Stony Brook University

Study of Recognition Step and Repair Efficiency in Human Nucleotide Excision Repair

A Dissertation Presented

by

Jung-Eun Yeo

to

The Graduate School

in Partial Fulfillment of the

Requirements

for the Degree of

Doctor of Philosophy

in

Chemistry

Stony Brook University

August 2011

Stony Brook University

The Graduate School

Jung-Eun Yeo

We, the dissertation committee for the above candidate for the Doctor of Philosophy degree, hereby recommend acceptance of this dissertation.

Orlando D. Schärer, Ph.D. – Dissertation Advisor
Professor, Department of Pharmacological Sciences

Peter J. Tonge, Ph.D. - Chairperson of Defense
Professor, Department of Chemistry

Iwao Ojima, Ph.D.
Distinguished Professor, Department of Chemistry

Carlos de los Santos, Ph.D.
Professor, Department of Pharmacological Sciences

This dissertation is accepted by the Graduate School

Lawrence Martin
Dean of the Graduate School

Abstract of the Dissertation

Study of Damage Recognition and Repair Efficiency in Human Nucleotide Excision Repair (NER)

by

Jung-Eun Yeo

Doctor of Philosophy

in

Chemistry

Stony Brook University

2011

Nucleotide excision repair (NER) is the main DNA repair pathway dealing with bulky adducts formed by UV irradiation and environmental carcinogens. The importance of NER is underscored by the inherited disorder xeroderma pigmentosum (XP) caused by defects in NER genes and characterized by an enhanced predisposition to skin cancer. Of over 30 proteins involved in NER, XPC-RAD23B is responsible for the recognition of diverse substrates, binding the non-damage strand of DNA and probing for the thermodynamic destabilization induced by the lesion. The helicase subunits in TFIIH are likely candidates to scan DNA to confirm the existence of a damage on DNA. Subsequent to the recognition and verification steps, a 24-32nt size oligonucleotide containing the lesion is excised. It has been found that the higher degrees of duplex destabilization induced by a lesion result in a higher repair rate. However, there are still many unresolved questions in understanding the correlation between the structure of a DNA lesion and its processing by NER.

Research in this thesis focused on various aspects of damage recognition in human NER, aming to understand how the structure of a lesion affects NER, acetylamino fluorene (AAF) and aminofluorene (AF) adducts of DNA, which create distinct degrees of distortion in various sequence contexts in duplex DNA were synthesized. These substrates were incorporated into plasmids and used in NER assay to study the repair efficiency in cell extracts, in binding assays to measure affinity to XPC-RAD23B, and in thermodynamic studies to measure the extent of DNA destabilization. These experiments showed that the degrees of destabilization of DNA duplex induced by AAF and AF indeed correlated with XPC-RAD23B binding affinity and NER efficiency. Similar

experimental methods were applied to DNA lesions induced by another mutagen, aristolochic acid (AA). A previous cellular study had suggested that AA is repaired by transcription coupled (TC) NER, but not global genome (GG) NER. Our studies showed that dA-ALII adducts of DNA are not repaired by GG-NER and not bound specifically by XPC-RA23B, providing a rationale for why ALII adducts persist in human cells.

Traditional NER assays require the use of radioactive materials to detect products and have the limitation that only products but not intermediates and unreacted substrates can be detected. To overcome this limitation, click chemistry was used to prepare fluorescently labeled NER substrates and these yielded promising first results in monitoring NER reactions. These fluorescent substrates should facilitate NER studies, by reducing assay time, to detect various NER reaction intermediates and to study NER reactions in a quantitative fashion.

TABLE of CONTENTS

LIST OF FIGURES

CHAPTER 1:

CHAPTER 2:

CHAPTER 3:

CHAPTER 4:

LIST OF ABBREVIATIONS

AABP	Acetylaminobenzophenone	AA	Aristolochic acid
AAF	Acetylaminofluorene	AAN	Aristolochic acid nephropathy
AF	Aminofluorene	AL	Aristolactam
cccDNA	Covalently closed circular DNA	Cy3 or 5	Cyanine
CS	Cockayne Syndrome	CEP	2-cyanoethyl-*N*,*N*-diisopropylphosphoramidite
DNA	Deoxyribonucleic acid	ds	Double stranded
EMSA	Electrophoretic Mobility Shift Assay	GGR	Global genome repair
GG-NER	Global genome – Nucleotide Excision Repair	XP-F	Xeroderma pigmentosum group F
NER	Nucleotide Excision Repair	NMR	Nuclear Magnetic Resonance
PCNA	Proliferating Cell Nuclear Antigen	Pol II	RNA polymerase II
ss	Single strand	TCR	Transcription coupled repair
TC –NER	Transcription coupled - Nucleotide Excision Repair	TFIIH	Transcription factor IIH
TTD	Trichothiodystrophy	wt	Wild type
XP	Xeroderma pigmentosum	CuAAC	copper-catalyzed azide-alkyne Huisgen type cycloaddition

Acknowledgments

I believe that my research in this dissertation and my life would not have been possible without the help and support of many people around me.

Above all, I would like to thank my advisor Orlando D. Schärer for allowing me to join his lab and giving me a variety of chances to do experiments in his laboratory. I really appreciate that he always opens his office door and his mind to listen eveything I tried to explain and talk as well. He has been supportive and patient during my research and encouraging when I was really depressed and had any problems. During my Ph.D., I have learned a lot of lessons which are not only technical skills/experiments but also the way to be a good scientist from him. I would also like to thank my thesis committee, Dr. Peter Tonge, Dr. Iwao Ojima, Dr. Carlos de los Santos for their advice, time, and support. I really appreciate their support for my presentation in a seminar class and attention to my research as well.

During my Ph.D., I have met many people when I joined Orlando's group. Lidija, Vinh, Barbara, Angelo, AJ, Banke in the lab were so friendly and warm, so that I never felt that I was alone. I really miss all moments we've talked about almost everything inside and outside lab. In addition, I thank Vinh for all his help for several bioassays I had never done, Barbara for NER assay, and Banke especially for Southern blot. Without them, I have never had a chance to learn a lot of experiments in the lab. I thank our group members, Andy, Yan, Burak, Shivam, and Alejandra for their support and conversation over years.

In my graduate school, I met friends who made my life much enjoyable and happy in USA. In my first year in Ph.D. program, I met Dongli, Suhui, and my classmates. Dongli and I discussed a lot about classes and life and Suhui's supported me a lot and we've shared great time for eating, studying, and working out together. I really appreciate her support and help. After I joined the lab, I have met Ilaria worked in Prof. Ojima's group.

I've shared a lot of my thoughts and tears with her when either of us had hard time. Even though she is in Europe now, we still talk and laugh together on skype. I am so happy that I met her and became a friend. I've also met Siyeon who is a good listener and thoughtful person. I really appreciate her help and support when I was depressed.

For aristolactam study in this thesis, we collaborated with Dr. Grollman's lab. I met Victoria in his group. I and Victoria were working together for Aristolactam study and I hope we will have a nice publication together soon. I thank Dr. Grollman as well for his support/encouragement in this research.

I also thank my previous advisor, Prof. Sangho Koo in MS program in South Korea. He is a thoughtful and generous person and I always visit him when I go back to Korea. Thank you, all friends and lab members in Korea for their support.

Finally, I would like to thank my parents and family. I know that my mother and father think about me all the time and I really appreciate their help and support. Without them, I would not reach this far. Thank Mom and Dad!! Also, Thank you, my sister, Tina and her husband, Jeff for their encouragement. I am really happy that they are my family. I would like to thank my fiancé, Minyong. We have been through a lot of events together for 10 years. I really appreciate that he always respects my decision and encourages me to do what I want. I never forget the moment, when I had a surgery about 2 years ago, you were there for me all the time and made me strong. You make my life so valuable.

CHAPTER 1:

GENERAL INTRODUCTION FOR NUCLEOTIDE EXCISION REPAIR

Introduction

Every day the DNA in each cell is harmed by a variety of exogenous and endogenous sources such as ultraviolet (UV) light, environmental mutagens, or reactive metabolites that generate lesions in DNA. The estimated number of DNA damage events is approximately about 10^5 -10^6 per cell per day [1]. The consequences of DNA damage are diverse and unpleasant, causing different responses and consequences in the cell (Figure 1). Cell-cycle arrest can be induced to give the cell some time to repair the damage before replication and cell division take place [2]. If unresolved, DNA damage can induce mutations that can result in carcinogenesis. If the damage load to a cell is too high, an apoptotic response can be triggered resulting in cell death (Figure 1). Since alterations in DNA can cause adverse results, DNA repair pathways to counteract this damage are indispensable for all cells. Figure 1 shows an overview of the most common DNA-damaging agents, the lesions they induce on DNA, and the repair pathways that can remove their lesions from the genomes of cells. There are five major pathways (direct repair, base excision repair, nucleotide excision repair, homologous recombination and non homologous end joining, and mismatch repair) (Figure 2) that are able to deal with different types of lesions. Among these, the work in this thesis was particularly concerned with investigations of nucleotide excision repair (NER), which removes bulky adducts generated by UV light or environmental mutagens from DNA.

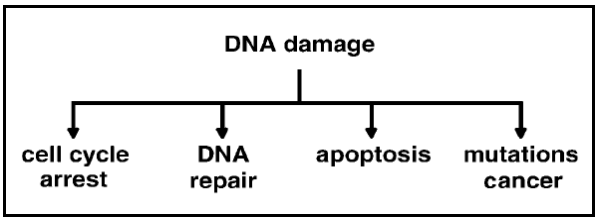


Figure 1: Responses to and consequences of DNA damage [7].

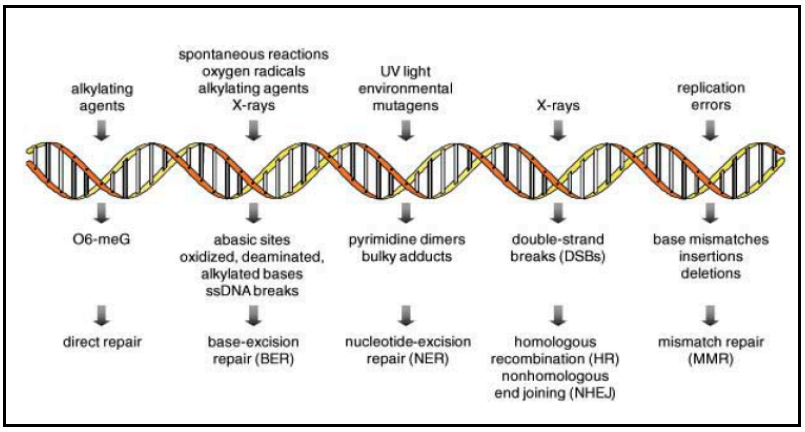


Figure 2: Most common DNA-damaging agents, lesions, and repair pathways [7].

Nucleotide Excision Repair

NER is the most versatile DNA repair process due to its ability to recognize a wide range of helix-distorting DNA lesions formed by UV light or environmental mutagens [3]. Figure 3 contains examples of lesions such as cyclobutane pyrimidine dimer (CPD), 6-4 photoproduct (6-4PP), generated by UV light and acetyl amino fluorene (dG-AAF) created by an environmental mutagen respectively. Although unrelated in primary chemical structure, the common feature of all lesions recognized by NER is that they cause some degree of distortion and thermodynamic destabilization of duplex DNA. The degree of distortion/thermodynamic destabilization induced in a DNA duplex is correlated with the rate of repair by NER [4].

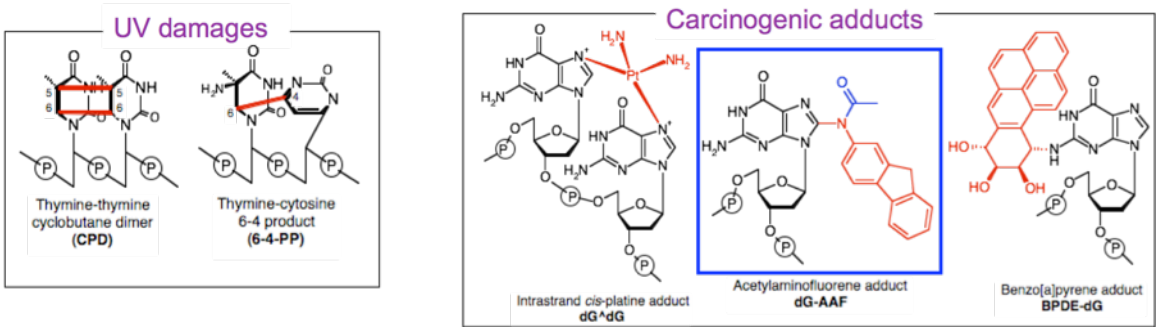


Figure 3: NER substrates. There are two main adducts formed by UV light or environmental mutagens. UV adducts contain thymine-thymine cyclobutane dimmer (CPD) and thymine-cytosine 6-4 product (6-4 PP). Chemically induced adducts involve intrastrand cis-platine adduct, acetylaminofluorene adduct (dG-AAF), and benzopyrene adduct (BPDE-dG).

Two distinguishable subpathways of NER exist: global genome NER (GG-NER), which is able to remove adducts in the entire genome and transcription-coupled NER

(TC-NER), which is responsible for the accelerated repair of lesions in transcribed DNA strands of active genes [3,4,5]. The notable difference between these two NER subpathways is that helix-distorting DNA lesions are recognized by specific NER factors in GG-NER, whereas TC-NER is believed to be triggered by a stalled RNA polymerase (elongating RNA Pol II). Defects in NER genes are associated with three inherited disorders: xeroderma pigmentosum (XP), Cockayne syndrome (CS), and trichothiodystrophy (TTD) [6]. In particular XP patients with defects in GG-NER and, in most cases also TC-NER, are highly sensitive to solar UV irradiation and have a 1000-fold increased predisposition to skin cancer. XP patients also show extreme sensitivity to UV radiation and severe skin, eye and in severe cases neurological abnormalities (Figure 4) and they are highly sensitive to UV irradiation, requiring patients to wear special protection from solar UV light such as the special suit made by NASA (Figure 4).

Figure 4: Images of XP patients. A patient on left figure shows skin abonormality and a child with XP wears UV-protective suit from NASA [8,9].

Mutations in CSA and CSB genes involved TC-NER are associated with the rare human disorder CS [24, 25]. CS and TTD [26] patients, have deficiencies in TC-NER and a mild defect in transcription, but are proficient in GG-NER. These patients also show UV sensitivity, but do not display a dramatic predisposition toward developing skin cancer. Instead, they suffer from various developmental and neurological abnormalities. Investigation of the NER process therefore contributes to our understanding of human disorders and carcinogenesis.

Molecular Mechanism of Nucleotide Excision Repair (NER)

GG-NER involves the concerted action of over 30 proteins: XPC-RAD23B, the ten subunit factor TFIIH, XPA and the RPA trimer are necessary for damage recognition and verification, the two endonucleases ERCC1-XPF and XPG incise the damaged DNA strand, and replication factors including the polymerases Polδ, Polε and Polκ, PCNA, RFC, DNA ligase I and DNA ligase IIIα/XRCC1, are involved in repair synthesis and ligation [7]. In GG-NER, the XPC-RAD23B heterodimer is the first factor that recognizes DNA lesions and initiates NER [10, 11] (Figure 5, step a). A recent structure of the yeast XPC ortholog Rad 4 bound to an oligonucleotide with a CPD lesion showed that this dimer binds to the non-damaged strand opposite the lesion rather than the lesion itself [12] and the recognition of non-damaged strand provides a mechanism to distinguish a wide variety of damaged bases from non-modified DNA. After the recognition of a lesion, XPC-RAD23B recruits the ten-subunit transcription and repair factor TFIIH to the damaged DNA. The two helicase subunits of TFIIH, XPB and XPD, partially unwind the DNA helix and verify the presence of the damaged residue [13, 14] (Figure 5, step b). Some studies have suggested the possible involvement of XPB and XPD in damage verification step in which ATPase activity of XPB pries apart the two DNA stands, while XPD scans along the DNA with a 5'-3' polarity until it encounters and stalls at a lesion [56, 57]. The stalling of XPD at a lesion ensures that NER is only complete if a chemical modification is present in the DNA. Full helix opening is then accomplished once RPA, XPA, and XPG are recruited to the pre-incision complex [15-17] (Figure 5, step c). This allows for the recruitment of the ERCC1-XPF endonuclease, which performs the incision 5' to the lesion [18. 19], whereas the endonuclease XPG cleaves 3' to the lesion [20] (Figure 5, step e). It was recently proposed that 5' incision by ERCC1-XPF takes place first and followed by intitiation of repair synthesis and 3' incision by XPG [21]. This result is consistent with the cut and patch mechanism originally proposed by Hanawalt and co-workers [22, 23] (Figure 5, step d and e). Finally repair synthesis is completed and ligation takes place to fill and seal the gap created by the incision step (Figure 5, step f).

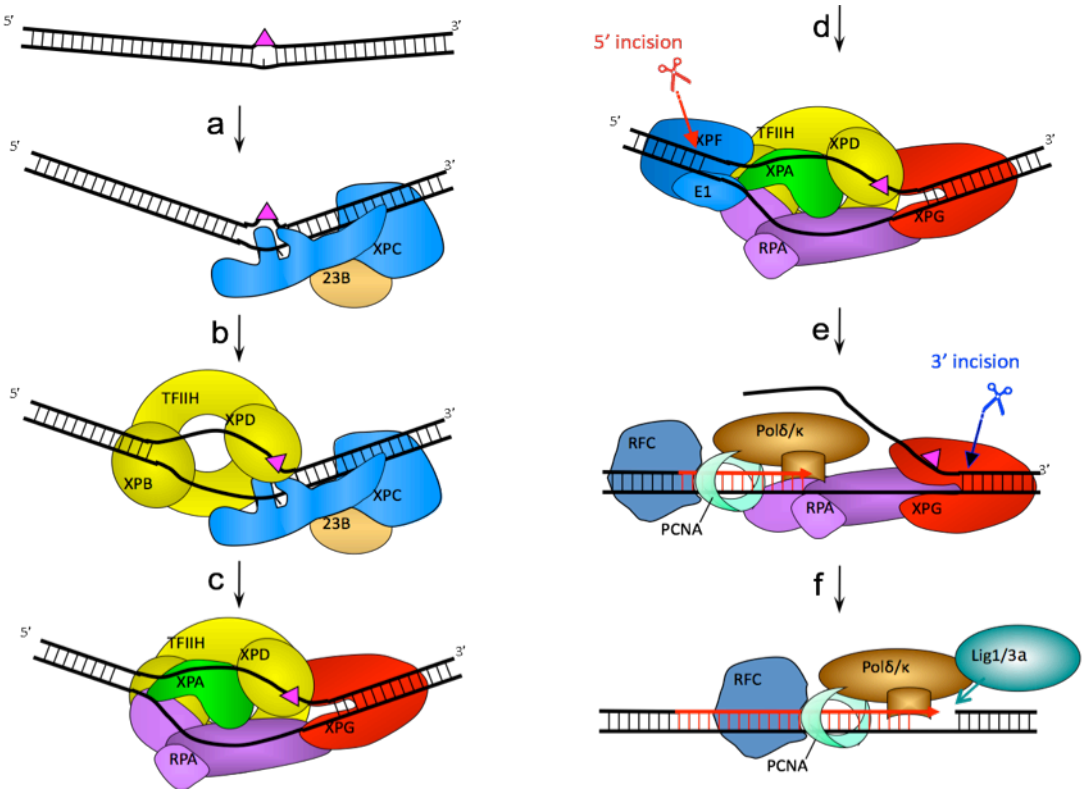


Figure 5: A Model of the Nucleotide Excision Repair pathway. Heterodimer, XPC-RAD23B binds to a lesion and recruit TFIIH. Two subunits, XPB and XPD of helicases, TFIIH unwind DNA around a lesion and recruit other proteins such as XPA, XPG, and RPA. The last factor, XPF for a pre-incision complex is an endonulcease. 5' incision by XPF takes place and subsequently, XPG incises the DNA containing a lesion to the 3' site. The replication machinery consisting of Polδ/κ, PCNA and the replication factor C (RFC) fills the gap and Lig I/III seals the nick.

