou XZ, Davies P, Lu KP. 1999. The prolyl isomerase Pin1 restores the function of Alzheimer-associated phosphorylated tau protein. Nature 399:784–788.

Lu X-J, Olson WK. 2008. 3DNA: a versatile, integrated software system for the analysis, rebuilding and visualization of three-dimensional nucleic-acid structures. Nat. Protoc. 3:1213–1227.

Lu Z, Hunter T. 2014. Prolyl isomerase Pin1 in cancer. Cell Res. 24:1033–1049.

Luch A. 2005. Nature and nurture - lessons from chemical carcinogenesis. Nat. Rev. Cancer 5:113–125.

Luijsterburg MS, Bornstaedt von G, Gourdin AM, Politi AZ, Moné MJ, Warmerdam DO, Goedhart J, Vermeulen W, van Driel R, Höfer T. 2010. Stochastic and reversible assembly of a multiprotein DNA repair complex ensures accurate target site recognition and efficient repair. J. Cell Biol. 189:445–463.

Luo C, Krishnasamy R, Basu AK, Zou Y. 2000. Recognition and incision of site-specifically modified C8 guanine adducts formed by 2-aminofluorene, N-acetyl-2-aminofluorene and 1-nitropyrene by UvrABC nuclease. Nucleic Acids Res. 28:3719–3724.

Luo X, Budihardjo I, Zou H, Slaughter C, Wang X. 1998. Bid, a Bcl2 interacting protein,

mediates cytochrome c release from mitochondria in response to activation of cell surface death receptors. Cell 94:481–490.

Lyubchenko YL, Shlyakhtenko LS, Appella E, Harrington RE. 1993. CA runs increase DNA flexibility in the complex of .lambda. Cro protein with the OR3 site. Biochemistry 32:4121–4127.

Ma J, Wang S, Zhao F, Xu J. 2013. Protein threading using context-specific alignment potential. Bioinformatics 29:i257–65.

Magnander K, Elmroth K. 2012. Biological consequences of formation and repair of complex DNA damage. Cancer Lett. 327:90–96.

Mah MC, Boldt J, Culp SJ, Maher VM, McCormick JJ. 1991. Replication of acetylaminofluorene-adducted plasmids in human cells: spectrum of base substitutions and evidence of excision repair. Proc. Natl. Acad. Sci. 88:10193–10197.

Malta E, Moolenaar GF, Goosen N. 2006. Base flipping in nucleotide excision repair. J. of Biological Chemistry 281:2184–2194.

Malta E, Verhagen CP, Moolenaar GF, Filippov DV, van der Marel GA, Goosen N. 2008. Functions of base flipping in E. coli nucleotide excision repair. DNA Repair 7:1647–1658.

Manju K, Muralikrishna B, Parnaik VK. 2006. Expression of disease-causing lamin A mutants impairs the formation of DNA repair foci. J. Cell Sci. 119:2704–2714.

Mao B, Hingerty BE, Broyde S, Patel DJ. 1998. Solution Structure of the Aminofluorene [AF]-Intercalated Conformer of the syn-[AF]-C 8-dG Adduct Opposite dC in a DNA Duplex †. Biochemistry 37:81–94.

Marteijn JA, Lans H, Vermeulen W, Hoeijmakers JHJ. 2014. Understanding nucleotide excision repair and its roles in cancer and ageing. Nat Rev Mol. Cell Biol. 15:465–481.

Matsuoka S, Ballif BA, Smogorzewska A, McDonald ER, Hurov KE, Luo J, Bakalarski CE, Zhao Z, Solimini N, Lerenthal Y, et al. 2007. ATM and ATR substrate analysis reveals extensive protein networks responsive to DNA damage. Science 316:1160–1166.

Mattout A, Goldberg M, Tzur Y, Margalit A, Gruenbaum Y. 2007. Specific and conserved sequences in D. melanogaster and C. elegans lamins and histone H2A mediate the attachment of lamins to chromosomes. J. Cell Sci. 120:77–85.

McAlear MA, Tuffo KM, Holm C. 1996. The large subunit of replication factor C (Rfc1p/Cdc44p) is required for DNA replication and DNA repair in Saccharomyces cerevisiae. Genetics 142:65–78.

McConnell KJ, Beveridge DL. 2001. Molecular dynamics simulations of B′-DNA: sequence effects on A-tract-induced bending and flexibility. J. Mol. Biol. 314:23–40.

McDonnell JM, Fushman D, Milliman CL, Korsmeyer SJ, Cowburn D. 1999. Solution structure of the proapoptotic molecule BID: a structural basis for apoptotic agonists and antagonists. Cell 96:625–634.

Meierhans D, Sieber M, Allemann RK. 1997. High affinity binding of MEF-2C correlates with DNA bending. Nucleic Acids Res. 25:4537–4544.

Mekhovich O, Tang MS, Romano LJ. 1998. Rate of incision of N-acetyl-2-aminofluorene and N-2-aminofluorene adducts by UvrABC nuclease is adduct- and sequence-specific: comparison of the rates of UvrABC nuclease incision and protein-DNA complex formation. Biochemistry 37:571–579.

Melchior WB, Marques MM, Beland FA. 1994. Mutations induced by aromatic amine DNA adducts in pBR322. Carcinogenesis 15:889–899.