Hanawalt and co-workers discovered 25 years ago that bulky DNA lesions are repaired more quickly in the transcribed strands of active genes, in a process called transcription-coupled NER [24, 27]. The mechanistic details of TC-NER have so far remained elusive [5]. TC-NER is independent of the recognition factor XPC-RAD23B, but instead relies on RNA polymerase II to detect a lesion and initiate the assembly of downstream NER factors in a process that requires the TC-NER specific factors CSA, CSB [25] and XAB2 [28, 29] (Figure 6). After the recognition of a damage on DNA by either XPC-HR23B or RNAPII, the two subpathways require the same NER core factors [28, 29] to complete the two pathways.

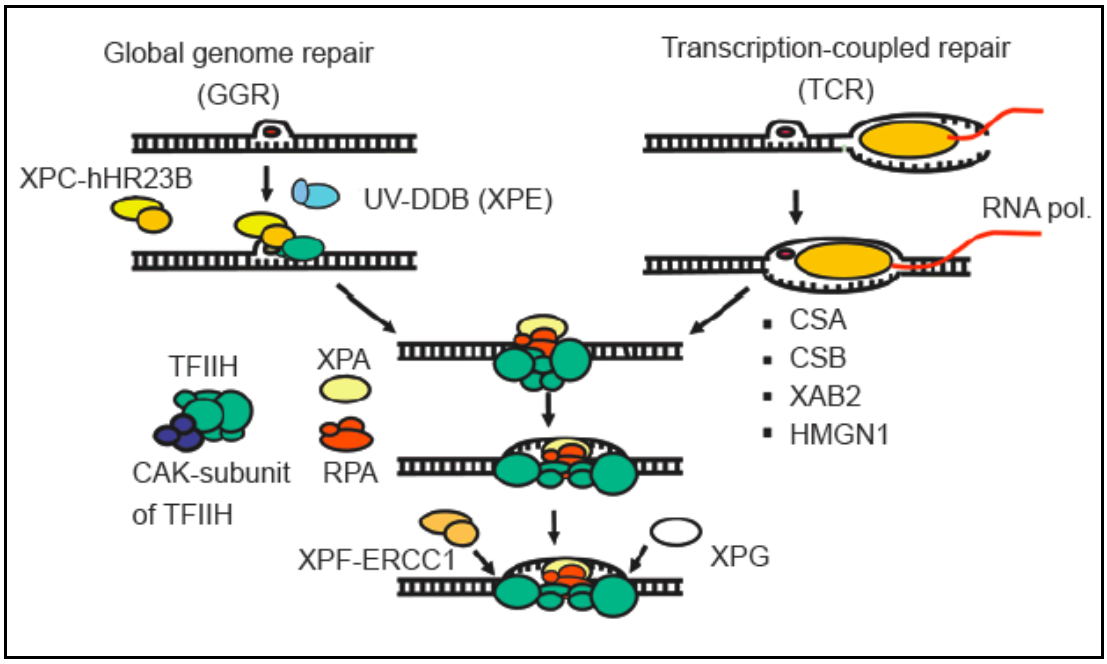


Figure 6: Two subpathways of mammalian NER. In GG-NER, XPC-RAD23B recognizes a lesion that destabilizes DNA and initiates GG-NER. The blockage of RNA polymerase is the first step to trigger TC-NER. After these recognition steps by either XPC-hHR23B or RNA polymerase (RNA pol II), TFIIH is recruited to verify the lesion and employs other NER factors to give a oligonucleotide containing a lesion and then subsequently a repair synthesis is followed [26].

Damages Repaired by NER

The most interesting feature of NER is its ability to recognize and excise a variety of DNA damages (Figure 3). There are two main sources responsible for introducing bulky lesions in DNA: one is UV light from sun and the other one includes carcinogenic compounds present in industrial exhaust, cigarette smoke and barbequed food. As shown in Figure 3, CPD and 6-4PP are typical UV lesions, whereas dG-AAF, and benzopyrene adducts are prime examples of adducts formed by known carcinogenic compounds. The common feature of the diverse NER substrates is their propensity for creating severe distortion in the DNA. The goal of the research in this dissertation was to determine how structural characteristics of DNA lesions influence how they are recognized and repaired by NER proteins.

UV damages

Solar UV irradiation is most detrimental to humans and other living organisms upon reaching earth. There are three types of UV that differ in their wavelength range: UVA (320-400nm), UVB (290-320nm), and UVC (200-290nm). Although UV of the shortest wavelength (UVC) is the strongest among the three, most of it is filtered (along with shorter-wavelength UVB) by the ozone layer. The UVA irradiation, although 1000 times weaker than UVB, can reach the earth 100 times more efficiently [30]. In addition, UVA can penetrate human skin more deeply due to its longer wavelength. However, the relationship between the depth penetrated by UV irradiation and photodamage and skin cancer is still not clear. While other organisms have additional pathways to address UV damage, NER is the only way to deal with UV damages in placental mammals [31].

Cyclobutane Pyrimidine Dimers (CPDs) are the most dominant products of UVB irradiation. As shown in figure 4, a [2+2] cycloaddition reaction takes place at the C5 – C6 double bonds of two adjacent pyrimidine bases. Although the prevalent product among CPDs is the TT pyrimidine dimer, there are other products such as TC, CT, or CC dimer are also formed depending on the DNA sequence and the wavelength of irradiation.

6-4 Photoproducts are the second most prevalent adducts of UV irradiation. These adducts are formed by a [2+2] cycloaddition between the C5-C6 double bond of a pyrimidine and the C4 carbonyl group of a thymine or the N4 imino group of a cytosine located 3’ to it.

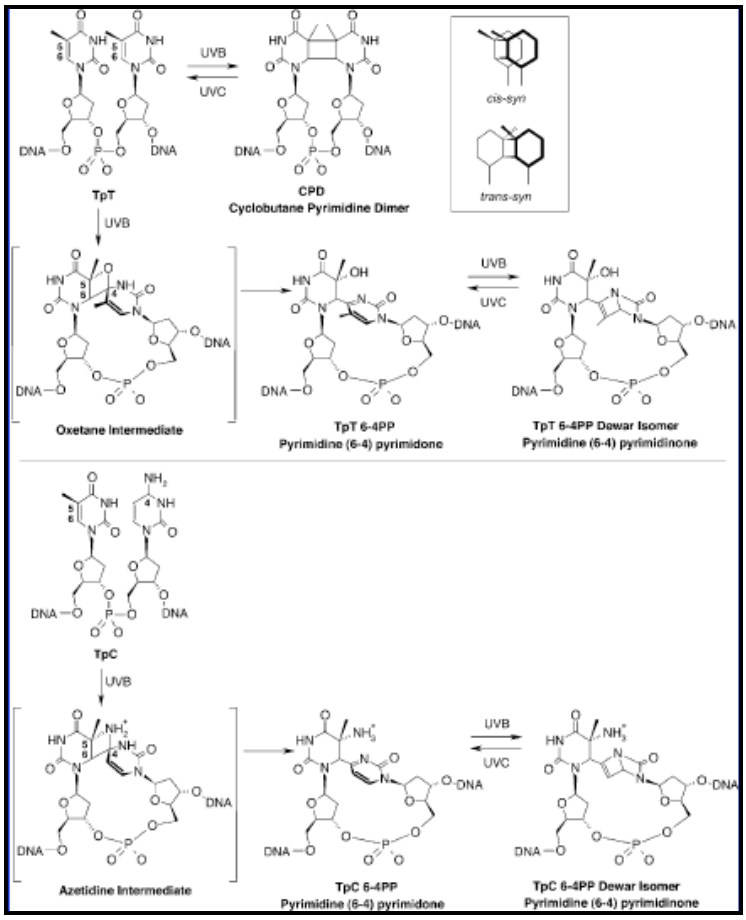


Figure 7: Ultraviolet TpT photoproducts. The structure and formation of CPD and 6-4pp.The rectangle box contains the different diastereoisomers of the CPD. The different wavelengths for each formation are indicated [4].

Although CPDs and 6-4PPs are produced by similar wavelengths, they show a difference in how they distort the structure of DNA. NMR and thermodynamic studies have shown that CPDs only create a small degree of distortion of a DNA duplex and maintain Watson-crick base pairing [32, 33]. On the other hand, 6-4 PPs cause severe distortion on DNA, leading to disruption of Watson-Crick base pairing [34]. The different degrees of DNA distortion induced by these two photolesions have been highly correlated with the repair efficiency, and 6-4PPs are much better NER substrates than CPDs [35].

DNA lesions formed by electrophilic molecules

Carcinogenic compounds are preponderantly electrophilic molecules reacting with the nucleophilic atoms present in the DNA such as the phosphodiester linkages, N2 and N7 of guanine or N3 and N6 of adenine. The most reactive position is N7 of guanine with high electronic density.

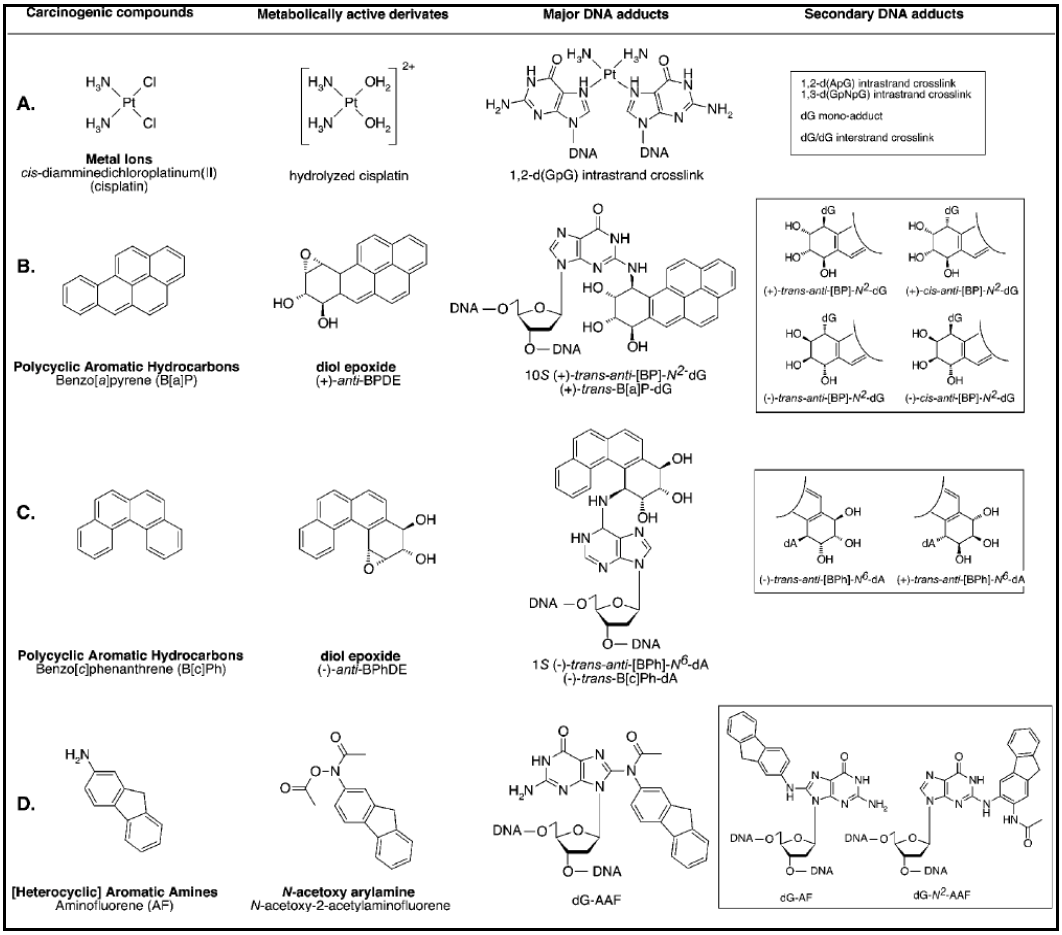


Figure 8: Examples of carcinogenic compounds and their DNA adducts (major/minor). Carcinogenic compounds, their activated derivatives, and the major and secondary adducts [4].

***cis*-Diaminodicholoroplatinum (cis-DDP or cisplatin, Figure 8A)** is one of the most commonly used chemotherapeutic drugs for treating various types of cancers. It is a neutral and square planar platinum (II) complex covalently binding to two inert ammonia groups and two chloride ligands in *cis* configuration [36]. Two chlorides are substituted by water or hydroxyl ions when it diffuses into cytoplasm of cells which

contain lower salt (chloride) concentration than the extracellular matrix. This replacement reaction results in formation of hydrolyzed cisplatin, an activated electrophilic derivative of cisplatin (Figure 8A). There are several possible intrastrand cross-link (N7-N7) products: a major adduct, 1,2-d(GpG) (up to 65% of all cisplatin-induced lesions) or 1,2-d(ApG) intrastrand cross-links (25% of all lesions), and 1,3-d(GpNpG) adducts (5-10%) containing one nucleotide(N) between two guanines [34]. These diverse cisplatin adducts form different structures within the DNA duplex: the major adduct, the 1, 2-d(GpG) intrastrand cross-link maintains intact base-pairing even though the cross-link creates a prominent kink [36]. Interestingly, 1, 3- intrastrand adduct shows substantial helix unwinding and base pairing disruption at the bases adjacent to the cisplatin adduct [36].

Aromatic Hydrocarbons (Figure 8B and 8C) are mostly found in cigarette smoke and industrial and vehicle emissions [37]. Without any modification, they are nonpolar, chemically unreactive. However, aromatic hydrocarbons become effective carcinogens as electrophilic epoxide derivatives that are formed as intermediates during metabolic detoxification in the body [38] (Figure 8B and 8C). The activated forms can easily react with exocyclic amino groups of guanine and adenine (Figure 8B and 8C for major adducts). In addition, various stereoisomers formed with DNA influence their conformation and thermodynamic stability within DNA, resulting in different degrees of repair efficiency [38].

Adducts of aromatic amines such as Aminofluorene (AF) and Acetyl Aminofluorene (AAF), two prototypical NER lesions

As shown in Figure 3, a variety of lesions can be removed by NER and among them, we were particularly interested in the dG-AAF and dG-AF lesions (Figure 9 and 10). These can be synthesized efficiently by a protocol developed by Ludovic C. J. Gillet in our laboratory [39] and their interesting structural features make them attractive targets for studying damage recognition in NER. AF and AAF were initially synthesized for potential use as insecticides. However, these compounds turned out to be powerful carcinogens and are no longer used as insecticides [40, 41]. AF and AAF have been

used intensively to study chemical mutagenesis and DNA repair [41], due to their structural characteristics (Figure 9 and 10) and carcinogenicity.

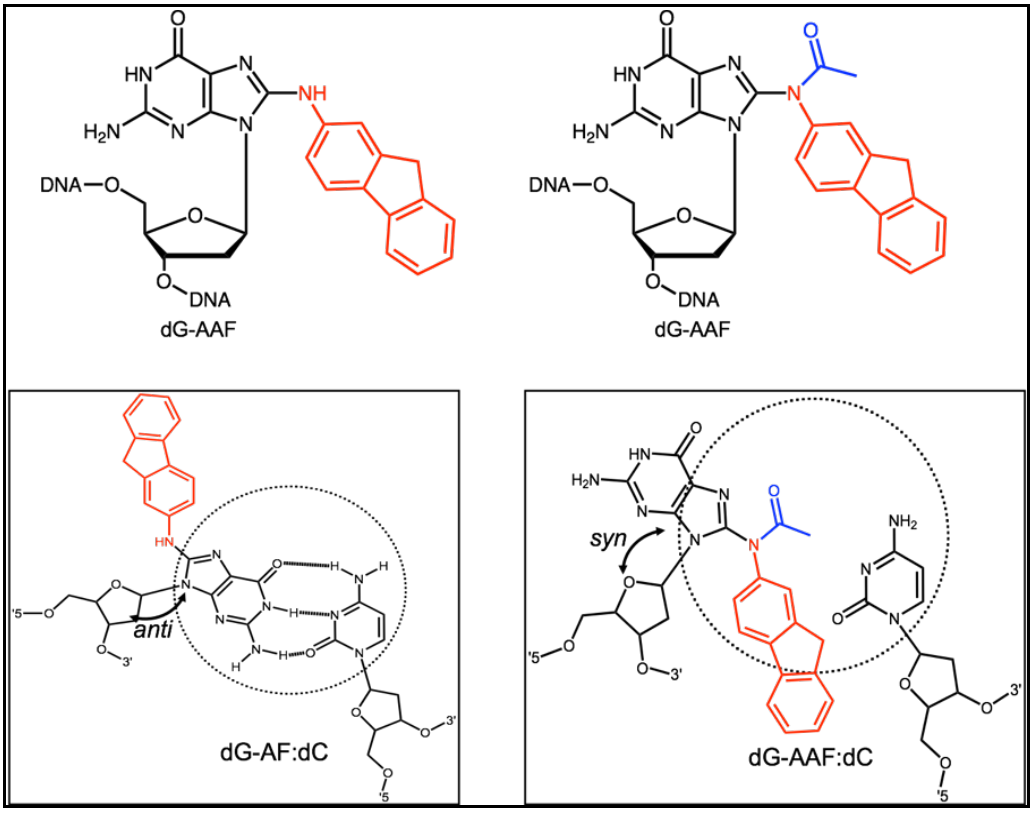


Figure 9: Structures of dG-AF and dG-AAF. (AF: aminofluorene, AAF: acetyl aminofluorene)

Although AF and AAF differ by only one acetyl group (Figure 9), they can induce distinct structural alterations in a DNA duplex, resulting in different biological effects. For example, it has been shown that the AAF adducts induce mainly frameshift mutations and block DNA synthesis whereas AF adducts lead predominantly to point mutations during replication in *E.coli* [40, 42-45]. With respect to NER, previous studies have found that AAF adducts are generally processed at higher rates, as they induce a higher degree of distortion and destabilization into the DNA helix [45, 46]. Figure 9 shows the conformations of dG-AF and dG-AAF in the DNA duplex. Even with AF present at C(8) of dG, the nucleoside can still assume the *anti* conformation around the glycosidic bond and the bases can form Watson-Crick base pairs. However, the steric hindrance between the acetyl group at the N8 position in AAF, and the sugar ring force dG-AAF to adapt a *syn* conformation, thereby preventing Watson-Crick base pair formation (Figure 9 and 10) [41, 49, 50]. Futhermore, AAF on dG generally creates more severe distortion

than AF, although the conformation of both adducts vary with the sequence context of the surrounding bases. Consistent with the notion that increased distortion correlates with higher repair in the NER pathway, dG-AAF adducts are generally repaired faster than dG-AF adducts [1].

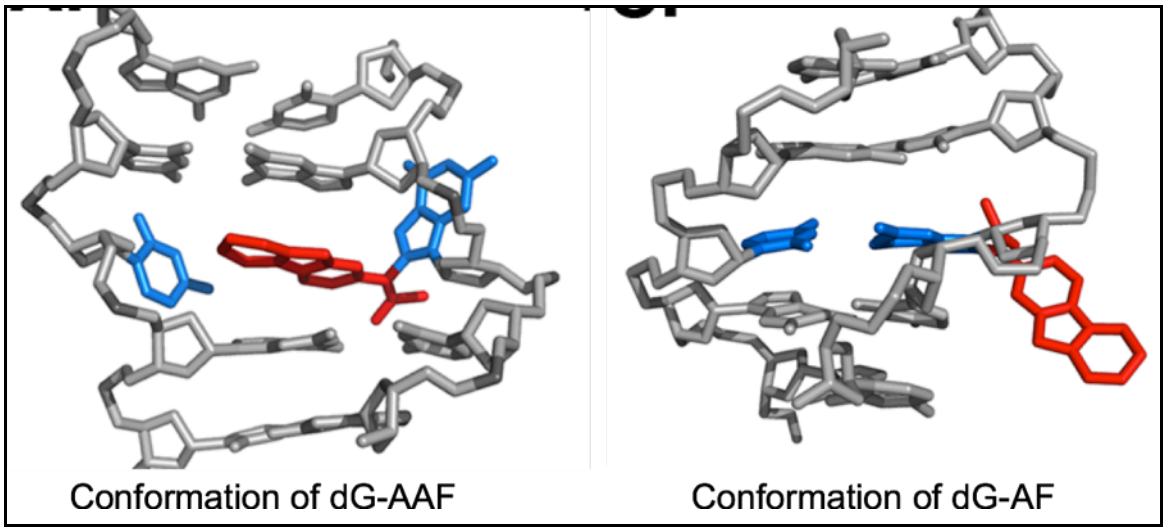


Figure 10: *syn* and *anti* conformations of dG-AF and dG-AAF. (Computational results from Prof. Suse Broyde in New York University)

Influence of Damages to Recognition Step and Repair Process in NER

As described above, NER can deal with a variety of lesions that share severe duplex distortion as a common feature. The crystal structure of the yeast XPC orthologue RAD4 in a complex with a DNA duplex containing a CPD lesion unveiled that two β-hairpins of RAD4 is inserted into the DNA duplex at the site of the lesion [12]. RAD4 recognizes the exposed nucleotides on the undamaged strand, without making contact with the damaged strand. The CPD lesion was found to be disordered and extruded from the helix. This finding strongly supports the notion that XPC recognizes the ssDNA character of the destabilized undamaged strand, rather than the damage itself. These observations provide structural basis for the ability of XPC to detect extraordinarily diverse lesions that cause thermodynamic destabilization of a DNA duplex.

The efficiency of NER of various damages has been studied intensively and these studies supported the idea that a higher degree of distortion/destabilization results in higher repair efficiency [46]. Naegeli and co-workers studied how the structure of many

different substrates influences NER. Interestingly, they found that the chemical modification and structural distortion can be separated. Modification of deoxyribose-phosphate backbones of DNA that maintain normal hydrogen bonding were not repaired by NER, unless those are placed in the context of a mismatch that destabilizes the DNA helix [46]. This finding lead then to propose a “bipartite substrate discrimination” model, suggesting that NER is triggered by lesions that induce DNA destabilization *and* contain a chemical modificaiton [47]. This proposed model nicely explains how the distortion created by lesions is responsible for recruitment of NER factors [48].

Efficiency of NER Can Vary with the Sequence Context for A Given Lesion

The efficiency of NER is not only dependent on the lesion, but can also greatly vary with the sequence context in which a lesion located. Such structure-function correlations have, for example, been established for benzo[α]pyrene adducts by Geacintov, Broyde and coworkers [51-54]. These studies have suggested that several factors, including disruption of base pairing, bending, and flexibility need to be considered for damage recognition in NER and supports the concept that thermodynamic stabilization influenced by all parameters is the key to controlling the NER process [54, 55]. In order to understand the relationship between NER efficiency, the structure of the damage and the sequence context in which the damage is located, we used dG-AAF and dG-AF lesions in the *Nar*I restriction enzyme recognition sequence ($G^1G^2CG^3CC$) in our investigations. The *Nar*I site initially attracted attention as mutational hotspot sequence in *E. coli* and has the remarkable property of inducing -2 frameshift mutations, leading to the sequence G-G-C-G, but only if damage is located on G^3. Moreover, AAF has a much greater propensity to induce frameshifts at this site than does AF [42]. Studies have shown that AF adducts on G^3, but not G^1or G^2, induces similar degree of structural distortion created by AAF [43]. Therefore, the AAF and AF lesions in the context of the *Nar*I sequence provide a good model system to study how structural alterations induced by DNA lesions in different sequence contexts influence NER activity.

PREVIEW

Work in this thesis is aimed at understanding how different degrees of structural destabilization created by DNA lesions affect damage recognition and efficiency of NER in a systematic fashion and to develop new ways of monitoring NER in a more quantitative fashion. Starting with the preparation of plasmids containing a site-specific dG-AAF and dG-AF lesion, we performed various sets of experiments to study the correlation between DNA structure and NER efficiency (Chapter 2). We also investigated the NER efficiency and recognition by XPC-RAD23B to adducts of the herbal mutagen aristolochic acid with DNA, providing a molecular understanding for why these adducts are only poorly repaired by NER (Chapter 3). Finally, we designed and prepared fluorescently modified substrates to monitor the NER reaction in a more quantitative and user friendly manner (Chapter 4).

CHAPTER 2:

THERMODYNAMIC DESTABILIZATION OF DNA DUPLEXES CORRELATES WITH EFFICIENCY OF THE RECOGNITION STEP AND REPAIR IN NUCLEOTIDE EXCISION REPAIR

Abstract

Nucleotide excision repair (NER) is a versatile pathway that can remove deleterious modifications to DNA caused by environmental mutagens or UV light. A common feature of good NER substrates is that they distort and/or thermodynamically destabilize the DNA helix. Moreover, there is evidence that more destabilizing lesions are repaired with higher efficiency by NER. We aimed to investigate the interdependence of the degree of structural distortion in DNA and repair efficiency by NER in a systematic fashion. For this purpose we introduced oligonucleotides containing site-specific acetylaminofluorene (AAF) and aminofluorene (AF) adducts at the C8 position of guanine on all positions of the *Nar*I ($G_1G_2CG_3$) sequence. We found that the efficiency of repair by NER as well as the binding affinity of the damage recognition factor XPC-RAD23B depended on which G within the *NarI* sequence AF or AAF was located. We determined the melting points and thermodynamic stabilities of *NarI* sequence with AAF or AF on G_1, G_2, and G_3 and found that the stability correlated with the XPC-RAD23 binding and NER efficiency. These studies demonstrate that the different degrees of thermodynamic destabilization induced in a DNA duplex correlated with the efficiency of the repair process by NER.

Introduction

As discussed in chapter 1, nucleotide excision repair (NER) is a repair pathway dealing with a variety of bulky adducts formed by UV irradiation and environmental carcinogens [3, 4]. Of more 30 proteins involved in NER, XPC-RAD23B is responsible for the recognition of diverse substrates [4]. XPC-RAD23B employs a general mechanism for damage recognition, binding to the non-damaged strand of DNA and probing for the thermodynamic destabilization induced by a lesion [12]. XPC-RAD23B has a binding affinity to lesions as well as other destabilized sites such as mismatches, making a verification step necessary to probe for the presence of a chemical alteration [11, 51]. The XPB and XPD helicases, subunits of TFIIH, are potential candidates to scan DNA to confirm the existence of a damaged site [56, 57]. The recognition and verification steps in NER are crucial to trigger the rest of NER process leading to the excision of 24-32nt size fragments containing the lesion.

Many studies have shown that a higher degree of distortion results in higher repair efficiency [46]. Naegeli and co-workers studied many different substrates that are recognized and repaired by NER, and have proposed a “bipartite substrate discrimination” model. This model suggests that lesions are processed by NER if they contain a chemical modification and impede base pairing [48]. Other studies have extended the “bipartite substrate discrimination” model, suggesting that a DNA duplex destabilization in a more general sense facilitates NER and may include additional factors such as bending and flexibility in addition to disrupted base pairing [51-53]. These findings resulted in the idea that thermodynamic destabilization of a DNA duplex influenced by a variety of parameters is the main trigger for NER [54, 55].

To gain more insights into the mechanisms of damage recognition in NER, we chose to use the DNA adducts of aminofluorene and acetylaminofluorene, dG-C8-AF (*N*-(2’-deoxyguanosine-8-yl)-2-aminofluorene) and dG-C8-AAF (*N*-(2’-deoxyguanosin-8-yl)-2-acetylaminofluorene) (Figure 1). AF and AAF, the prototypical aromatic amines, were originally developed as a pesticide [41], but were found to be potent mutagens and carcinogens [4]. dG-AF and dG-AAF adducts have structurally and biologically captivating properties. These adducts can assume two main conformations in DNA that

depend on the conformation of the glycosidic bond. While dG-AF can assume a *syn* or *anti* conformation, dG-AAF is primarily restricted to the *syn* conformation due to the bulk at the N (8) position (Figure 1) [41, 50]. Whereas the adducts in the *anti* conformation can undergo normal Watson-Crick base pairing, and have the AF located in the major groove of DNA, adducts in the *syn* conformation have the base extruded from the helix and the aromatic ring of the AAF intercalated in DNA duplexes. Furthermore, ^{19}F NMR studies of the conformation of AF on dG confirmed that AF has a more flexible conformation than AAF and contains a 1:1 ratio mixture of *syn* and *anti* [58].

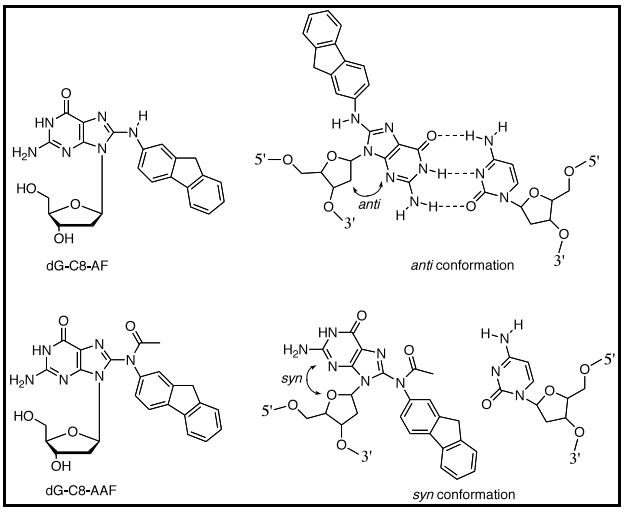


Figure 1: Structures of dG-C8-AF; *N*-(deoxyguanosin-8-yl)-2-aminofluorene and dG-C8-AAF ; *N*-(deoxyguanosin-8-yl)-2-acetylaminofluorene and *anti* and *syn* conformations. dG-C8-AF is believed to have the base mainly in the *syn* conformation, whereas dG-C8-AAF favorably retains *anti* conformation. The hydrogen bonds are shown as dashed lines.

The aim in this research is to investigate the relationship between the degree of thermodynamic destabilization induced by AF or AAF in DNA and the efficiency of NER in a systematic fashion (Figure 2). For this purpose, we synthesized oligonucleotides containing site-specific acetylaminofluorene (AAF) and aminofluorene (AF) adducts at the C8 position of guanine on all positions of the *NarI* ($G_1G_2CG_3$) sequence. Depending on which G the lesion was located, various degrees of distortion were induced. In this chapter, we explored how the extent of destabilization correlates with NER activity and binding affinity to the damage recognition protein XPC-RAD23B. These findings are consistent with the idea that more destabilizing adducts are repaired more readily as they display higher binding affinity to the damage sensor XPC-RAD23B.

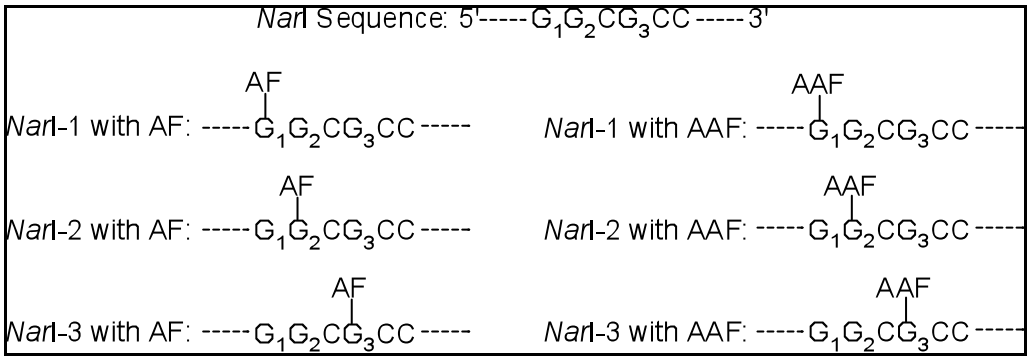


Figure 2: Six modified oligonucleotides with either AF or AAF on *Nar*I sequence (GGCG) used in this study. 12mer, 24mer and 44mer containing site-specific dG-AF or dG-AAF lesions were prepared from DNA synthesizer. AF; aminofluorene, AAF; acetylaminofluorene.

RESULTS

Syntheses of dG-AAF and dG-AF

This study requires the availability of defined substrates containing site-specifically damaged DNA. Many potent compounds that induce mutagenesis and cancer are heterocyclic aromatic amines that form adducts with C8 of 2'-deoxyguanosine (dG) [59]. One of the most potent NER substrates is 8-(*N*-acetylaminofluorene)-2'-deoxyguanosine (dG-AAF). The damaged base, dG-AAF and its relative dG-AF were prepared and incorporated into DNA using solid-phase synthesis for our studies using a synthetic route developed by L.C.J. Gillet, a former lab member [39, 60].

In order to synthesize dG-AAF, a fully protected Br-dG that is an essential starting material, was employed in a Büchwald-Hartwig coupling reaction with aminofluorene (Figure 3). To obtain the fully protected Br-dG, dG was first brominated (Figure 3a) at the C8 position by a radical reaction, followed by protection of the 5' and 3' alcohols with *tert*-butyldimethylsilyl chloride and a Mitsunobu reaction to protect O^6 position with benzyl alcohol. The last protecting reaction was N^2 position with DMTr-Cl, which is stable under basic condition used for the coupling reaction with the amine. Using this fully protected Br-dG, a Büchwald-Hartwig Coupling reaction with aminofluorene was used to provide protected dG-AF. The acetyl group was selectively introduced at the

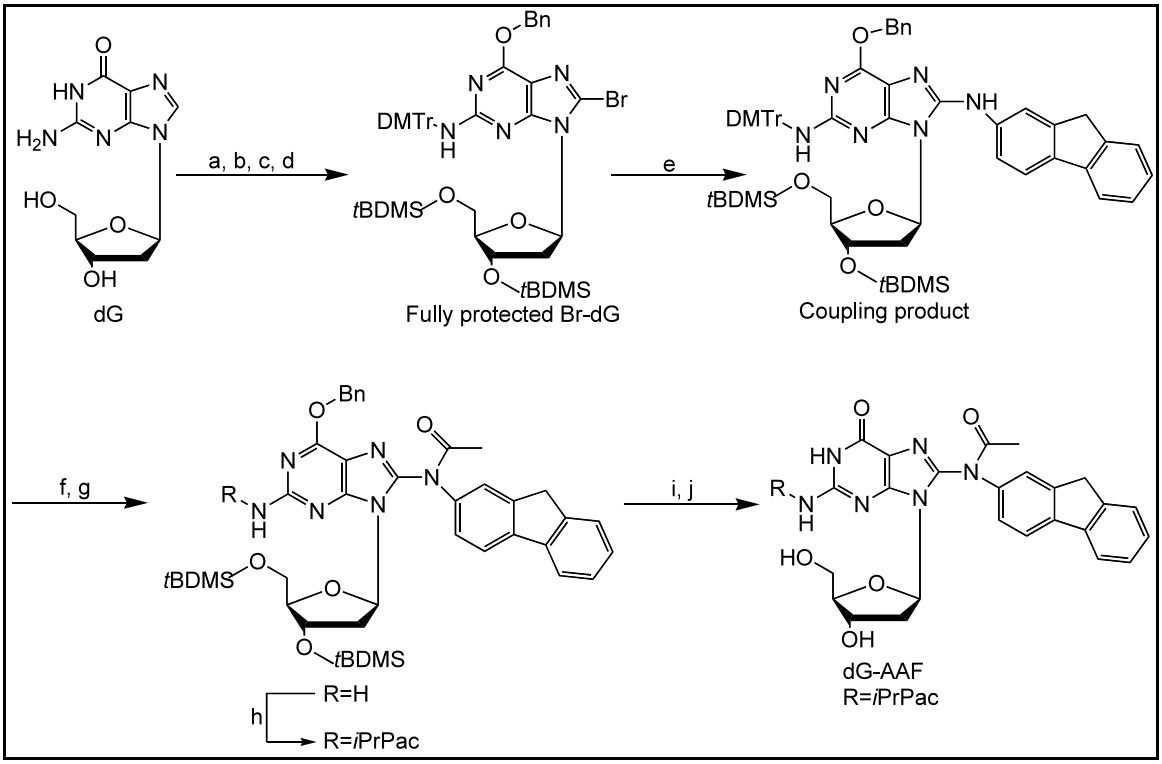


Figure 3: Preparation of an acetylamine adduct of 2'-deoxyguanosine (dG-AAF). Reaction conditions: (a) NBS, water, acetonitrile, rt, 2hrs (80%); (b) *t*BDMS-Cl, imidazole, DMF, rt, 5hrs (98%); (c) Bn-OH, PPh_3, DIAD, dioxane, rt, 5hrs (78%); (d) DMTr-Cl, pyridine, rt, overnight (94%); (e) 2-aminofluorene, $Pd_2(dba)_3$, BINAP, toluene, NaO*t*Bu, 90-100°C, 1hr and 40min. (72%) (f) Ac_2O, Et_3N, DMAP, pyridine, rt, 5hrs; (g) HCl 0.01M, MeOH, rt, overnight (78% for f-g); (h) iPrPac-Cl, pyridine, rt, overnight (97%); (i) CH_3COOH, TBAF, THF, 0°C→ rt, overnight (97%); (j) 10%Pd/C, H_2, DCM, 80°C, 4hrs (78%).