Meneni S, Liang F, Cho BP. 2007. Examination of the long-range effects of aminofluorene-induced conformational heterogeneity and its relevance to the mechanism of translesional DNA synthesis. J. Mol. Biol. 366:1387–1400.

Meneni S, Shell SM, Zou Y, Cho BP. 2007. Conformation-specific recognition of carcinogen-DNA adduct in escherichia coli nucleotide excision repair. Chem. Res. Toxicol. 20:6–10.

Meneni SR. 2006. Sequence effects of aminofluorene-modified DNA duplexes: thermodynamic and circular dichroism properties. Nucleic Acids Res. 34:755–763.

Meneni SR, Shell SM, Gao L, Jurecka P, Lee W, Sponer J, Zou Y, Chiarelli MP, Cho BP. 2007. Spectroscopic and theoretical insights into sequence effects of aminofluorene-induced conformational heterogeneity and nucleotide excision repair. Biochemistry 46:11263–11278.

Méndez J, Stillman B. 2000. Chromatin association of human origin recognition complex, cdc6, and minichromosome maintenance proteins during the cell cycle: assembly of prereplication complexes in late mitosis. Mol. Cell Biol. 20:8602–8612.

Min J-H, Pavletich NP. 2007. Recognition of DNA damage by the Rad4 nucleotide excision repair protein. Nature 449:570–575.

Mirkin EV, Mirkin SM. 2007. Replication fork stalling at natural impediments. Microbiol. Mol. Biol. Rev. 71:13–35.

Missura M, Buterin T, Hindges R, Hübscher U, Kaspárková J, Brabec V, Naegeli H. 2001. Double-check probing of DNA bending and unwinding by XPA-RPA: an architectural function in DNA repair. EMBO J. 20:3554–3564.

Mocquet V, Kropachev K, Kolbanovskiy M, Kolbanovskiy A, Tapias A, Cai Y, Broyde S, Geacintov NE, Egly JM. 2007. The human DNA repair factor XPC-HR23B distinguishes stereoisomeric benzo[a]pyrenyl-DNA lesions. EMBO J. 26:2923–2932.

Moir RD, Spann TP, Goldman RD. 1995. The dynamic properties and possible functions of nuclear lamins. Int. Rev. Cytol. 162B:141–182.

Moir RD, Spann TP, Herrmann H, Goldman RD. 2000. Disruption of nuclear lamin organization blocks the elongation phase of DNA replication. J. Cell Biol. 149:1179–1192.

Morafraile EC, Diffley JFX, Tercero JA, Segurado M. 2015. Checkpoint-dependent RNR induction promotes fork restart after replicative stress. Sci. Rep. 5:7886.

Moriel-Carretero M, Aguilera A. 2010. A postincision-deficient TFIIH causes replication fork breakage and uncovers alternative Rad51- or Pol32-mediated restart mechanisms. Mol. Cell 37:690–701.

Mu D, Bertrand-Burggraf E, Huang JC, Fuchs RP, Sancar A, Fuchs BP. 1994. Human and E.coli excinucleases are affected differently by the sequence context of acetylaminofluorene-guanine adduct. Nucleic Acids Res. 22:4869–4871.

Mu H, Geacintov NE, Zhang Y, Broyde S. 2015 Aug 13. Recognition of damaged DNA for nucleotide excision repair: a correlated motion mechanism with a mismatched cis-syn thymine dimer lesion. Biochemistry.

Mu H, Kropachev K, Wang L, Zhang L, Kolbanovskiy A, Kolbanovskiy M, Geacintov NE, Broyde S. 2012. Nucleotide excision repair of 2-acetylaminofluorene- and 2-aminofluorene-(C8)-guanine adducts: molecular dynamics simulations elucidate how lesion structure and base sequence context impact repair efficiencies. Nucleic Acids Res. 40:9675–9690.

Musich PR, Zou Y. 2009. Genomic instability and DNA damage responses in progeria arising from defective maturation of prelamin A. Aging 1:28–37.

Musich PR, Zou Y. 2011. DNA-damage accumulation and replicative arrest in Hutchinson-Gilford progeria syndrome. Biochem. Soc. Trans. 39:1764–1769.

Myers K, Gagou ME, Zuazua-Villar P, Rodriguez R, Meuth M. 2009. ATR and Chk1 suppress a caspase-3-dependent apoptotic response following DNA replication stress. PLoS Genet. 5:e1000324.

Nakamura K, Greenwood A, Binder L, Bigio EH, Denial S, Nicholson L, Zhou XZ, Lu KP. 2012. Proline isomer-specific antibodies reveal the early pathogenic tau conformation in Alzheimer's disease. Cell 149:232–244.

Nam EA, Zhao R, Glick GG, Bansbach CE, Friedman DB, Cortez D. 2011. Thr-1989

phosphorylation is a marker of active ataxia telangiectasia-mutated and Rad3-related (ATR) kinase. J. Biol. Chem. 286:28707–28714.

Nasir M, Ahmad N, Sieber CMK, Latif A, Malik SA, Hameed A. 2013. In silico characterization of a novel pathogenic deletion mutation identified in XPA gene in a Pakistani family with severe xeroderma pigmentosum. J. Biomed. Sci. 20:70.

Navarro CL, Cadiñanos J, De Sandre-Giovannoli A, Bernard R, Courrier S, Boccaccio I, Boyer A, Kleijer WJ, Wagner A, Giuliano F, et al. 2005. Loss of ZMPSTE24 (FACE-1) causes autosomal recessive restrictive dermopathy and accumulation of Lamin A precursors. Human Molecular Genetics 14:1503–1513.

Neumann H-G. 2007. Aromatic amines in experimental cancer research: tissue-specific effects, an old problem and new solutions. Crit. Rev. Toxicol. 37:211–236.

Niedernhofer LJ, Garinis GA, Raams A, Lalai AS, Robinson AR, Appeldoorn E, Odijk H, Oostendorp R, Ahmad A, Van Leeuwen W, et al. 2006. A new progeroid syndrome reveals that genotoxic stress suppresses the somatotroph axis. Nature 444:1038–1043.

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

Nouspikel T. 2009. DNA repair in mammalian cells : Nucleotide excision repair: variations on versatility. Cell. Mol. Life Sci. 66:994–1009.

O'Driscoll M. 2009. Mouse models for ATR deficiency. DNA Repair 8:1333–1337.

O'Handley SF, Sanford DG, Xu R, Lester CC, Hingerty BE, Broyde S, Krugh TR. 2002. Structural characterization of an N-acetyl-2-aminofluorene (AAF) modified DNA oligomer by NMR, energy minimization, and molecular dynamics. Biochemistry 32:2481–2497.

Ogi T, Limsirichaikul S, Overmeer RM, Volker M, Takenaka K, Cloney R, Nakazawa Y, Niimi A, Miki Y, Jaspers NG, et al. 2010. Three DNA polymerases, recruited by different mechanisms, carry out NER repair synthesis in human cells. Mol. Cell 37:714–727.

Oh EY, Grossman L. 1986. The effect of Escherichia coli Uvr protein binding on the topology of supercoiled DNA. Nucleic Acids Res. 14:8557–8571.

Olive M, Harten I, Mitchell R, Beers JK, Djabali K, Cao K, Erdos MR, Blair C, Funke B, Smoot L, et al. 2010. Cardiovascular pathology in Hutchinson-Gilford progeria: correlation with the vascular pathology of aging. Arterioscler. Thromb. Vasc. Biol. 30:2301–2309.

Orren DK, Sancar A. 1990. Formation and enzymatic properties of the UvrB.DNA complex. J.

of Biological Chemistry 265:15796–15803.

Pabla N, Huang S, Mi Q-S, Daniel R, Dong Z. 2008. ATR-Chk2 signaling in p53 activation and DNA damage response during cisplatin-induced apoptosis. J. of Biological Chemistry 283:6572–6583.

Pace CN, Scholtz JM. 1998. A helix propensity scale based on experimental studies of peptides and proteins. Biophysj. 75:422–427.

Pacheco LM, Gomez LA, Dias J, Ziebarth NM, Howard GA, Schiller PC. 2014. Progerin expression disrupts critical adult stem cell functions involved in tissue repair. Aging 6:1049–1063.

Pakotiprapha D, Inuzuka Y, Bowman BR, Moolenaar GF, Goosen N, Jeruzalmi D, Verdine GL. 2008. Crystal structure of Bacillus stearothermophilus UvrA provides insight into ATP-modulated dimerization, UvrB interaction, and DNA binding. Mol. Cell 29:122–133.

Pakotiprapha D, Samuels M, Shen K, Hu JH, Jeruzalmi D. 2012. Structure and mechanism of the UvrA-UvrB DNA damage sensor. Nature 19:291–298.

Papamichos-Chronakis M, Peterson CL. 2013. Chromatin and the genome integrity network. Nat. Rev. Genet. 14:62–75.

Park C-J, Choi B-S. 2006. The protein shuffle. Sequential interactions among components of the human nucleotide excision repair pathway. FEBS J. 273:1600–1608.

Park CH, Sancar A. 1994. Formation of a ternary complex by human XPA, ERCC1, and ERCC4(XPF) excision repair proteins. Proc. Natl. Acad. Sci. 91:5017–5021.

Parrinello M, Rahman A. 1980. Crystal Structure and Pair Potentials: A Molecular-Dynamics Study. Physical Review Letters 45:1196–1199.

Patnaik S, Cho BP. 2010. Structures of 2-acetylaminofluorene modified DNA revisited: insight into conformational heterogeneity. Chem. Res. Toxicol. 23:1650–1652.

Pearl LH, Schierz AC, Ward SE, Al-Lazikani B, Pearl FMG. 2015. Therapeutic opportunities within the DNA damage response. Nat. Rev. Cancer 15:166–180.

Peng J, Xu J. 2011a. RaptorX: exploiting structure information for protein alignment by statistical inference. Proteins 79 Suppl 10:161–171.

Peng J, Xu J. 2011b. A multiple-template approach to protein threading. Proteins 79:1930–1939.

Petermann E, Maya-Mendoza A, Zachos G, Gillespie DAF, Jackson DA, Caldecott KW. 2006. Chk1 requirement for high global rates of replication fork progression during normal vertebrate S phase. Mol. Cell Biol. 26:3319–3326.

Petermann E, Woodcock M, Helleday T. 2010. Chk1 promotes replication fork progression by controlling replication initiation. Proc. Natl. Acad. Sci. 107:16090–16095.

Petit C, Sancar A. 1999. Nucleotide excision repair: from E. coli to man. Biochimie 81:15–25.