N^8 position of dG-AF to yield the dG-AAF derivative with the trityl group preventing acetylation of N^2 in this reaciton. The trityl group was then removed under weakly acidic conditions and isopropylphenoxyacetyl group was introduced at N^2 position for protection during DNA synthesis. It has previously been shown that this protecting group can be removed after oligonucleotide synthesis under mildly basic conditions that ensure the integrity of the acetyl group at N^8 [60]. The last steps required to obtain dG-AAF were the removal of the silyl protecting group using TBAF buffered with acetic acid to avoid depurination and of the benzyl group using a palladium-catalyzed hydrogenolysis. For the purpose of oligonucleotide synthesis, demethoxytritylation of the 5' alcohol group and phosphitylation of the 3' alcohol group yield the phosphoramidite of iPrPac-dG-AAF needed for incorporation into oligonucleotides (Figure 4).

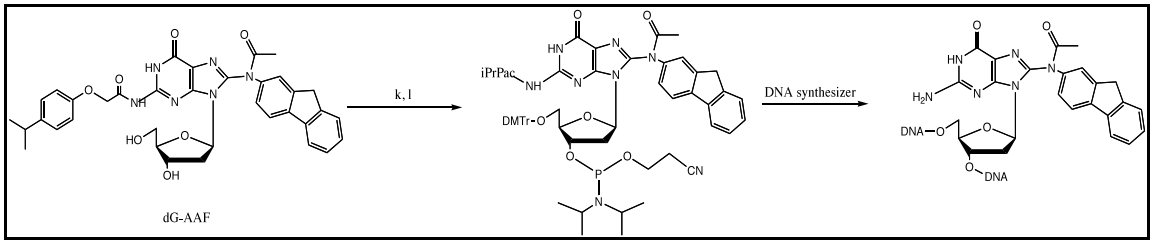


Figure 4: DNA synthesis with dG-AAF. Reaction conditions: (k) DMTr-Cl, pyridine, rt, 2hrs (76%); (l) *N*-ethyldiisopropylamine, 2-cyanoethoxydiisopropylchlorophosphoramidite, CH_2Cl_2, rt, 2hrs.

Preparation of Oligodeoxynucleotides

In general, the preparation of dsDNA containing a defined site-specific lesion is the one of the most important, but also laborious steps in the study of DNA repair processes. In this study, maintenance of the base-labile acetyl group in dG-AAF is critical during oligonucleotide synthesis. Previous studies established that the N^8 acetyl group was lost upon treatment with ammonia [61] or K_2CO_3/MeOH, standard oligonucleotide deprotecting conditions. To preserve the integrity of the acetyl group, our group had published a general solid-phase DNA synthesis protocol employing ultra-mild conditions [60]. In the modified protocol, an isopropylphenoxyacetyl group was used to protect NH_2 of exocyclic N^2 and this protecting group allows us to use the deprotecting condition (5% diisopropylamine solution in methanol), retaining the integrity of the acetyl group at N^8 [60].

dG-AAF was incorporated into oligonulceotides using the building block shown in scheme 3 under ultra mild solid phase DNA synthesis conditions. Figure 5 shows the reaction cycle of the DNA synthesizer. In this cycle, there are 5 steps, detritylation to reveal the 5' OH group as a substrate for the coupling step with the phosphoramidite to add one nucleotide, capping to prevent chain elongation of any uncoupled nucleotides, oxidation to convert to the pentavalent protected phophodiester group, and deprotection to yield the final oligonucleotide products.

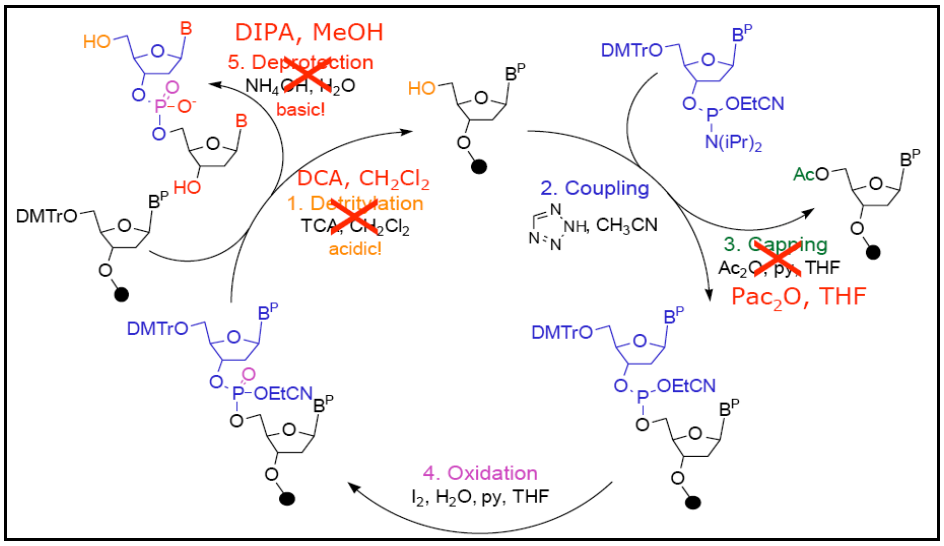


Figure 5: Modified DNA synthesis of AAF oligos. DCA: dichloro acetic acid, Pac_2O: phenoxyacetic anhydride, and DIPA: diisopropyl amine are used for an ultra mild condition. TCA: trichloro acetic acid, Ac_2O: acetic anhydride and NH_4OH are used for a general condition and these reagents are too acidic or basic for ultra mild conditions.

Table 1 below shows the sequences of the oligonucleotides that were prepared for further investigations. All oligonucleotides were purified by reverse phase HPLC and characterized by mass spectrometry (See supplementary data).

Name	Sequence
12 *Nar*I-1	5'-TACC**X**GCGCCAC-3'
12 *Nar*I-2	5'-TACCG**X**CGCCAC-3'
12 *Nar*I-3	5'-TACCGGC**X**CCAC-3'
24 *Nar*I-1	5'-TAGCTATTACC**X**GCGCCACATGTC-3'
24 *Nar*I-2	5'-GCTATTACCG**X**CGCCACATGTCAG-3'
24 *Nar*I-3	5'-CTATTACCGGC**X**CCACATGTCAGC-3'
44*Nar*I-1	5'-CCCTAGCTAGAGCTACGTAGCTATTACC**X**GCGCCACATGTCAGC-3'
44*Nar*I-2	5'-CCCTAGCTAGAGCTACGTAGCTATTACCG**X**CGCCACATGTCAGC-3'
44*Nar*I-3	5'-CCCTAGCTAGAGCTACGTAGCTATTACCGGC**X**CCACATGTCAGC-3'

Table 1. Sequences of the oligonucleotides used in this study. The 12-mer oligonucleotides used for thermodynamic stability assays. The 24-mer oligonucleotides used for the preparation of plasmids for NER assays. The 44-mer oligonucleotides were used for binding assays with XPC-RAD23B. X= either acetylaminofluorene or aminofluorene (AAF/AF placed on C8 of a deoxyguanosine).

Preparation of Plasmids and In Vitro NER Assay

The 24mer oligonucleotides (Table 1) were incorporated into plasmids for NER assays. The 24mers were annealed to a complementary single-stranded 3kb plasmid that was prepared using R408 helper phage from a modified pBluescirpt vector (Figure 6). Primer extension and ligation were performed using T4 DNA polymerase and T4 DNA ligase to yield closed circular dsDNA. The plasmid was purified using two additional purification steps; cesium chloride/ethidium bromide density gradient centrifugation and sucrose gradient centrifugation. These substrates were incubated with NER-proficient HeLa whole cell extract and the NER products, 24-32 nucleotide-long fragments containing the lesion, were annealed to a complementary strand containing an overhang of 4G that can be filled in with [α-^{32}P]dCTP to radioactively label the products. The excised NER fragments were run on a 14% sequencing gel and detected by a phosphor imaging [62].

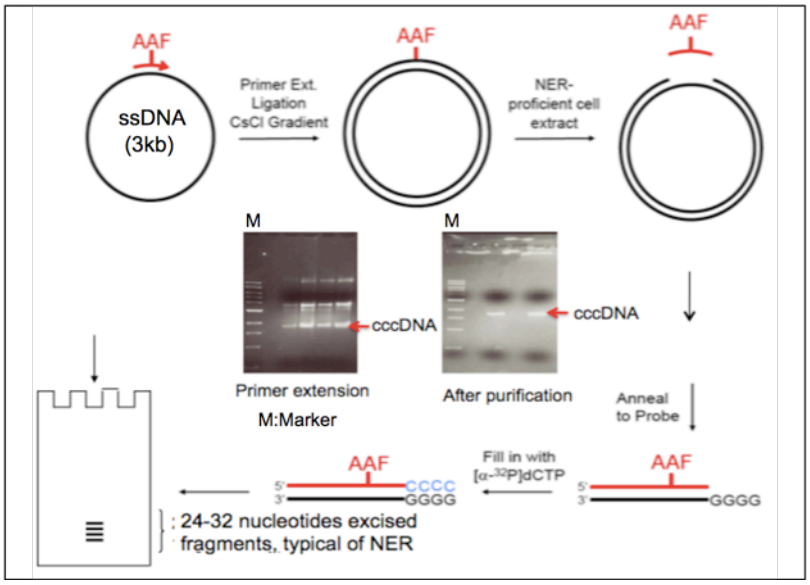


Figure 6: A scheme for primer extension and NER assay. A plasmid for NER assay was prepared by annealing of a 5' phosphorylated oligonucleotide containing a site-specific lesion to a ssDNA. Primer extension was performed with T4 DNA polymerase and ligase. CsCl gradient purification was achieved to purify the ccc plasmid. A purified plasmid containing a site-specific AAF was incubated with HeLa cell extract. The excised DNA fragments of 24-32 oligonucleotides were detected by annealing a complementary oligonucleotide containing a non-complementary 4G overhang and filling in with [α-^{32}P]dCTP. AAF: 2-acetylaminofluorene. 1% agarose gels show cccDNA (3kb) prepared by this procedure, indicated in the gels. cccDNA: covalently closed-circlular DNA.

Different Degrees of DNA Distortion Influence NER Efficiency of dG-AF or AAF Substrates in NarI Sequence

All substrates with either AAF or AF in the *Nar*I sequence (substrates 1-6 in Figure 7) were incubated with a HeLa cell-free extract and NER products detected by annealing the product to an oligonucleotide with a 4G overhang and a fill in reaction with polymerase and ^{32}P-dCTP. The level of product from incubation of substrates (1-3 in Figure 7) with AF on G1, G2, and G3 positions was lower than that resulting from incubation of the AAF substrates. The relative intensities of repair averaged from three independent experiments were 100% for 3-AAF, 60% for 2-AAF and 40% for 1-AAF, whlie the relative intensities for the AF substrates was 20% (3-AF), 7% (1-AF) and 2% (2-AF). In line with our observations it is known that dG-AAF creates a more severe distortion of DNA than dG-AF [40, 42-44]. The degree of DNA duplex destabilization is believed to be correlated with the binding affinity of XPC-RAD23B for damaged sites [12] and hence the recruitment of other NER factors and overall NER efficiency. An intriguing result is that the repair efficiencies of substrates 1-6 not only different between AAF and AF, but also within the sequence context, consistent with findings that the structural features of AAF and AF adducts are not only depended on the nature of the lesion, but also the sequence context.

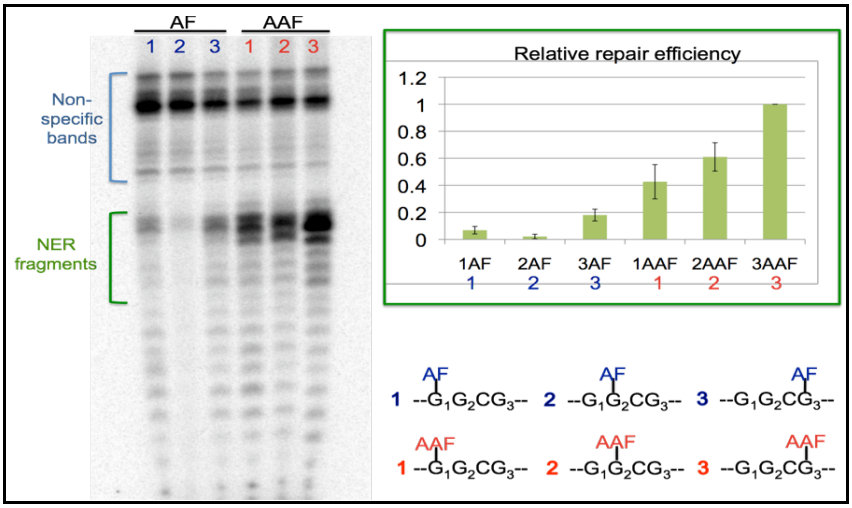


Figure 7: A NER sequencing gel and relative repair efficiency (14% denaturing PAGE). Plasmids containing site-specific dG-AF or dG-AAF residues were incubated with cell extracts prepared from HeLa cells. The 24mer to 32mer excision products containing dG-AF or dG-AAF were detected by annealing to a complementary oligonucleotide with a 5'-GpGpGpG overhang, which served as a template for end-labeling with [α-^{32}P]dCTP with sequenase. The reaction products were resolved on a 14% denaturing polyacrylamide gel. A low molecular weight DNA ladder from NEB was used as a size marker and the position of the 25 and 34nt bands are indicated. Relative repair efficiency was quantified by ImageQuant program.

The Binding Efficiency of XPC-RAD23B Correlates with Repair Efficiency of AAF and AF Adducts

Consistent with the influence of structure on relative repair efficiencies, structural and biochemical studies have indicated that thermodynamic destabilization induced by a lesion is a significant factor for the recognition by XPC-RAD23B [12]. To investigate how the various structures of the dG-AF and dG-AAF substrates influence binding by XPC-RAD23B, we used electrophoretic mobility shift assays (EMSA). 44mers oligodeoxynucleotides containing either a central AF or AAF lesion (Table 1) were annealed to a complementary strand fluorescently labeled with Cy5 at the 5' end and incubated with increasing amounts of XPC-RAD23B. Bound and unbound fractions were separated on a native 5% polyacrylamide gel (PAGE) (Figure 8). Comparison of the binding patterns revealed that the 3AAF substrate binds XPC-RAD23B protein at the lowest concentration followed by the 1-AAF and 2-AAF (Figure 9). AF substrates generally showed a lower affinity for XPC-HR23B, with 3AF binding more tightly than 2-AF and 1-AF (Figure 9). To quantify the EMSA data, the experiments were carried out in triplicate and intensities of the bands for free and XPC-RAD23B-bound 44mer oligonucleotides were quantified (Table 2). The K_d value for each sample was calculated using the Hill equation from Sigmaplot program, $y=[L]^n / (Kd+[L]^n)$, (Y= % of bound substrate; L: concentration (nM) of XPC-RAD23B; n: Hill coefficient, describing cooperativity). K_d values of *Nar*I-1AF and *Nar*I-2AF could not be accurately calculated since the binding efficiencies do not reach the saturation at the highest protein

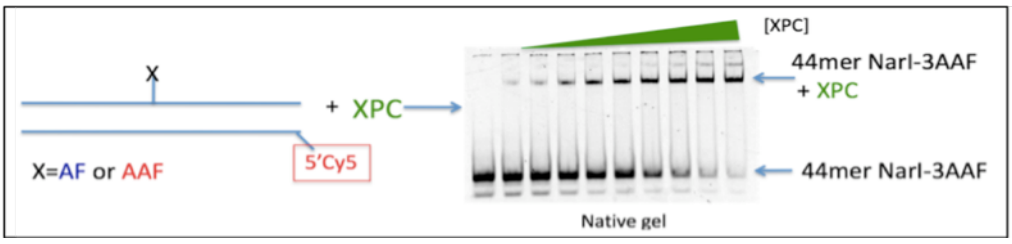


Figure 8: A Scheme for EMSA with XPC-HR23B. 44mer oligonucleotides containing a site-specific lesion were annealed to complementary oligonucleotides containing 5' Cy5. The annealed substrates were incubated with different concentrations (0-150nM) of XPC/RAD23B for 30min in a binding buffer. The reactions in mixtures were loaded on to a 5% native polyacrylamide gel. Shifted bands (DNA+XPC-RAD23B) and non-shifted bands (DNA) were quantified using imagequant TL v2005 program from GE healthcare life sciences.