Poirier MC. 2012. Chemical-induced DNA damage and human cancer risk. Discov. Med 14:283–288.

Postel-Vinay S, Vanhecke E, Olaussen KA, Lord CJ, Ashworth A, Soria J-C. 2012. The potential of exploiting DNA-repair defects for optimizing lung cancer treatment. Nat. Rev. Clin. Oncol. 9:144–155.

Puumalainen M-R, Lessel D, Rüthemann P, Kaczmarek N, Bachmann K, Ramadan K, Naegeli H. 2014. Chromatin retention of DNA damage sensors DDB2 and XPC through loss of p97 segregase causes genotoxicity. Nat. Commun. 5:3695.

Radomski JL. 1979. The primary aromatic amines: their biological properties and structure-activity relationships. Annu. Rev. Pharmacol. Toxicol. 19:129–157.

Ragnauth CD, Warren DT, Liu Y, McNair R, Tajsic T, Figg N, Shroff R, Skepper J, Shanahan CM. 2010. Prelamin A acts to accelerate smooth muscle cell senescence and is a novel biomarker of human vascular aging. Circulation 121:2200–2210.

Ray A, Milum K, Battu A, Wani G, Wani AA. 2013. NER initiation factors, DDB2 and XPC, regulate UV radiation response by recruiting ATR and ATM kinases to DNA damage sites. DNA Repair 12:273–283.

Reardon JT, Mu D, Sancar A. 1996. Overproduction, purification, and characterization of the XPC subunit of the human DNA repair excision nuclease. J. of Biological Chemistry 271:19451–19456.

Reardon JT, Sancar A. 2005. Nucleotide excision repair. Prog. Nucleic Acid Res. Mol. Biol. 79:183–235.

Rees JR, Stukel TA, Perry AE, Zens MS, Spencer SK, Karagas MR. 2007. Tea consumption and basal cell and squamous cell skin cancer: results of a case-control study. J. Am. Acad. Dermatol. 56:781–785.

Rhéaume E, Cohen LY, Uhlmann F, Lazure C, Alam A, Hurwitz J, Sékaly RP, Denis F. 1997. The large subunit of replication factor C is a substrate for caspase-3 in vitro and is cleaved by a caspase-3-like protease during Fas-mediated apoptosis. EMBO J. 16:6346–6354.

Ricke RM, Bielinsky A-K. 2005. Easy detection of chromatin binding proteins by the Histone Association Assay. Biol. Proced Online 7:60–69.

Riedl T, Hanaoka F, Egly JM. 2003. The comings and goings of nucleotide excision repair

factors on damaged DNA. EMBO J. 22:5293–5303.

Ross MK, Said B, Shank RC. 2000. DNA-damaging effects of genotoxins in mixture: modulation of covalent binding to DNA. Toxicol. Sci. 53:224–236.

Rouse J, Jackson SP. 2002. Interfaces between the detection, signaling, and repair of DNA damage. Science 297:547–551.

Roy A, Fuchs RPP. 1994. Mutational spectrum induced in Saccharomyces cerevisiae by the carcinogen N-2-acetylaminofluorene. Mol. Gen. Genet. 245:69–77.

Ruan Q, Liu T, Kolbanovskiy A, Liu Y, Ren J, Skorvaga M, Zou Y, Lader J, Malkani B, Amin S, et al. 2007. Sequence context- and temperature-dependent nucleotide excision repair of a benzo[a]pyrene diol epoxide-guanine DNA adduct catalyzed by thermophilic UvrABC proteins. Biochemistry 46:7006–7015.

Sancar A, Hearst JE. 1993. Molecular matchmakers. Science 259:1415–1420.

Sancar A, Lindsey-Boltz LA, Unsal-Kaçmaz K, Linn S. 2004. Molecular mechanisms of mammalian DNA repair and the DNA damage checkpoints. Annu. Rev. Biochem. 73:39–85.

Sarkaria JN, Busby EC, Tibbetts RS, Roos P, Taya Y, Karnitz LM, Abraham RT. 1999. Inhibition of ATM and ATR kinase activities by the radiosensitizing agent, caffeine. Cancer Res. 59:4375–4382.

Scaffidi P, Misteli T. 2006. Lamin A-dependent nuclear defects in human aging. Science 312:1059–1063.

Scharer OD. 2013. Nucleotide excision repair in eukaryotes. Cold Spring Harb. Perspect. Biol. 5:a012609–a012609.

Schoppy DW, Ragland RL, Gilad O, Shastri N, Peters AA, Murga M, Fernandez-Capetillo O, Diehl JA, Brown EJ. 2012. Oncogenic stress sensitizes murine cancers to hypomorphic suppression of ATR. J. Clin. Invest. 122:241–252.

Schumacher B, Garinis GA, Hoeijmakers JHJ. 2008. Age to survive: DNA damage and aging. Trends Genet. 24:77–85.

Sedelnikova OA, Horikawa I, Zimonjic DB, Popescu NC, Bonner WM, Barrett JC. 2004. Senescing human cells and ageing mice accumulate DNA lesions with unrepairable double-strand breaks. Nat. Cell Biol. 6:168–170.

Seeberg E, Fuchs RP. 1990. Acetylaminofluorene bound to different guanines of the sequence -GGCGCC- is excised with different efficiencies by the UvrABC excision nuclease in a pattern not correlated to the potency of mutation induction. Proc. Natl. Acad Sci. 87:191–194.

Shechter D, Costanzo V, Gautier J. 2004. Regulation of DNA replication by ATR: signaling in response to DNA intermediates. DNA Repair 3:901–908.

Shell SM, Hawkins EK, Tsai M-S, Hlaing AS, Rizzo CJ, Chazin WJ. 2013 Sep 14. Xeroderma pigmentosum complementation group C protein (XPC) serves as a general sensor of damaged DNA. DNA Repair 1–7.

Shell SM, Hess S, Kvaratskhelia M, Zou Y. 2005. Mass spectrometric identification of lysines involved in the interaction of human replication protein a with single-stranded DNA. Biochemistry 44:971–978.

Shell SM, Li Z, Shkriabai N, Kvaratskhelia M, Brosey C, Serrano MA, Chazin WJ, Musich PR, Zou Y. 2009. Checkpoint kinase ATR promotes nucleotide excision repair of UV-induced DNA damage via physical interaction with xeroderma pigmentosum group A. J. Biol. Chem. 284:24213–24222.

Shell SM, Zou Y. 2008. Other proteins interacting with XP proteins. Adv. Exp. Med. Biol. 637:103–112.

Shen Z-J, Esnault S, Schinzel A, Borner C, Malter JS. 2009. The peptidyl-prolyl isomerase Pin1 facilitates cytokine-induced survival of eosinophils by suppressing Bax activation. Nat. Immunol. 10:257–265.

Shi Q, Thresher R, Sancar A, Griffith J. 1992. Electron microscopic study of (A)BC excinuclease. J. Mol. Biol. 226:425–432.

Shikazono N, Pearson C, O'Neill P, Thacker J. 2006. The roles of specific glycosylases in determining the mutagenic consequences of clustered DNA base damage. Nucleic Acids Res. 34:3722–3730.

Shiloh Y. 2003. ATM and related protein kinases: safeguarding genome integrity. Nat. Rev. Cancer 3:155–168.

Shivji MK, Moggs JG, Kuraoka I, Wood RD. 1999. Dual-incision assays for nucleotide excision repair using DNA with a lesion at a specific site. Methods Mol. Biol. 113:373–392.

Shkriabai N, Patil SS, Hess S, Budihas SR, Craigie R, Burke TR, Le Grice SFJ, Kvaratskhelia M. 2004. Identification of an inhibitor-binding site to HIV-1 integrase with affinity acetylation and mass spectrometry. Proc. Natl. Acad. Sci. 101:6894–6899.

Shumaker DK, Solimando L, Sengupta K, Shimi T, Adam SA, Grunwald A, Strelkov SV, Aebi U, Cardoso MC, Goldman RD. 2008. The highly conserved nuclear lamin Ig-fold binds to PCNA: its role in DNA replication. J. Cell Biol. 181:269–280.

Sibson R. 1973. SLINK: An optimally efficient algorithm for the single-link cluster method. Computer J. 16:30–34.

Sorrentino G, Comel A, Mantovani F, Del Sal G. 2014. Regulation of mitochondrial apoptosis by Pin1 in cancer and neurodegeneration. Mitochondrion 19 Pt A:88–96.

Spann TP, Moir RD, Goldman AE, Stick R, Goldman RD. 1997. Disruption of nuclear lamin organization alters the distribution of replication factors and inhibits DNA synthesis. J. Cell Biol. 136:1201–1212.

Steger M, Murina O, Hühn D, Ferretti LP, Walser R, Hänggi K, Lafranchi L, Neugebauer C, Paliwal S, Janscak P, et al. 2013. Prolyl isomerase PIN1 regulates DNA double-strand break repair by counteracting DNA end resection. Mol. Cell 50:333–343.

Sugasawa K, Akagi J-I, Nishi R, Iwai S, Hanaoka F. 2009. Two-step recognition of DNA damage for mammalian nucleotide excision repair: Directional binding of the XPC complex and DNA strand scanning. Mol. Cell 36:642–653.

Sugasawa K, Ng JM, Masutani C, Iwai S, van der Spek PJ, Eker AP, Hanaoka F, Bootsma D, Hoeijmakers JH. 1998. Xeroderma pigmentosum group C protein complex is the initiator of global genome nucleotide excision repair. Mol. Cell 2:223–232.

Sugasawa K, Okamoto T, Shimizu Y, Masutani C, Iwai S, Hanaoka F. 2001. A multistep damage recognition mechanism for global genomic nucleotide excision repair. Genes Dev. 15:507–521.

Sugitani N, Shell SM, Soss SE, Chazin WJ. 2014 Jul 24. Redefining the DNA-Binding Domain of Human XPA. J. Am. Chem. Soc.:140724070100004.

Suh M, Ariese F, Small GJ, Jankowiak R, Liu TM, Geacintov NE. 1995. Conformational studies of the (+)-trans, (-)-trans, (+)-cis, and (-)-cis adducts of anti-benzo[a]pyrene diolepoxide to N2-dG in duplex oligonucleotides using polyacrylamide gel electrophoresis and low-temperature fluorescence spectroscopy. Biophys. Chem. 56:281–296.

Tait SWG, Green DR. 2010. Mitochondria and cell death: outer membrane permeabilization and beyond. Nat Rev Mol. Cell Biol. 11:621–632.