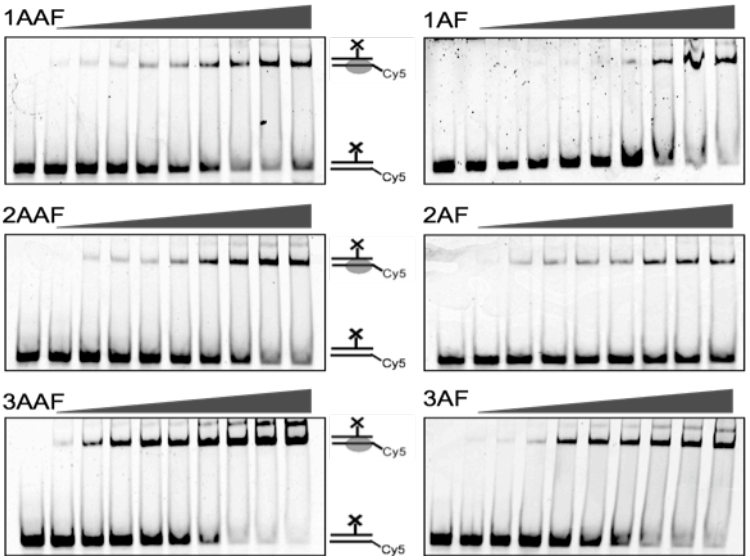
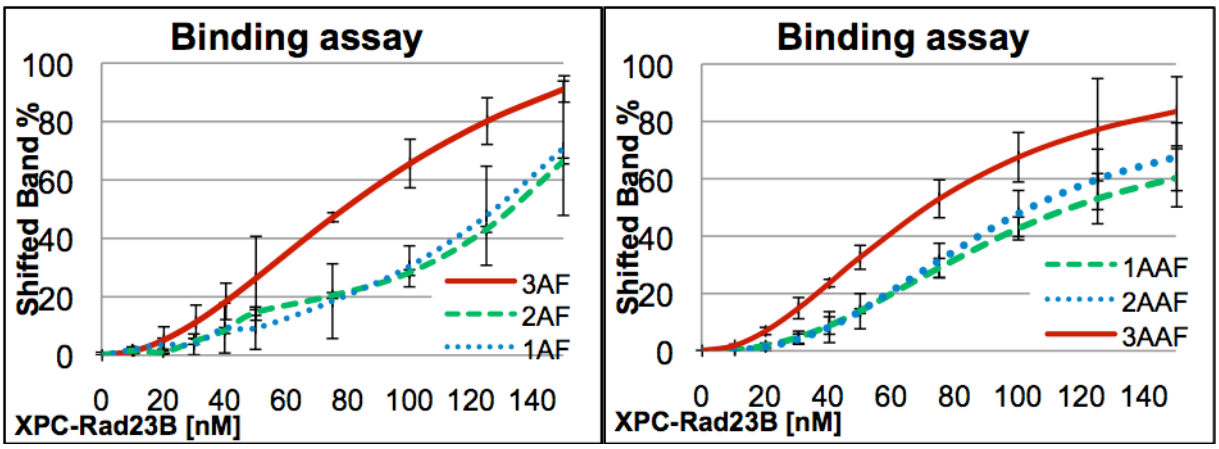


Figure 9: EMSA results with various dG-AAF and dG-AF substrates. DNA binding was analyzed using electrophoretic mobility shift assays. 44mer oligonucleotides containing a site-specific lesion were annealed to complementary oligonucleotides containing 5' Cy5. The annealed substrates were incubated with different concentrations ((left to right): 0, 10, 20, 30, 40, 50, 75, 100, 125, 150nM) of XPC/RAD23B for 30min in a binding buffer containing 25mM Tris-HCl (pH 7.5), 0.1mg/mL BSA, and 10% glycerol. The reactions in mixtures were loaded on to a 5% native polyacrylamide gel. Shifted bands (DNA+XPC/RAD23B) and non-shifted bands (DNA) were quantified using imagequant TL v2005 program from GE healthcare life sciences.

concentrations used. The *Narl*-3 AAF and *Narl*-3 AF substrates have the lowest K_d values. The low K_d value for *Nar*I-3AAF correlates well with the results indicating that *Nar*I-3AAF is the best NER substrate (Figure 7). The K_d value for *Nar*I-3AF substrate is 67±7nM which was low compared to the other AF and the AAF-2 and AAF-3 sequences, considering its moderate repair rate by NER. This substrate is therefore the only one that does not fit the overall trend that higher repair efficiency correlates with higher binding affinity to XPC-RAD23B. Interestingly, in the case of *Nar*I-3AF it has been shown that AF adducts at this position is consistent with the observation that AF on G3 position behaves like AAF [43], explaining why it has a similar binding affinity with XPC-RAD-23B as 3AAF.



Substrate	1AF	2AF	3AF	1AAF	2AAF	3AAF
K_d (nM)	N/A from Sigmaplot	N/A from Sigmaplot	102 ± 5	95 ± 3	91 ± 5	72 ± 6

Table 2: EMSA assay graphs and calculated dissociation constants (K_d). Bound (DNA+XPC/RAD23B) and unbound (DNA) from figure 9 were quantified using imagequant TL v2005 program from GE healthcare life sciences. K_d values were calculated from Sigma Plot program and graphs were modified by Sigma Plot program. Graphs show the quantification of three independent experiments. Standard deviations are indicated by error bars.

Thermodynamic Stabilities from Tm Values Correlated with Binding and NER Efficiency

The EMSA provided information about the binding affinity of XPC-RAD23B to modified DNA and we wished to correlate the binding and NER efficiencies with the thermodynamic stabilities of the 6 different substrates. We therefore measured the melting temperatures of 12mer AAF and AF oligos to study their thermodynamic stabilities. Two 12mer oligos, one strand containing modification (either AF or AAF on *Nar*I sequence) and the complementary strand, were incubated together in UV-VIS spectrophotometer measuring the UV absorbance while heating the sample at a rate of 0.2°C/min. Denaturation (single stranded DNA) of the duplex results in an increase of the UV absorbance and the 50% point of the total increase is where ssDNA and dsDNA are present in a 1:1 ratio and is defined as the melting temperature (Figure 10).

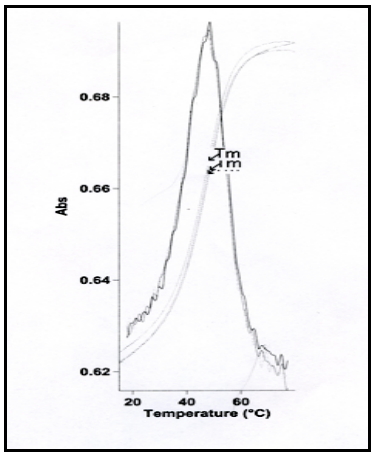


Figure 10: A diagram to measure melting temperatures. Y axis represents UV absorbance and X axis represents temperature. Tm is the corresponding point where 50% of ssDNA and 50% of dsDNA are present.

Based on thermodynamic stability experiments, the presence of either AF or AAF decreases thermal and thermodynamic stability of DNA duplex (Table 3). The melting temperature of the control oligonucleotide is between 10 and 18 °C higher than that of AF or AAF-containing DNA duplexes. *Nar*I-3 AAF had the lowest melting temperature (around 48.18°C in Table 3) and the highest ΔG (Gibbs free energy, -9.77) value, indicating that this substrate has the lowest stability among the 6 duplexes, correlating well to the fact it is the best NER substrate. The rank order (high to low) of ΔG of 6 different substrates is: NarI-3AAF>NarI-2AAF>NarI-1AAF>NarI-3AF>NarI-2AF≈NarI-1AF>control (Table 3). These thermal and thermodynamic stability data therefore correlate well with the efficiency of XPC-RAD23B binding and overall NER activity.

Substrates	Tm (°C)[a]	ΔH (kcalmol^{-1})	ΔS (calmol^{-1}k^{-1})	ΔG (kcalmol^{-1})[b]
Control	66±1	-102±0.7	-278±1	-16±0.4
*Nar*I-1 AF	55±0.4	-110±0.3	-310±0.4	-13±0.1
*Nar*I-2 AF	56±0.4	-105±0.3	-294±0.8	-14±0.1
*Nar*I-3 AF	51±1	-123±0.7	-354±1	-13±0.4
*Nar*I-1 AAF	54 ±1	-91±0.6	-254±0.9	-12±0.3
*Nar*I-2 AAF	50±0.9	-113±0.6	-325±1	-12±0.3
*Nar*I-3 AAF	48±1	-67±2	-178±0.7	-10±0.2

Table 3: *T*m values and thermodynamic stabilities. a: melting temperatures were measured at 5.8×10^{-6} M. b: Gibbs free energy at 37°C, Reported errors are standard deviations.

DISCUSSION

To address the question how the different degrees of duplex distortion and thermodynamic destabilization influence NER, we synthesized oligonucleotides containing dG-AF and dG-AAF lesions in the *Nar*I sequence, where AF and AAF adducts adopt distinctive structural conformations and used them to study various NER parameters. In assays of overall NER efficiency, we observed consistent differences in repair efficiencies. A plasmid containing *Nar*I-3 AAF showed higher repair efficiency than *Nar*I-2 and *Nar*I-3 AAF. It is interesting that these substrates contain the same modification, AAF, but placed at a different position in the same sequence results in different efficiencies of repair. AF substrates showed lower repair efficiency than AAF substrates. In order to investigate whether these differences can be traced to the lesion recognition step, we performed binding assays with XPC-RAD23B and found that NarI-3 AAF substrates were also recognized by XPC-RAD23B with the highest affinity. This result supports the idea that the efficiency of initial damage recognition correlates well with the overall NER efficiency and likely represents the rate-limiting step in NER. We also determined how thermodynamic destabilization induced by a lesion, influenced the repair process by measuring melting temperatures of 12mer oligos containing the AAF and AF lesions in the same sequence context. This study showed that NarI-3 AAF substrate is the least stable substrate among the ones we studied. Our study is therefore consistent with the idea that several factors that affect helix stability, including base-pairing disruption, bending, and flexibility need to be considered for damage recognition in NER and also leads to support the idea that thermodynamic stabilization influenced by all parameters that determine the efficiency of the NER process [53, 54].

MATERIALS AND METHODS

Reagents and equipments

Chemicals and solvents were purchased from Fluka-Sigma-Aldrich. Reagents for DNA synthesis were purchased from Glen Research. The “ultra-mild” phosphoramidites were available from Glen Research and the 1000 Å “Q-columns” from Biosearch Technologies. DNA syntheses were performed on a PerSeptive Biosystems Expedite 8900 DNA synthesizer. HPLC purifications were performed on a JASCO system equipped with a Phenomenex Clarity Oligo-RP Semi-prep column: C18, 5µm, 50×10.00 mm. The C18-SepPak cartridges were from Waters. T4 PNK, T4 polymerase, and T4 DNA ligase were from New England Biolabs. The sequenase enzyme (T7 DNA polymerase) was purchased from USB and [α-^{32}P] dCTP from PerkinElmer. Some oligonucleotides were purchased from Integrated DNA Technologies (IDT). NMR spectra were recorded on a Varian 300MHz. The 12-mer and 24-mer oligonucleotides were analyzed by Agilent 1100 ESI and 44 mer nucleotides were analyzed by MALDI.

The synthetic method of a C8-dG-AAF adduct was previously reported to a published procedure [39] and all NMR and mass spectrum data were confirmed by pulished data [39].

8-bromo-2'-deoxyguanosine (Scheme 2a)

To a suspension of 2'-deoxyguanosine (2.0g, 7.5mmol) in a mixture of acetonitrile (80ml) and water (20ml) was added *N*-bromosuccinimide (2.0g, 11.2mmol) slowly. The reaction mixture was stirred for 30' at room temperature. The solvent was removed and the precipitate suspended in acetone (40ml), stirred for 2 hours at room temperature and cooled overnight at -20°C. The precipitate was collected by filtration, extensively washed with cold acetone and dried under vacuum to give 2.02g (5.83mmol, 78%) of a slightly yellow powder.

^{1}H NMR(DMSO-d_6) δ (ppm) 10.78 (s, 1H, N_1-H), 6.74 (s, 2H, N^2-H_2), 6.16 (dd, J=7.3, 7.3, 1H, $C_{1'}$-H), 4.40 (ddd, J=3.0, 3.1, 6.0, 1H, $C_{3'}$-H), 3.80 (ddd, J=3.1, 5.4, 5.5, 1H, $C_{4'}$-

H), 3.63 (dd, J=5.3, 11.6, 1H, $C_{5'}$-H), 3.50 (dd, J=5.8, 11.6, 1H, $C_{5'}$-H), 3.16 (ddd, J=6.6, 7.6, 13.7, 1H, $C_{2'}$-H), 2.10 (ddd, J=3.0, 6.5, 13.3, 1H, $C_{2'}$-H)

8-bromo-3',5'-O-bis(tert-butyldimethylsilyl)-2'-deoxyguanosine (Scheme 2b)

8-bromo-2'-deoxyguanosine (1.97g, 5.54mmol) and imidazole (1.72g, 25.0mmol) were suspended in DMF (10ml) and treated with a 1M solution of tert-butyldimethylsilyl chloride in THF (15.24ml, 15.24mmol). The reaction mixture was stirred for 5hours at room temperature, concentrated, diluted in water (30ml) and extracted with EtOAc (3×20ml). The organic layer was dried over Na_2SO_4, concentrated and purified by chromatography on silica gel to give 2.76g (4.8mmol, 87%) of a white powder.
^{1}H NMR(DMSO-d_6) δ (ppm) 10.78 (s, 1H, N_1-H), 6.38 (s, 2H, N^2-H_2), 6.15 (dd, J=7.0, 7.1, 1H, $C_{1'}$-H), 4.59 (ddd, J=3.2, 3.3, 6.2, 1H, $C_{3'}$-H), 3.76 (m, 2H, $C_{5'}$-H), 3.65 (m, 1H, $C_{4'}$-H), 3.38 (ddd, J=6.5, 6.6, 13.4 1H, $C_{2'}$-H), 2.15 (ddd, J=3.6, 7.0, 13.2, 1H, $C_{2'}$-H), 0.89 (s, 9H, tBu-H), 0.83 (s, 9H, tBu-H), 0.11 (s, 6H, 2×CH_3), -0.01 (s, 3H, CH_3), -0.023 (s, 3H, CH_3)

O^6-benzyl-8-bromo-3',5'-O-bis(tert-butyldimethylsilyl)-2'-deoxyguanosine (Scheme 2c)

To a suspension of 8-bromo-3',5'-*O*-bis(*tert*-butyldimethylsilyl)-2'-deoxyguanosine (0.58g, 1.01mmol) and triphenylphosphine (0.4g, 1.5mmol) in dioxane (15ml) were added benzylalcohol (0.16ml, 1.5mmol) and diisopropyl azodicarboxylate (0.31ml, 1.5mmol). The reaction mixture was stirred for 5hours at room temperature. After concentration, the crude product was purified by chromatography on silica gel to provide 0.52g (0.78mmol, 78%) of a white foam.
^{1}H NMR(DMSO-d_6) δ (ppm) 7.45-7.51 (m, 2H, Bz-H), 7.30-7.43 (m, 3H, Bz-H), 6.43 (s, 2H, N^2-H_2), 6.19 (dd, J=6.8, 6.9, 1H, $C_{1'}$-H) 5.48 (AB quart, J=12.3, 2H, Bz-CH_2), 4.68 (ddd, J=3.7, 3.8, 6.5, $C_{3'}$-H), 3.77 (m, 2H, $C_{5'}$-H), 3.63 (ddd, J=5.2, 7.6, 7.8, 1H, $C_{4'}$-H) 3.52 (ddd, J=6.5, 6.5, 13.1, 1H, $C_{2'}$-H), 2.19 (ddd, J=4.0, 7.0, 13.2, 1H, $C_{2'}$-H), 0.90 (s, 9H, tBu-H), 0.80 (s, 9H, tBu-H), 0.13 (s, 6H, 2×CH_3), -0.04 (s, 3H, CH_3), -0.06 (s, 3H, CH_3)

O^6-benzyl-8-bromo-3',5'-O-bis(tert-butyldimethylsilyl)-N^2-dimethoxytrityl-2'-deoxyguanosine (Scheme 2d)

To a solution of O^6-benzyl-8-bromo-3',5'-*O*-bis(*tert*-butyldimethylsilyl)-2'-deoxyguanosine(1.09g, 1.64mmol) in pyridine (15ml) was added 4,4'-dimethoxytrityl chloride (0.8g, 2.46mmol) and the mixture was stirred overnight at room temperature. The solvent was evaporated and the residue purified by chromatography on aluminum oxide to provide 1.4g (1.47mmol, 90%) of the product as a white foam.

^{1}H NMR(DMSO-d_6) δ (ppm) 7.06-7.36 (m, 15H, Bz-H_5+DMTr-H_9+N^2-H), 6.78-6.84 (m, 4H, DMTr-H_4), 6.07 (m, 1H, $C_{1'}$-H), 4.85 (br s, 2H, Bz-CH_2), 4.63 (m, 1H, $C_{3'}$-H), 3.65-3.75 (m, 8H, 2×OMe+ $C_{5'}$-H_2), 3.56(m, 2H, $C_{4'}$-H+ $C_{2'}$-H), 2.03 (m, 1H, $C_{2'}$-H), 0.87 (s, 9H, tBu-H), 0.77 (s, 9H, tBu-H), 0.10 (s, 6H, 2×CH_3), -0.05 (s, 3H, CH_3), -0.10 (s, 3H, CH_3)

8-(N-2-aminofluorene)-O^6-benzyl-3',5'-O-bis(tert-butyldimethylsilyl)-N^2-dimethoxytrityl-2'-deoxyguanosine (Scheme 2e)

O^6-benzyl-8-bromo-3',5'-*O*-bis(*tert*-butyldimethylsilyl)-N^2-dimethoxytrityl-2'-deoxyguanosine (0.72g, 0.75mmol), 2-aminofluorene (0.21g, 1.13mmol), tris(dibenzylidenaceton)-dipalladium ($Pd_2(dba)_3$, 42mg, 39.3μmol) and rac-2,2'-Bis(diphenylphosphino)-1,1'-binaphtyl (BINAP, 77mg, 0.11mmol) were preheated in dried toluene (8ml) at 90°C. After 40', NaOtBu (0.12g, 1.13mmol) was added and the reaction mixture was stirred 1 additional hour at 100°C. The reaction was cooled, diluted with diethylether and filtered. The filtrate was concentrated and purified by chromatography on aluminium oxide (Al_2O_3) to provide 0.54g of the product as a slightly yellow powder.

^{1}H NMR(DMSO-d_6) δ (ppm) 8.85 (s, 1H, N^8-H), 7.85 (m, 1H, AF-C_1-H), 7.72-7.77 (m, 2H, AF-H_2), 7.59 (m, 1H, AF-H), 7.50 (m, 1H, AF-H), 7.12-7.35 (m, 16H, AF-H_2+Bz-H_5+DMTr-H_9), 6.81 (m, 4H, DMTr-H_4), 6.53 (s, 1H, N^2-H), 6.22 (dd, J=6.7, 6.7, 1H, $C_{1'}$-H), 4.88 (br s, 2H, Bz-CH_2), 4.60 (m, 1H, $C_{3'}$-H), 3.76-3.88 (m, 4H, AF-C_9-H+ $C_{5'}$-H_2),

3.71 (s, 6H, 2×OMe), 3.64(m, 1H, $C_{4'}$-H), 3.42 (m, 1H, $C_{2'}$-H), 2.02 (m, 1H, $C_{2'}$-H), 0.88 (s, 9H, tBu-H), 0.80 (s, 9H, tBu-H), 0.10 (s, 6H, 2×CH_3), -0.03 (s, 3H, CH_3), -0.05 (s, 3H, CH_3)

8-(N-acetyl-2-aminofluorene)-O^6-benzyl-3',5'-O-bis(tert-butyldimethylsilyl)- 2'-deoxyguanosine (Scheme 2f, g)

To a solution of 8-(*N*-2-aminofluorene)-O^6-benzyl-3',5'-*O*-bis(*tert*-butyldimethylsilyl)-N^2-dimethoxytrityl-2'-deoxyguanosine (0.36g, 0.34mmol) in pyridine (5ml) were added triethylamine (0.23ml, 1.67mmol), 4-dimethylaminopyridine (21mg, 0.17mmol) and acetic anhydride (0.16ml, 1.68mmol). After stirring 5hours at room temperature, the reaction mixture was evaporated, diluted in EtOAc (50ml) and exacted with water (3×50ml). The organic layer was dried over Na_2SO_4, evaporated to dryness and treated with a 0.01M solution of HCl in MeOH (1.67ml, 16.7μmol). The mixture was stirred overnight at room temperature, concentrated and purified by chromatography on silica gel to provide 0.22g (0.22mmol, 67% overall) of the product as a slightly yellow powder.