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

Tang H, Hilton B, Musich PR, Fang DZ, Zou Y. 2012. Replication factor C1, the large subunit of replication factor C, is proteolytically truncated in Hutchinson–Gilford progeria syndrome. Aging Cell 11:363–365.

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. J. Biol. Chem. 279:19074–19083.

The PyMOL Molecular Graphics System. 2010. *The PyMOL Molecular Graphics System*. Schrodinger, LLC.

Thoma BS, Vasquez KM. 2003. Critical DNA damage recognition functions of XPC-hHR23B and XPA-RPA in nucleotide excision repair. Mol. Carcinog. 38:1–13.

Toko H, Konstandin MH, Doroudgar S, Ormachea L, Joyo E, Joyo AY, Din S, Gude NA, Collins B, Völkers M, et al. 2013. Regulation of cardiac hypertrophic signaling by prolyl isomerase Pin1. Circulation Research 112:1244–1252.

Trego KS, Turchi JJ. 2006. Pre-steady-state binding of damaged DNA by XPC-hHR23B reveals a kinetic mechanism for damage discrimination. Biochemistry 45:1961–1969.

Truglio JJ, Croteau DL, Van Houten B, Kisker C. 2006. Prokaryotic nucleotide excision repair: the UvrABC system. Chem. Rev. 106:233–252.

Tsao H, Mao B, Zhuang P, Xu R, Amin S, Geacintov NE. 1998. Sequence dependence and characteristics of bends induced by site-specific polynuclear aromatic carcinogen-deoxyguanosine lesions in oligonucleotides. Biochemistry 37:4993–5000.

Tsodikov OV, Ivanov D, Orelli B, Staresincic L, Shoshani I, Oberman R, Scharer OD, Wagner G, Ellenberger T. 2007. Structural basis for the recruitment of ERCC1-XPF to nucleotide excision repair complexes by XPA. EMBO J. 26:4768–4776.

Ukekawa R, Miki K, Fujii M, Hirano H, Ayusawa D. 2007. Accumulation of multiple forms of lamin A with down-regulation of FACE-1 suppresses growth in senescent human cells. Genes Cells 12:397–406.

Vaidyanathan VG, Xu L, Cho BP. 2012. Binary and ternary binding affinities between exonuclease-deficient Klenow fragment (Kf-exo(-)) and various arylamine DNA lesions characterized by surface plasmon resonance. Chem. Res. Toxicol. 25:1568–1570.

Vaidyanathan VG, Xu L, Cho BP. 2013 May 22. Binding kinetics of DNA-protein interaction using surface plasmon resonance. Nature.

Van Houten B. 1990. Nucleotide excision repair in Escherichia coli. Microbiol. Rev. 54:18–51.

Van Houten B, Croteau DL, DellaVecchia MJ, Wang H, Kisker C. 2005. “Close-fitting sleeves”: DNA damage recognition by the UvrABC nuclease system. Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis 577:92–117.

Van Houten B, Kad N. 2014 Jan 25. DNA Repair 1–8.

Van Houten B, Snowden A. 1993. Mechanism of action of theEscherichia coli UvrABC

nuclease: Clues to the damage recognition problem. BioEssays 15:51–59.

van Steeg H, Mullenders LH, Vijg J. 2000. Mutagenesis and carcinogenesis in nucleotide excision repair-deficient XPA knock out mice. Mutat. Res. 450:167–180.

Varela I, Cadiñanos J, Pendas AM, Gutiérrez-Fernández A, Folgueras AR, Sánchez LM, Zhou Z, Rodríguez FJ, Stewart CL, Vega JA, et al. 2005. Accelerated ageing in mice deficient in Zmpste24 protease is linked to p53 signalling activation. Nature 437:564–568.

Varela I, Pereira S, Ugalde AP, Navarro CL, Suárez MF, Cau P, Cadiñanos J, Osorio FG, Foray N, Cobo J, et al. 2008. Combined treatment with statins and aminobisphosphonates extends longevity in a mouse model of human premature aging. Nat. Med. 14:767–772.

Volker M, Moné MJ, Karmakar P, van Hoffen A, Schul W, Vermeulen W, Hoeijmakers JH, van Driel R, van Zeeland AA, Mullenders LH. 2001. Sequential assembly of the nucleotide excision repair factors in vivo. Mol. Cell 8:213–224.

Walkup GK, You Z, Ross PL, Allen EKH, Daryaee F, Hale MR, O'Donnell J, Ehmann DE, Schuck VJA, Buurman ET, et al. 2015. Translating slow-binding inhibition kinetics into cellular and in vivo effects. Nat. Chem. Biol. 11:416–423.

Wang X, Ira G, Tercero JA, Holmes AM, Diffley JFX, Haber JE. 2004. Role of DNA replication proteins in double-strand break-induced recombination in Saccharomyces cerevisiae. Mol. Cell Biol. 24:6891–6899.

Ward JF. 1988. DNA Damage Produced by Ionizing Radiation in Mammalian Cells: Identities, Mechanisms of Formation, and Reparability. Vol. 35. Elsevier. (Progress in Nucleic Acid Research and Molecular Biology). pp. 95–125; 31 p.

Weber AM, Ryan AJ. 2014 Dec 13. ATM and ATR as therapeutic targets in cancer. Pharmacol. Ther.

Westphal D, Kluck RM, Dewson G. 2014. Building blocks of the apoptotic pore: how Bax and Bak are activated and oligomerize during apoptosis. Cell Death Differ. 21:196–205.

Wood RD. 1999. DNA damage recognition during nucleotide excision repair in mammalian cells. Biochimie 81:39–44.

Wu X, Shell SM, Yang Z, Zou Y. 2006. Phosphorylation of nucleotide excision repair factor xeroderma pigmentosum group A by ataxia telangiectasia mutated and Rad3-related-dependent checkpoint pathway promotes cell survival in response to UV irradiation. Cancer Res. 66:2997–3005.

Xu L, Vaidyanathan VG, Cho BP. 2014. Real-time surface plasmon resonance study of biomolecular interactions between polymerase and bulky mutagenic DNA lesions. Chem. Res. Toxicol. 27:1796–1807.

Yakovchuk P, Protozanova E, Frank-Kamenetskii MD. 2006. Base-stacking and base-pairing contributions into thermal stability of the DNA double helix. Nucleic Acids Res. 34:564–574.

Yang J, Yan R, Roy A, Xu D, Poisson J, Zhang Y. 2015. The I-TASSER Suite: protein structure and function prediction. Nat. Methods 12:7–8.

Yang Z, Colis LC, Basu AK, Zou Y. 2005. Recognition and incision of gamma-radiation-induced cross-linked guanine-thymine tandem lesion G[8,5-Me]T by UvrABC nuclease. Chem. Res. Toxicol. 18:1339–1346.

Yang Z, Roginskaya M, Colis LC, Basu AK, Shell SM, Liu Y, Musich PR, Harris CM, Harris TM, Zou Y. 2006. Specific and efficient binding of xeroderma pigmentosum complementation group A to double-strand/single-strand DNA junctions with 3"-and/or 5-"ssDNA branches. Biochemistry 45:15921–15930.

Yang Z-G, Liu Y, Mao LY, Zhang J-T, Zou Y. 2002. Dimerization of human XPA and formation of XPA2-RPA protein complex. Biochemistry 41:13012–13020.

Yeo J-E, Khoo A, Fagbemi AF, Scharer OD. 2012. The efficiencies of damage recognition and excision correlate with duplex destabilization induced by acetylaminofluorene adducts in human nucleotide excision repair. Chem. Res. Toxicol. 25:2462–2468.

Yeung AT, Mattes WB, Grossman L. 1986. Protein complexes formed during the incision reaction catalyzed by the Escherichia coli UvrABC endonuclease. Nucleic Acids Res. 14:2567–2582.

You J-S, Wang M, Lee S-H. 2003. Biochemical analysis of the damage recognition process in nucleotide excision repair. J. of Biological Chemistry 278:7476–7485.

Zacchi P, Gostissa M, Uchida T, Salvagno C, Avolio F, Volinia S, Ronai Z, Blandino G, Schneider C, Del Sal G. 2002. The prolyl isomerase Pin1 reveals a mechanism to control p53 functions after genotoxic insults. Nature 419:853–857.

Zappacosta F, Pessi A, Bianchi E, Venturini S, Sollazzo M, Tramontano A, Marino G, Pucci P. 1996. Probing the tertiary structure of proteins by limited proteolysis and mass spectrometry: the case of Minibody. Protein Sci. 5:802–813.

Zeman MK, Cimprich KA. 2014. Causes and consequences of replication stress. Nat. Cell Biol. 16:2–9.

Zhang L-K, Gross ML. 2000. Matrix-assisted laser desorption/ionization mass spectrometry methods for oligodeoxynucleotides: Improvements in matrix, detection limits, quantification, and sequencing. J. of the American Society for Mass Spectrometry 11:854–865.

Zheng H, You H, Zhou XZ, Murray SA, Uchida T, Wulf G, Gu L, Tang X, Lu KP, Xiao Z-XJ. 2002.

The prolyl isomerase Pin1 is a regulator of p53 in genotoxic response. Nature 419:849–853.

Zhou BB, Elledge SJ. 2000. The DNA damage response: putting checkpoints in perspective. Nature 408:433–439.

Zhou L, Rajabzadeh M, Traficante DD, Cho BP. 1997. Conformational Heterogeneity of Arylamine-Modified DNA: 19F NMR Evidence. J. Am. Chem. Soc. 119:5384–5389.

Zinkel SS, Hurov KE, Ong C, Abtahi FM, Gross A, Korsmeyer SJ. 2005. A role for proapoptotic BID in the DNA-damage response. Cell 122:579–591.

Zou L, Liu D, Elledge SJ. 2003. Replication protein A-mediated recruitment and activation of Rad17 complexes. Proc. Natl. Acad. Sci. 100:13827–13832.

Zou Y, Bassett H, Walker R, Bishop A, Amin S, Geacintov NE, Van Houten B. 1998. Hydrophobic forces dominate the thermodynamic characteristics of UvrA-DNA damage interactions. J. Mol. Biol. 281:107–119.

Zou Y, Liu TM, Geacintov NE, Van Houten B. 1995. Interaction of the UvrABC nuclease system with a DNA duplex containing a single stereoisomer of dG-(+)- or dG-(-)-anti-BPDE. Biochemistry 34:13582–13593.