^{1}H NMR(DMSO-d_6) δ (ppm) 7.84-7.94 (m, 2H, AF- H), 7.46-7.60 (m, 4H, AF-H), 7.26-7.44 (m, 6H, Bz-H+AF-H), 6.41 (s, 2H, N^2-H_2), 6.10 (m, 1H, $C_{1'}$-H), 5.50 (br s, 2H, Bz-CH_2), 4.48 (m, 1H, $C_{3'}$-H), 3.62-3.94 (m, 5H, AF-C_9H_2+ $C_{4'}$-H+$C_{5'}$-H_2), 3.25 (m, 1H, $C_{2'}$-H), 2.07 (s, 3H, N^8-Ac), 1.99 (m, 1H, $C_{2'}$-H), 0.86 (s, 9H, tBu-H), 0.76 (s, 9H, tBu-H), 0.08 (s, 6H, 2×CH_3), -0.07 (s, 6H, 2×CH_3)

8-(N-acetyl-2-aminofluorene)-O^6-benzyl-3',5'-O-bis(tert-butyldimethylsilyl)- N^2-isopropylphenoxyacetyl-2'-deoxyguanosine (Scheme 2h)

To 8-(*N*-acetyl-2-aminofluorene)-O^6-benzyl-3',5'-*O*-bis(*tert*–butyldimethylsilyl)-2'-deoxyguanosine (0.22g, 0.27mmol) dissolved in pyridine (6ml) was added isopropylphenoxyacetyl chloride[1] (0.24g, 1.11mmol). After stirring overnight at room temperature, the reaction mixture was concentrated, diluted in EtOAc (50ml) and extracted with water (2× 50ml). The organic layer was dried over Na_2SO_4, concentrated and purified by chromatography on silica gel to provide 0.23g (0.23mmol, 86%) of the product as a slightly yellow powder.

^{1}H NMR(DMSO-d_6) δ (ppm) 10.55 (s, 1H, N^2-H), 7.90 (m, 2H, AF- H), 7.50-7.70 (m, 4H, AF-H+Bz-H), 7.26-7.48 (m, 6H, AF-H + Bz-H), 7.12 (d, J=8.6, 2H, Pac-H_2), 6.85 (d, J=8.6, 2H, Pac-H_2), 6.25 (m, 1H, $C_{1'}$-H), 5.61 (br s, 2H, Bz-CH_2), 4.87 (s, 3H, Pac-CH_2+$C_{3'}$-H), 3.91 (s, 2H, AF-C_9-H_2), 3.68-3.86 (m, 3H, $C_{4'}$-H+ $C_{5'}$-H_2), 3.21 (m, 1H, $C_{2'}$-H),2.81 (sept, J=6.8, 1H, iPr-H), 2.08 (s, 3H, N^8-Ac), 1.90 (m, 1H, $C_{2'}$-H), 1.16 (s, 3H, iPr-CH_3), 1.14 (s, 3H, iPr-CH_3), 0.83 (s, 9H, tBu-H), 0.72 (s, 9H, tBu-H), 0.024 (s, 3H, CH_3), 0.00 (s, 3H, CH_3) -0.14 (m, 6H, 2×CH_3)

8-(N-acetyl-2-aminofluorene)-O^6-benzyl-N^2-isopropylphenoxyacetyl-2'-deoxyguanosine (Scheme 2i)

8-(*N*-acetyl-2-aminofluorene)-O^6-benzyl-3',5'-*O*-bis(*tert*-butyldimethylsilyl)-N^2-isopropylphenoxyacetyl-2'-deoxyguanosine (0.44g, 0.44mmol) in THF (10ml) at 0°C were added acetic acid (0.15ml, 2.68mmol) and a 1M solution of tetrabutylamonium fluoride in THF (1.28ml, 1.28mmol). The reaction mixture was allowed to reach room temperature and stirred overnight. After concentration, the residue was purified by chromatography on silica gel to provide 0.3g (0.4mmol, 88%) of the product as a white powder.

^{1}H NMR(DMSO-d_6) δ (ppm) 10.58 (s, 1H, N^2-H), 7.97 (d, J=7.8, 1H, AF- H), 7.91 (d, J=7.2, 1H, AF-H), 7.67 (m, 1H, AF-H), 7.59 (d, J=7.2, 1H, AF-H), 7.49-7.56 (m, 2H, AF-H), 7.30-7.43 (m, 5H, Bz-H), 7.15 (d, J=8.6, 2H, Pac-H_2), 6.86 (d, J=8.7, 2H, Pac-H_2), 6.25 (dd, J=6.9, 6.9, 1H, $C_{1'}$-H), 5.60 (br s, 2H, Bz-CH_2), 5.20 (m, 1H, $C_{3'}$-OH), 4.95 (s, 2H, Pac-CH_2), 4.85 (m, 2H, $C_{5'}$-OH+ $C_{3'}$-H), 3.95 (s, 2H, AF-C_9-H), 3.48 (m, 1H, $C_{4'}$-H), 3.69 (m, 1H, $C_{5'}$-H), 3.55 (m, 1H, $C_{5'}$-H), 3.17 (m, 1H, $C_{2'}$-H) 2.82 (sept, J=6.9, 1H, iPr-H), 2.10 (s, 3H, N^8-Ac), 2.08 (m, 1H, $C_{2'}$-H), 1.17 (s, 3H, iPr-CH_3), 1.15 (s, 3H, iPr-CH_3)

8-(N-acetyl-2-aminofluorene)-N^2–isopropylphenoxyacetyl-2' deoxyguanosine (Scheme 2j)

To 8-(*N*-acetyl-2-aminofluorene)-O^6-benzyl-2'-deoxyguanosine (0.3g, 0.4mmol) dissolved in EtOH (15ml) were added a 10% powder of Pd/C (36mg) and cyclohexene (910μl). The reaction mixture was refluxed for 4 hours, cooled and filtered. The filtrate

was concentrated and purified by chromatography on silica gel to provide 0.23g (0.34mmol, 87%) of the product as a slightly yellow powder.

^{1}H NMR(DMSO-d_6) δ (ppm) 11.91 (s, 1H, N_1-H), 11.47 (s, 1H, N^2-H), 7.97 (d, J=7.8, 1H, AF- H), 7.92 (d, J=7.2, 1H, AF-H), 7.69 (m, 1H, AF-H), 7.59 (d, J=7.2, 1H, AF-H), 7.49 (m, 1H, AF-H), 7.41 (dd, J=1.0, 6.3, 1H, AF-H), 7.36 (dd, J=1.3, 4.9, 1H, AF-H), 7.32 (dd, J=1.2, 7.3, 1H, AF-H), 7.17 (d, J=8.6, 2H, Pac-H_2), 6.90 (d, J=8.7, 2H, Pac-H_2), 6.22 (dd, J=6.9, 7.0, 1H, $C_{1'}$-H), 5.24 (d, J=4.6, 1H, $C_{3'}$-OH), 4.86 (s, 2H, Pac-CH_2), 4.71 (m, 1H, $C_{5'}$-OH), 4.48 (m, 1H, $C_{3'}$-OH), 3.96 (s, 2H, AF-C_9-H), 3.86 (m, 1H, $C_{4'}$-H), 3.67 (m, 1H, $C_{5'}$-H), 3.54 (m, 1H, $C_{5'}$-H), 3.08 (m, 1H, $C_{2'}$-H) 2.83 (sept, J=6.9, 1H, iPr-H), 1.99-2.09 (m, 4H, N-Ac+ $C_{2'}$-H), 1.18 (s, 3H, iPr-CH_3), 1.15 (s, 3H, iPr-CH_3)

(5'-O-dimethoxytrityl)-N-[N^2-isopropylphenoxyacetyl]-deoxyguanosin-8-yl]-2-acetylaminofluorene (Scheme 3k)

To 8-(N-acetyl-2-amino-fluorene) - N^2-isopropylphenoxyacetyl -2'-deoxyguanosine (170mg, 0.26mmol) dissolved in pyridine (2ml) was added 4, 4-dimethoxytrityl chloride 112mg, 0.33mmol). The reaction mixture was stirred for 2 hours at room temperature, concentrated and purified by chromatography on silica gel to provide 185mg (0.19mmol, 74%) of a slightly yellow powder.

^{1}H NMR(DMSO-d_6) δ (ppm) 11.76 (s, 1H, N_1-H), 11.30 (s, 1H, N^2-H), 7.84-7.95 (m, 2H, AF- H_2), 7.44-7.78 (m, 3H, AF-H_3), 7.25-7.41 (m, 4H, AF-H_2+DMTr-H_2), 7.05-7.21 (m, 9H, Pac-H_2+DMTr-H_7) 6.89 (d, J=8.7, 2H, Pac-H_2), 6.56-6.73 (m, 4H, DMTr-H_4), 6.32 (dd, J=6.9, 7.2, 1H, $C_{1'}$-H), 5.19 (d, J=5.3, 1H, $C_{3'}$-OH), 4.79 (s, 2H, Pac-CH_2), 4.61 (m, 1H, $C_{3'}$-OH), 4.07 (m, 1H, $C_{4'}$-H), 3.86 (m, 2H, AF-C_9-H), 3.66 (s, 3H, OMe), 3.64 (s, 3H, OMe), 3.39 (m, 1H, $C_{5'}$-H), 3.16 (m, 1H, $C_{5'}$-H), 2.99 (m, 1H, $C_{2'}$-H), 2.85 (sept, J=6.9, 1H, iPr-H), 2.30 (m, 1H, $C_{2'}$-H), 2.11 (br s, 3H, N-Ac), 1.18 (s, 3H, iPr-CH_3), 1.16 (s, 3H, iPr-CH_3)

[3'-O-(2''- cyanoethoxydiisopropylaminophosphino)]-(5'-O-dimethoxytrityl)-N-[(N^2-isopropylphenoxyacetyl)-deoxyguanosin-8-yl]-2-acetylaminofluorene (Scheme 3l)

To (5'-O-dimethoxytrityl)-N-[N^2-isopropylphenoxyacetyl)-deoxyguanosin-8-yl]-2-acetylaminofluorene in CH_2Cl_2 was added N-ethyldiisopropylamine and 2-cyanoethoxydiisopropylchloro-phosphoramidite. The reaction mixture was stirred for 2 hours at room temperature, concentrated to dryness and redissolved in the minimum amount of CH_2Cl_2, 40ml of hexane were then added under smooth stirring and the product was further precipitated for 2hours at -20°C. The clear supernatant was then removed and the residue was purified by chromatography on silica gel to provide the product of a slight yellow powder.

Preparation of Isopropylphenoxyacetyl acid was obtained from isopropylphenol according to a published procedure [63] and reacted with 1.2 equivalents of $SOCl_2$ to lead after distillation to isopropylphenoxyacetyl chloride in 75% yield.

XPC-HR23B expression and purification and HeLa whole cell extract

Polyhistidine-tagged RAD23B was expressed in the E. coli BL21(DE3)LysS using the expression vector pET-24d and purified on nickel beads (Qiagen). Polyhistidine and MBP-tagged hXPC was expressed in Sf9 cells using the expression vector pFastBac1. The cells were lysed as described [62] and S3 was combined with partially purified RAD23B. The correctly folded heterodimer was further purified through nickel beads (Qiagen), gel filteration (Pharmacia) and heparin (Amersham) columns. Expression and purification were performed by Adebanke Fagbemi. HeLa whole-cell extract was prepared as described [65].

Preparation of Oligodeoxynucleotides containing a site-specific AAF or AF

12-AAF, 24-AAF, and 44-AAF (Table 1) were synthesized based on modified ‘ultra-mild’ DNA synthesis according to a published procedure [60]. All DNA syntheses were performed on 1μM scale using 1000 Å “Q-columns”. The “ultra-mild” phosphoramidites (T, Ac-dC, Pac-dA, iPrPac-dG) were dissolved to 0.1M in CH_3CN. The iPrPac-dG-AAF phosphoramidite was dissolved in CH_2Cl_2 to give a 0.1M solution. The final 5’-DMTr protective group was retained for all the syntheses (‘DMT-on’ synthesis). After completion of oligonucleotide synthesis, the solid support was dried and incubated

overnight at 55°C with a solution containing 10% (V/V) of diisopropylamine (iPr_2NH) and 0.25M of β-mercaptoethanol in MeOH. The supernatant containing the oligonucleotide was decanted and concentrated. This concentrated oligonucleotide was dissolved in 1ml of 1M triethylamoniumacetate (TEAA) buffered at pH 7 and this solution was filtered through 0.45μm filter and purified by standard Reversed-Phase HPLC purification. The peak of 'DMT-ON' oligonucleotide was collected, concentrated, and treated with 80% acetic acid solution for 40 min at room temperature to remove DMT group. Oligonucleotides containing AF were prepared by treatment of the dG-AAF oligos with 1M NaOH as described in a published procedure [61, 66]. After purification, all oligonucleotides were desalted using Sep-Pak columns (Waters) according to the manufacturer's instructions.

12-AAF/AF and 24-AAF/AF were identified by ESI and 44-AAF/AF were identified by MALDI.

Preparation of single strand (ss) DNA for plasmid preparation (NarI sequence)

The preparation of ssDNA was carried out according to a published procedure [60]. 2ml of LB was inoculated with single colony of XL1Blue containing p98 (*Nar*I MOCK in pBlkp II SK+) in presence of ampicillin for 6h. After inoculation, 0.5 ml R408 helper phage was added and incubated for 10min at room temperature. The 2ml culture was then transferred to 600ml LB-Amp for overnight. The 600ml culture was centrifuged at 4000 rpm for 40 min. and the supernatant was transferred to the other centrifuge tube and this step was repeated. The clarified supernatant was treated with 200μl of 5mg/ml RNase A and 50μl of DNase I (10U/μl) and incubated at 37°C for 1h 30min. After incubation, the supernatant was treated 20% PEG (8000); 2.5M NH_4OAc and incubated at 4°C for overnight to precipitate the virus particles, centrifuged at 4000rpm for 40min at 4°C and the pellet dissolved in 10mL of 10mM tris-HCl (pH8). This aqueous solution was extracted with Phenol/chloroform several times until no more white precipitant was visible at the interface. Once additional extraction was performed with 10ml chloroform and the 3M NaOAc (pH4.8) (0.1 volume of sample) and 100% cold EtOH (2.2 or 2.5 eq. volume of sample) were added to the water layer. The mixed sample was incubated at -80°C for overnight, centrifuged at 4000 rpm for 40min at 4°C and the pellet was washed

with 70% EtOH (0.5 volume of sample) and the pellet air dried. The dried pellet was resuspended in 10mM tris-HCl (pH8) and the concentration and purity checked by Nano drop and an agarose gel.

Preparation of plasmids containing a site-specific AAF or AF (primer extenstion) and In vitro NER assay

100 pmol of the 24-AAF (or 24-AF) oligonucleotides was 5'-phosphorylated by incubation with 20 units of T4 PNK enzyme and 2mM of ATP at 37 °C for 2hours. After annealing with 30pmol of single-stranded p98 (*Nar*I), the annealed primer was incubated with dNTPs (800 μM), T4 DNA polymerase (90 units) and T4 DNA ligase (100 units) yielded to covalently closed circular DNA containing a single AAF (or AF) adduct. The closed circular DNA was purified by cesium chloride/ethidium bromide density gradient centrifugation, followed by consecutive butanol extractions to remove the ethidium bromide and finally concentrated on a Centricon YM-30 (Millipore). The collected covalently closed circular DNA was further purified by sucrose gradient centrifugation to remove the rest of ethidium bromide and traces of nicked DNA. The plasmids containing dG-AAF and dG-AF were aliquoted and stored at -80 °C. The in vitro NER assay was performed as described [62]. HeLa (2 μl) cell extract [91], 2 μl of 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) (final NaCl concentration was 70 mM) in a total volume of 10 μl were pre-incubated at 30°C for 10 min. 1 μL of a covalently-closed circular DNA plasmid (50ng/μl) containing either the AAF or the AF adduct was added before incubating the mixture at 30 °C for 45 min. After placing the samples on ice, 0.5 μl of 1 μM of a 3'-phophorylated oligonucleotide: d(GGGGCATGTGGCGCCGGTAATAGCTAC GTAGCTC) was added and the mixture heated at 95 °C for 5 min. The samples were allowed to cool down to room temperature for 15 min to allow the DNA to anneal. 1 μL of a Sequenase/[α-^{32}P]-dCTP mix (0.25 units of Sequenase and 2.5 μCi of [α-^{32}P]-dCTP per reaction) was added before incubating at 37 °C for 3 min, followed by addition of 1.2 μl of dNTP mix (100 μM of each dATP, dTTP, dGTP; 50 μM dCTP) and incubated for another 12 min. The

reactions were stopped by adding 12 µl of loading dye (80% formamide/10 mM EDTA) and heating at 95 °C for 5 min. The samples were run on a 14% sequencing gel (0.5x TBE) at 45 W for 2.5 hrs. The reactions products were visualized using a PhosphorImager (Typhoon 9400, Amersham Biosciences). NER assay was performed three times with each substrate.

EMSA to measure binding to XPC-RAD23B (Gel Shift Binding assay)

AAF/AF-modified 44mer templates (4nM) were annealed to the 5' Cy5 labeled primer (1.3nM) in the presence of 10mM Tris-HCl (pH8), containing 6.6mM NaCl and 0.66mM MgCl. The binding reactions were containing 25mM Tris-HCl (pH7.5), 0.1mg/mL BSA and 10% glycerol at 25 °C for 30min in a 15µL reaction containing 4nM non-modified annealed 44mer competitor, increasing amounts of XPC-RAD23B (0-150nM), and 300mM NaCl to adjust salt concentration to 40mM as final concentration. The reaction mixture were loaded onto a native 5% polyacrylamide gel pre-equilibrated with 0.5 X TBE buffer, which were run at 4 °C for 50 min at 20 mA. Gels were scanned using a Typhoon 9400 imager. The band intensities of free annealed 44mer oligonucleotides and the complex formed with XPC-RAD23B were determined by Image Quant TL program from Amersham Biosciences. EMSA with each substrate was performed three times to give average binding data. Sigmaplot software was used to calculate *Kd* values and to draw graphs. The eqilibrium dissociation constants (K_d) for the XPC-RAD23B and annealed 44mer oligonucleotides containing either AAF or AF and 44mer complementary strand were determined by Sigmaplot program with Hill equation, 1/Y=(Kd + L) / L, (Y= % of binding between a substrate and protein, L: concentration (nM) of XPC (protein), and n: Hill coefficient, describing cooperativity)

Measuring melting temperatures and thermodynamic stabilities

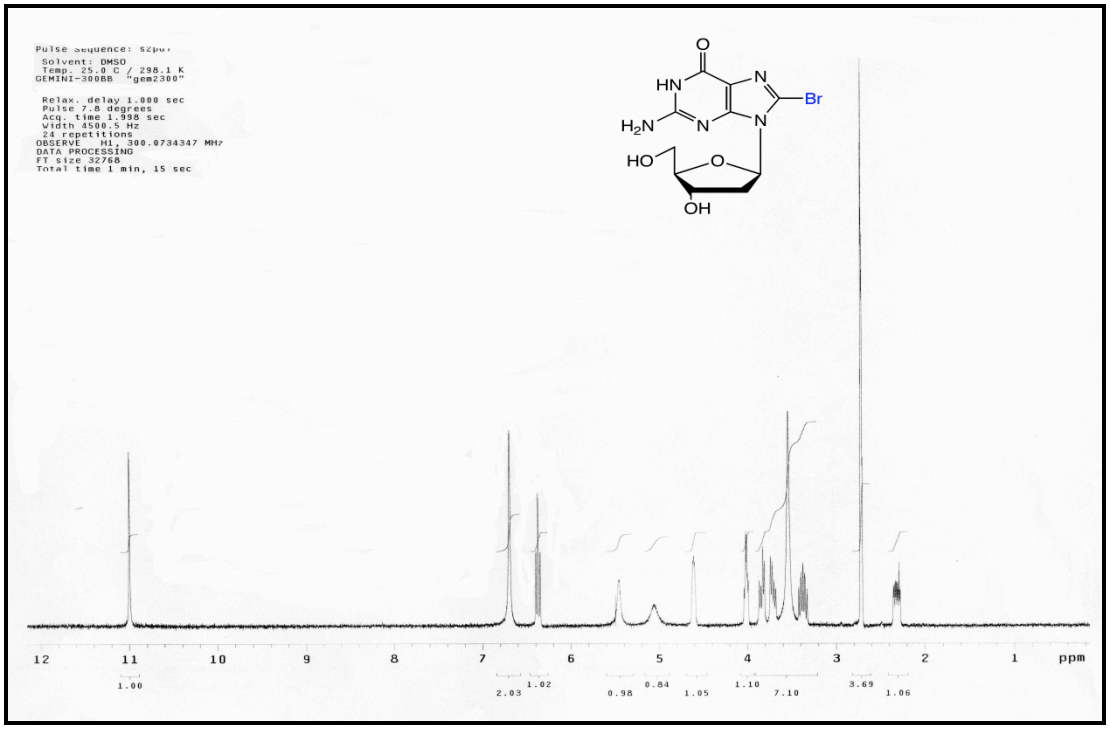
UV-absorption thermal denaturation experiments were carried out as described [67] and using the 'thermal' program of a CARY 100 Bio UV-VIS spectrophotometer equipped with a multicell block temperature regulation unit and a fluid conduction thermal regulation enhancement attachment (Varian, Inc.). Four duplex samples at a

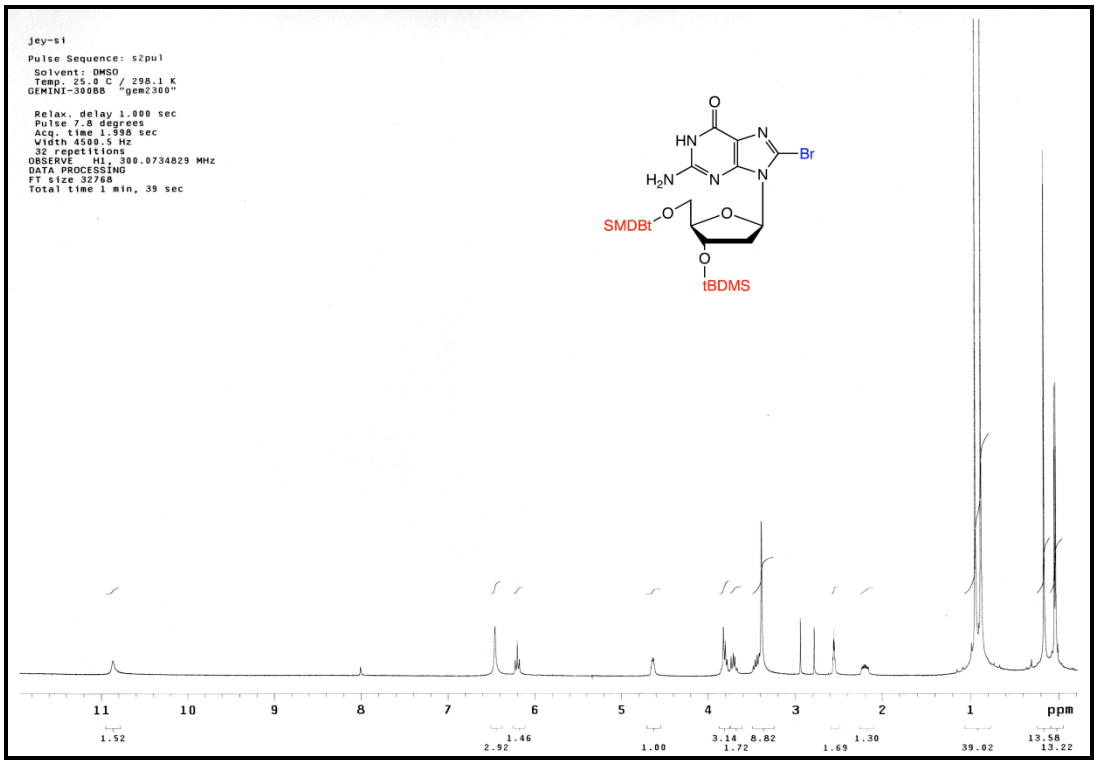
time, plus a blank, were placed in stoppered cuvettes and inserted into the spectrophotometer. The pre-run temperature was allowed to equilibrate for at least 10min to either 15 or 80°C or decreased from 80 to 15°C at a rate of 0.2°C/min for complete temperature equilibration at each point. DNA samples consisted of between 0.3 and 1.2 OD_{260} units of duplex dissolved in 1mL of 25mM sodium phosphate buffer solution, pH 6.8, containing 100mM NaCl and 0.5mM EDTA. Eight different sample concentrations and three independent melting profiles at each concentration composed a set of 24 independent determinations performed on each duplex. Thermodynamic stabilities were calculated by a protocol from the published articles [68, 69].

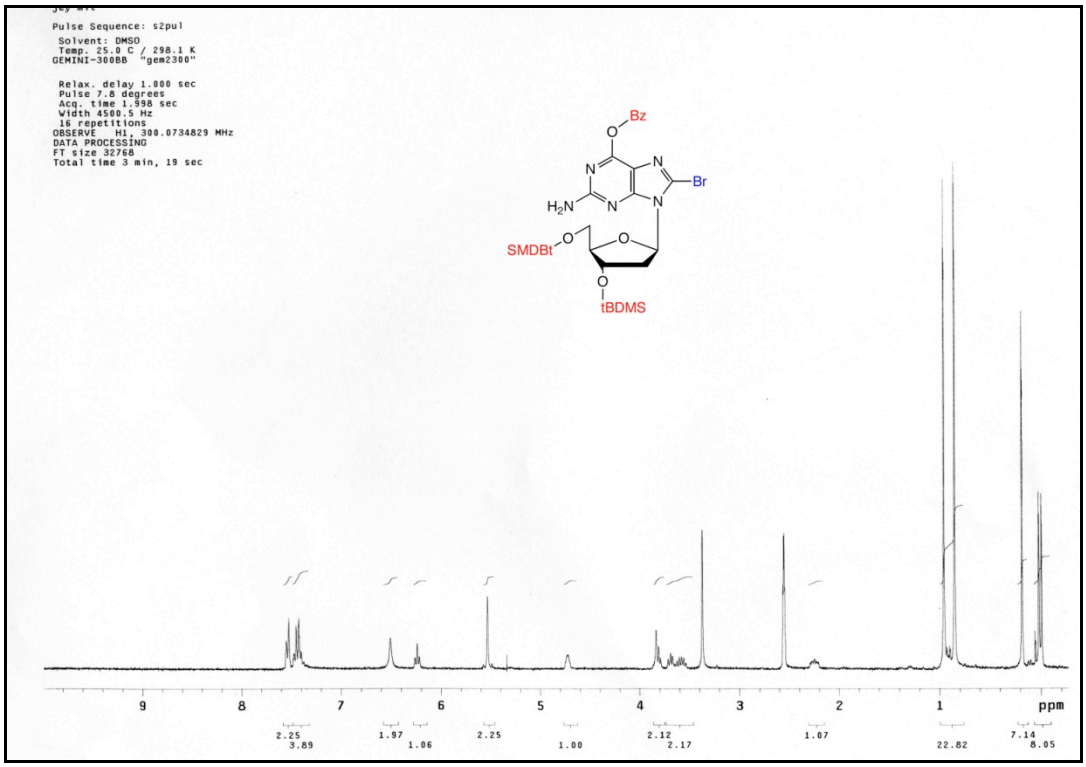
SUPPLEMENTARY DATA

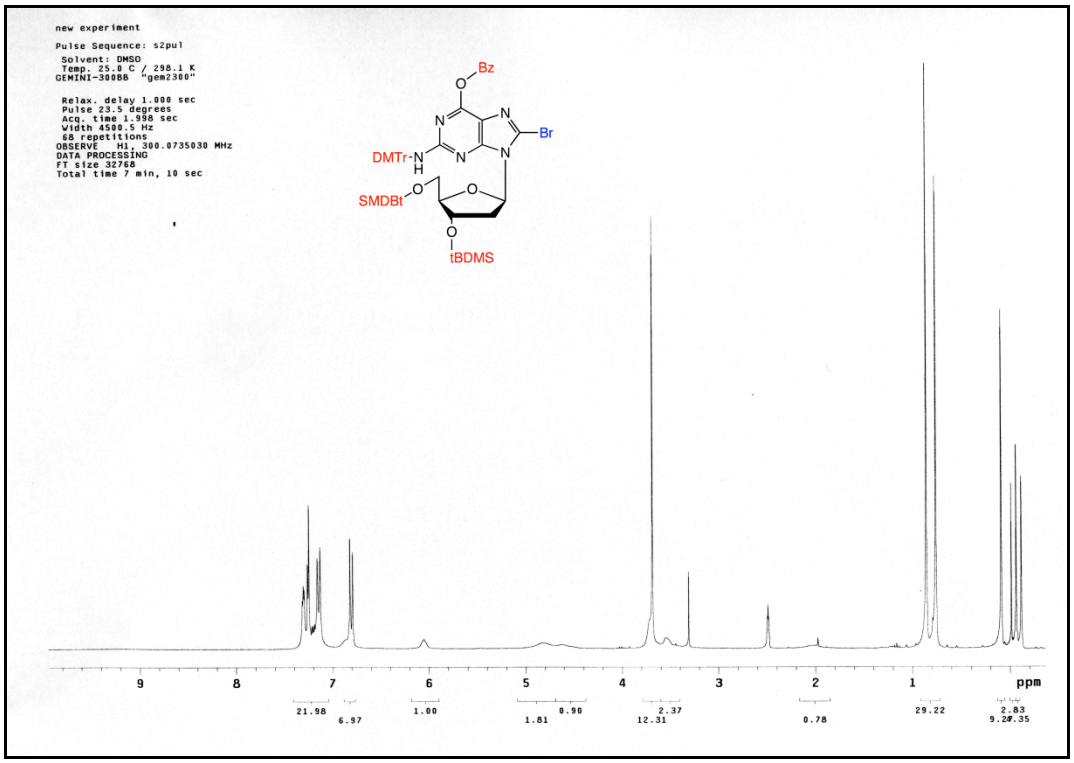
NMR data

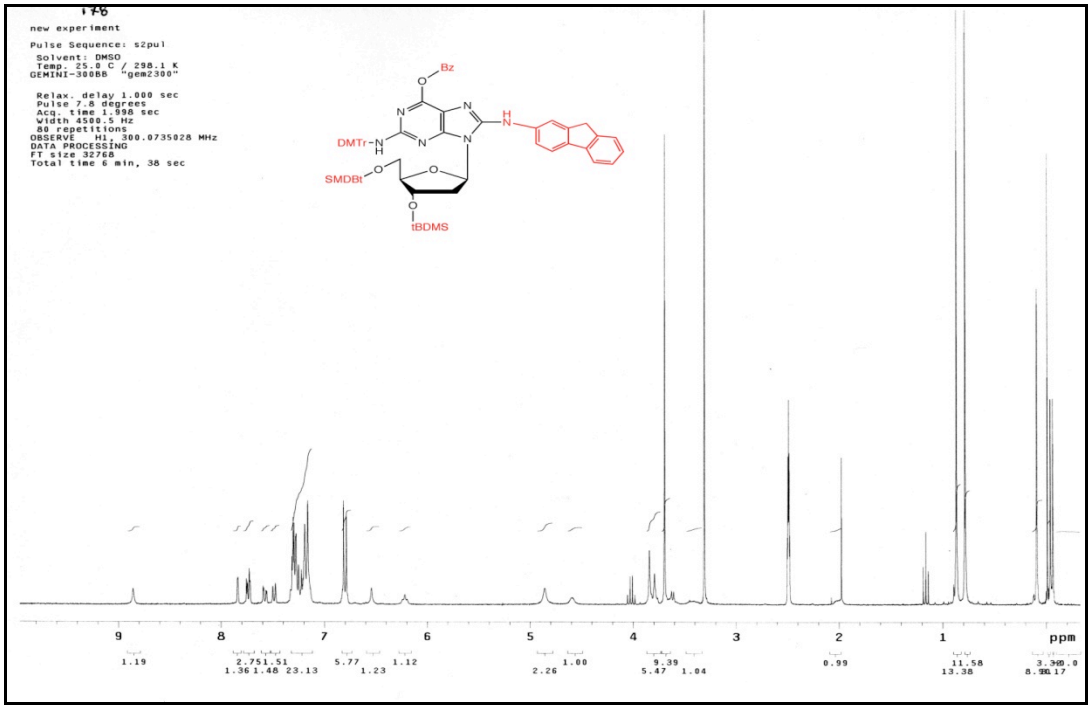
Supplementary Date- NMR data

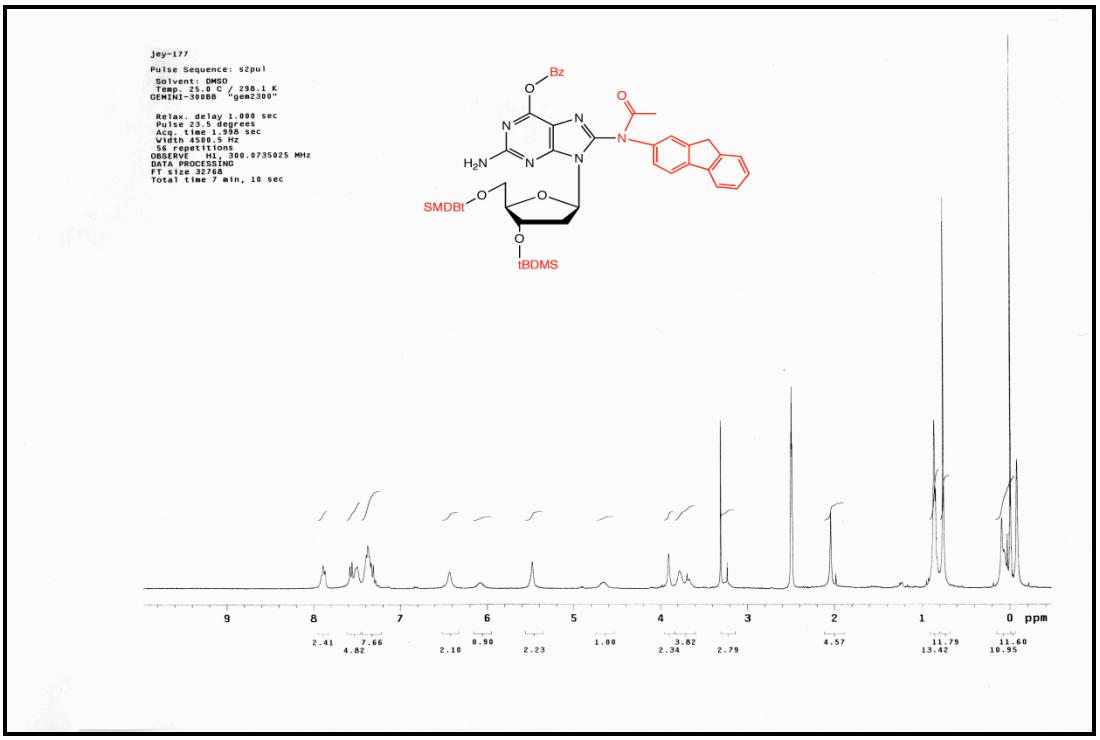


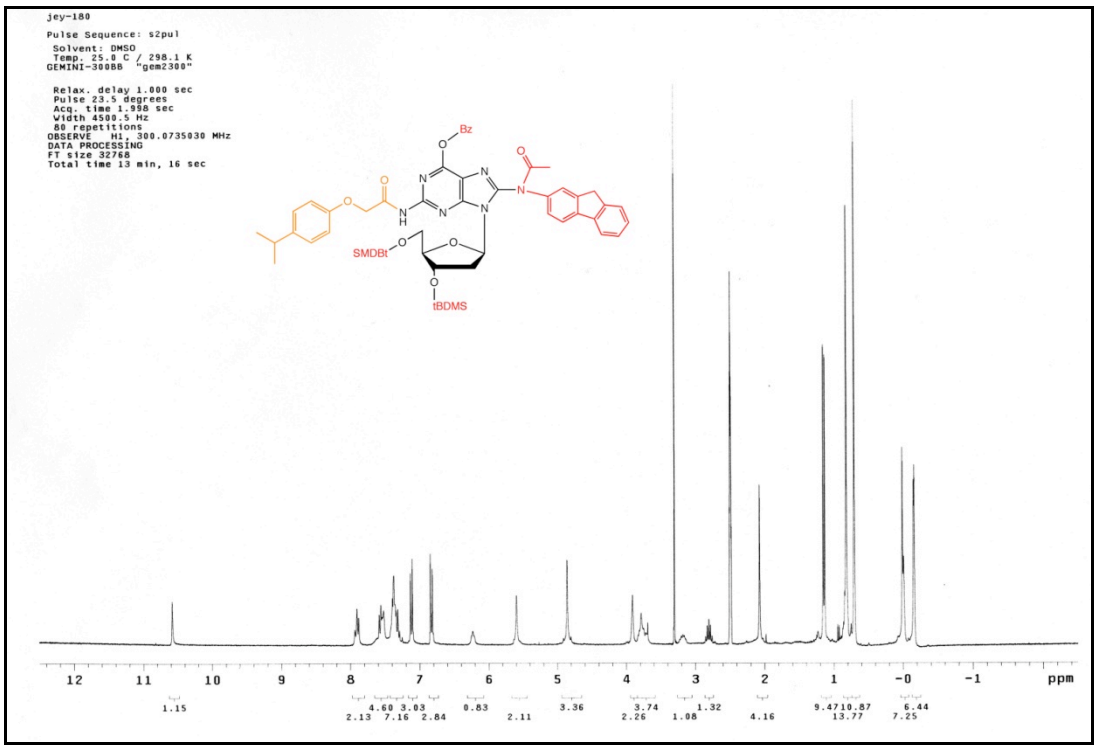

jey-si
Pulse Sequence: s2pul
Solvent: DMSO
Temp. 25.0 C / 298.1 K
GEMINI-300BB "gem2300"
Relax. delay 1.000 sec
Pulse 7.8 degrees
Acq. time 1.998 sec
Width 4500.5 Hz
32 repetitions
OBSERVE H1, 300.0734829 MHz
DATA PROCESSING
FT size 32768
Total time 1 min, 39 sec
HN
H2N
N
Br
SMDBt
O
tBDMS
11 10 9 8 7 6 5 4 3 2 1 ppm
1.52 2.92 1.46 1.00 3.14 8.82 1.72 1.69 1.30 39.02 13.58 13.22