Zou Y, Liu Y, Wu X, Shell SM. 2006. Functions of human replication protein A (RPA): from DNA replication to DNA damage and stress responses. J. Cell. Physiol. 208:267–273.

Zou Y, Luo C, Geacintov NE. 2001. Hierarchy of DNA Damage Recognition in EscherichiacoliNucleotide Excision Repair †. Biochemistry 40:2923–2931.

Zou Y, Ma H, Minko IG, Shell SM, Yang Z, Qu Y, Xu Y, Geacintov NE, Lloyd RS. 2004. DNA damage recognition of mutated forms of UvrB proteins in nucleotide excision repair. Biochemistry 43:4196–4205.

Zou Y, Shell SM, Utzat CD, Luo C, Yang Z, Geacintov NE, Basu AK. 2003. Effects of DNA adduct structure and sequence context on strand opening of repair intermediates and incision by UvrABC nuclease. Biochemistry 42:12654–12661.

Zou Y, Van Houten B. 1999. Strand opening by the UvrA(2)B complex allows dynamic recognition of DNA damage. EMBO J. 18:4889–4901.

Zou Y, Van Houten B, Farrell N. 1994. Sequence specificity of DNA-DNA interstrand cross-link formation by cisplatin and dinuclear platinum complexes. Biochemistry 33:5404–5410.

VITA

BENJAMIN A. HILTON

Education: Dobyns-Bennett High School, Kingsport, TN
B.S., Biological Sciences, Clemson University, Clemson, SC
Ph.D., Biomedical Sciences, East Tennessee State University, Johnson City, TN

Professional Experience: Research Associate, East Tennessee State University, Johnson City, TN

Publications:

Benjamin A. Hilton*, Zhengke Li*, Phillip R. Musich, Hui Wang, Brian M. Cartwright, Moises Serrano, Xiao Zhen Zhou, Kun Ping Lu, and Yue Zou. ATR Plays a Direct Antiapoptotic Role at Mitochondria, which Is Regulated by Prolyl Isomerase Pin1. *Molecular Cell* (2015). 60(1): 35-46.

Wang, Y., **Hilton, B.A.**, Cui, K., and Zhu, M.-Y. Effects of Antidepressants on DSP4/CPT-Induced DNA Damage Response in Neuroblastoma SH-SY5Y Cells. *Neurotoxicity Research* (2015). 28(2): 154-170.

Hilton, B., Shkriabai, N., Musich, P. R., Kvaratskhelia, M., Shell, S., & Zou, Y. A new structural insight into XPA-DNA interactions. *Bioscience Reports* (2014). 34: art:e00162.

Jain, V.*, **Hilton, B.***, Lin, B.R.*, Jain, A., Mackerell, A.D., Zou, Y., and Cho, B.P. Structural and thermodynamic insight into *E. coli* UvrABC mediated incision of cluster di-acetylaminofluorene adducts on the *Nar*I sequence. *Chemical Research in Toxicology* (2013). 26(8): 1251–1262.

Jain, V.*, **Hilton, B.***, Lin, B.*, Patnaik, S., Liang, F., Darian, E., Zou, Y., Mackerell, A.D., and Cho, B.P. Unusual sequence effects on nucleotide excision repair of arylamine lesions: DNA bending/distortion as a primary recognition factor. *Nucleic Acids Research* (2013). 41(2): 869-880.

Jain, V., **Hilton, B.**, Patnaik, S., Zou, Y., Chiarelli, M.P., and Cho, B.P. Conformational and thermodynamic properties modulate the nucleotide excision repair of 2-aminofluorene and 2-acetylaminofluorene dG adducts in the *Nar*I sequence. *Nucleic Acids Research* (2012). 40(9): 3939-3951.

Tang, H.*, **Hilton, B.***, Musich, P.R., Fang, D.Z., and Zou, Y. Replication factor C1, the large subunit of replication factor C, is proteolyticly truncated in Hutchinson–Gilford progeria syndrome. *Aging Cell* (2012). 11: 363–365.

Jasti, V.P., Das, R.S., **Hilton, B.A.**, Weerasooriya, S., Zou, Y., and Basu, A.K. (5'S)-8,5‘-Cyclo-2’-deoxyguanosine is a strong block to replication, a potent pol V-dependent mutagenic lesion, and is inefficiently repaired in *Escherichia coli*. *Biochemistry* (2011). 50(19): 3862–3865.

In Press:

Hilton, B., Sathyaraj, G., Xu, L., Musich, P., Cho, B., Zou, Y. Dynamic dissociation of the XPC-HR23B complex from damaged DNA is a determining factor of NER efficiency. Under revision: *PLoS One*

In preparation:

Benjamin Hilton, Yiyong Liu, Ji Liu, Rowdy Jones, Hui Tang, Nahid Mehraban, McKayla Johnson, Youjie Wang, Maya Breitman, Antonio Rusinol, Phillip Musich, and Yue Zou. Progerin Sequesters PCNA Resulting in Replication Fork Collapse in Hutchinson-Gilford Progeria Syndrome.

*Equally contributing author

Honors and Awards:

Graduate Student Outstanding Oral Presentation, 15th Annual Midwest DNA Repair Symposium

James H. Quillen Graduate Scholar, ETSU

Graduate Student Outstanding Oral Presentation, 17th Annual Midwest DNA Repair Symposium