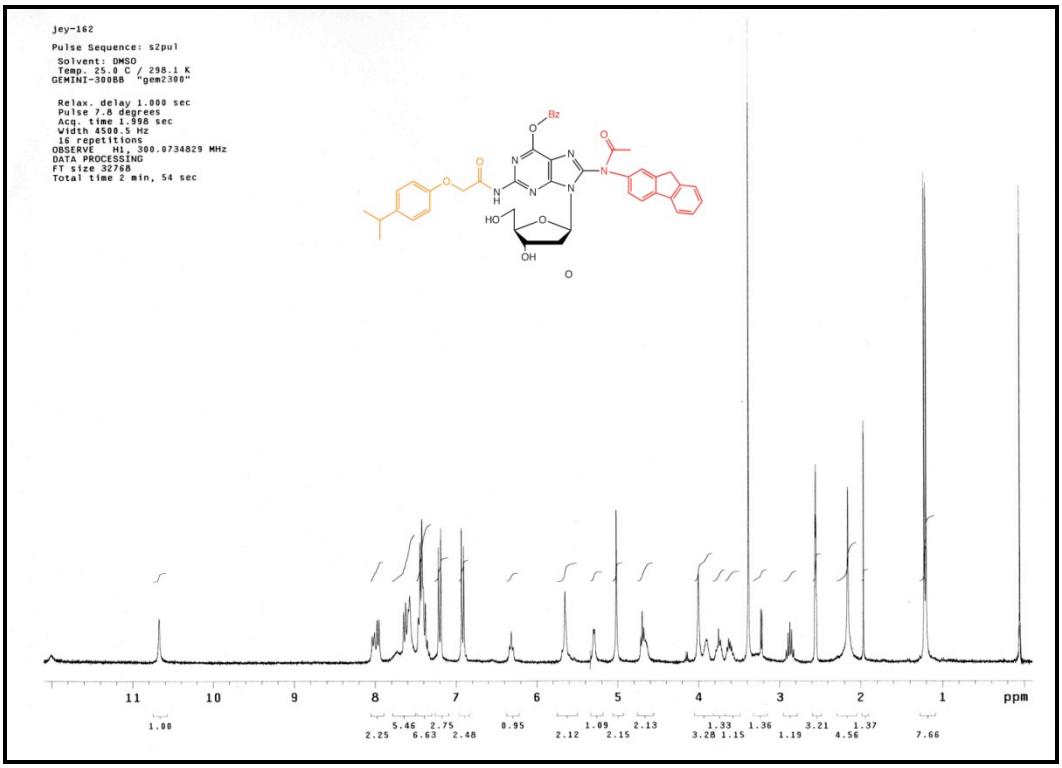

Pulse Sequence: s2pul
Solvent: DMSO
Temp. 25.0 C / 298.1 K
GEMINI-300BB "gem2300"
Relax. delay 1.000 sec
Pulse 7.8 degrees
Acq. time 1.998 sec
Width 4500.5 Hz
16 repetitions
OBSERVE H1, 300.0734829 MHz
DATA PROCESSING
FT size 32768
Total time 3 min, 19 sec
Bz
O
N
Br
H2N
SMDBt
tBDMS
9 8 7 6 5 4 3 2 1 ppm
2.25 3.89 1.97 1.06 2.25 1.00 2.12 2.17 1.07 22.82 7.14 8.05


new experiment
Pulse Sequence: s2pul
Solvent: DMSO
Temp. 25.0 C / 298.1 K
GEMINI-300BB "gem2300"
Relax. delay 1.000 sec
Pulse 23.5 degrees
Acq. time 1.998 sec
Width 4500.5 Hz
68 repetitions
OBSERVE H1, 300.0735030 MHz
DATA PROCESSING
FT size 32768
Total time 7 min, 10 sec
Bz
Br
DMTr-N
H
SMDBt
tBDMS
ppm


new experiment
Pulse Sequence: s2pul
Solvent: DMSO
Temp. 25.0 C / 298.1 K
GEMINI-300BB "gem2300"
Relax. delay 1.000 sec
Pulse 7.8 degrees
Acq. time 1.998 sec
Width 4500.5 Hz
80 repetitions
OBSERVE H1, 300.0735028 MHz
DATA PROCESSING
FT size 32768
Total time 6 min, 38 sec
Bz
H
N
DMTr-N
H
SMDBt
tBDMS
ppm


jey-177
Pulse Sequence: s2pul
Solvent: DMSO
Temp. 25.0 C / 298.1 K
GEMINI-300BB "gem2300"
Relax. delay 1.000 sec
Pulse 23.5 degrees
Acq. time 1.998 sec
Width 4500.5 Hz
56 repetitions
OBSERVE H1, 300.0735025 MHz
DATA PROCESSING
FT size 32768
Total time 7 min, 10 sec
Bz
SMDBt
tBDMS
ppm


jey-180
Pulse Sequence: s2pul
Solvent: DMSO
Temp. 25.0 C / 298.1 K
GEMINI-300BB "gem2300"
Relax. delay 1.000 sec
Pulse 23.5 degrees
Acq. time 1.998 sec
Width 4500.5 Hz
80 repetitions
OBSERVE H1, 300.0735030 MHz
DATA PROCESSING
FT size 32768
Total time 13 min, 16 sec
Bz
SMDBt
tBDMS
ppm



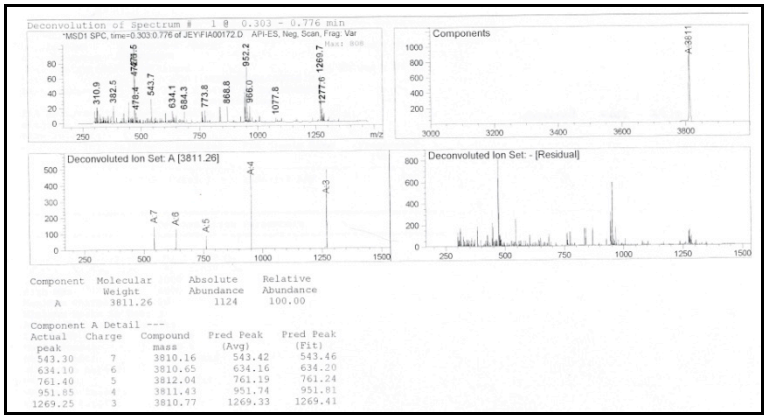


Figure 1: Mass Spectrometry (MS) result of 12mer NarI-1 AAF, expected molecular weight: 3812, obtained molecular weight: 3811

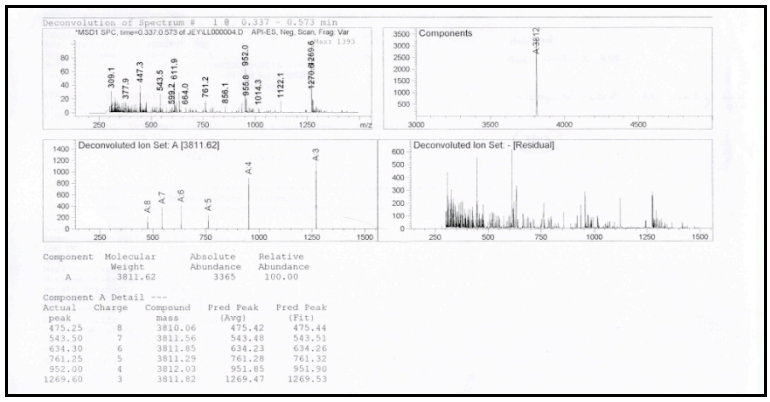


Figure 2: MS result of 12mer NarI-2 AAF, expected molecular weight: 3812, obtained molecular weight: 3812

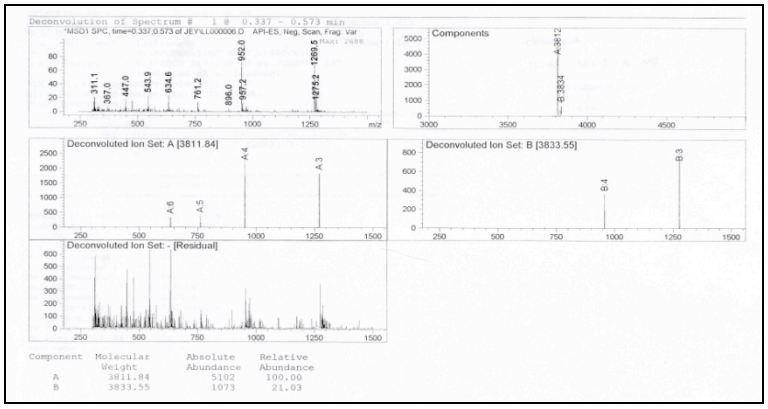


Figure 3: MS result of 12mer NarI-3AAF, expected molecular weight: 3812, obtained molecular weight: 3812

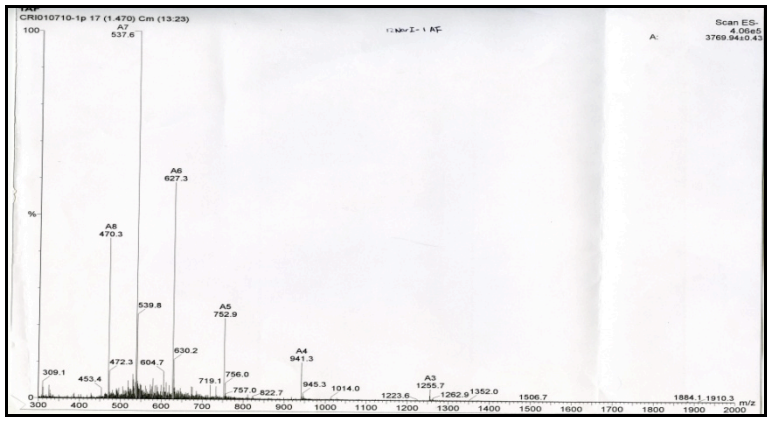


Figure 4: MS result of 12mer NarI-1 AF, expected molecular weight: 3770, obtained molecular weight: 3769.94

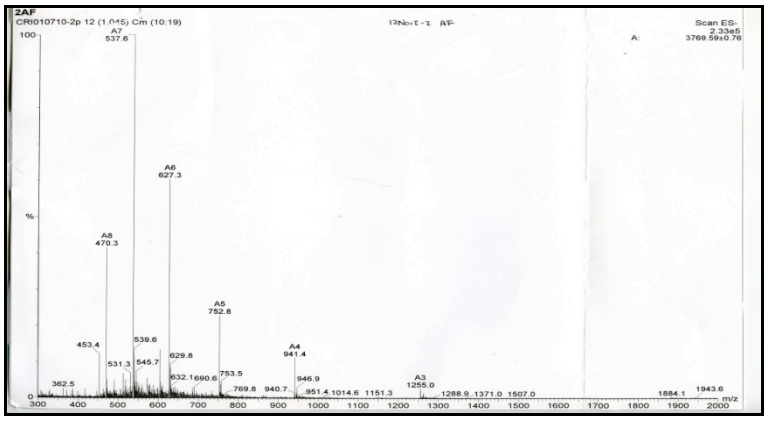


Figure 5: MS result of 12mer NarI-2 AF, expected molecular weight: 3770, obtained molecular weight: 3769.59

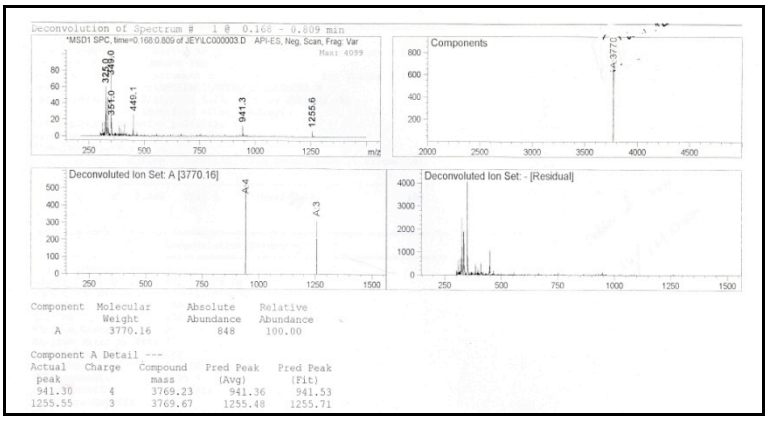


Figure 6: MS result of 12mer NarI-3 AF, expected molecular weight: 3770, obtained molecular weight: 3770

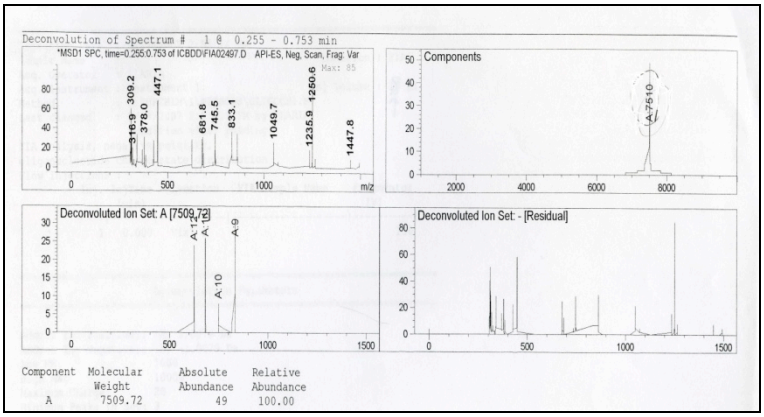


Figure 7: MS result of 24mer NarI-1 AAF, expected molecular weight: 7509.8, obtained molecular weight: 7510

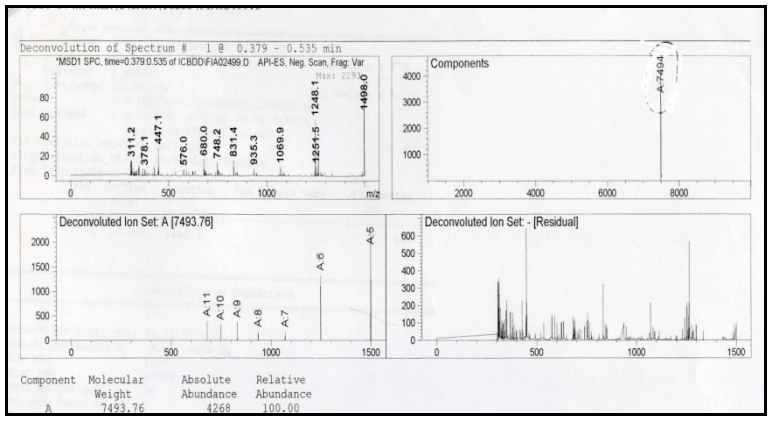


Figure 8: MS result of 24mer NarI-3AAF, expected molecular weight: 7494.8,obtained molecular weight: 7493.76

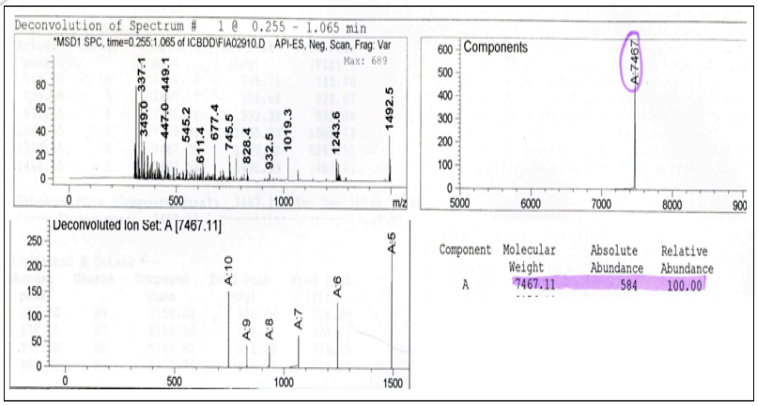


Figure 9: MS result of 24mer NarI-1AF, expected molecular weight: 7467.8,obtained molecular weight: 7467

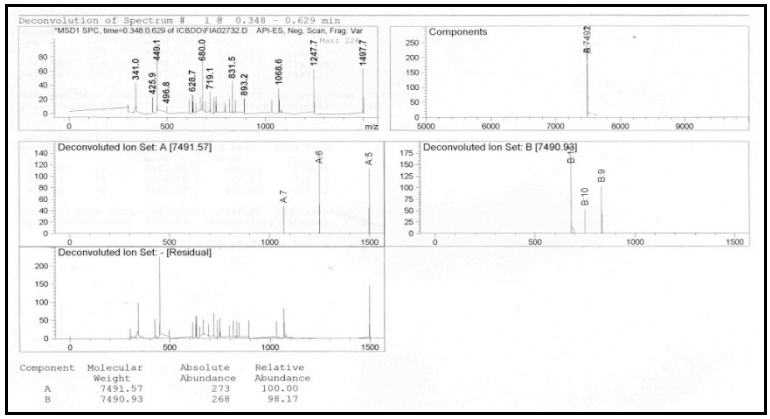
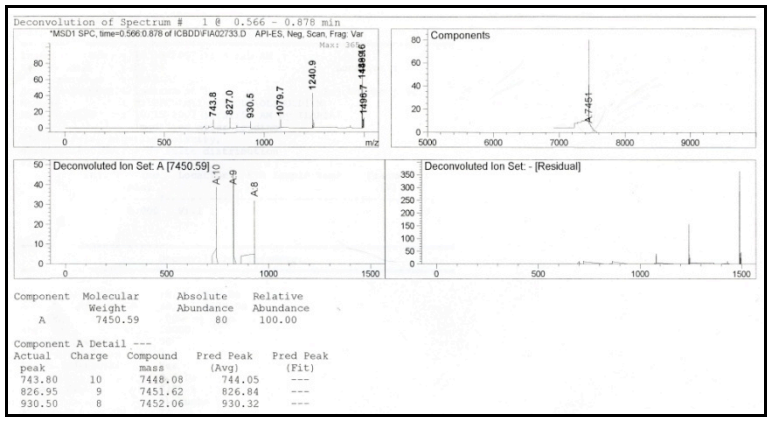


Figure 10: MS result of 24mer NarI-2AF, expected molecular weight: 7492.83,obtained molecular weight: 7491.57



Component	Molecular Weight	Absolute Abundance	Relative Abundance
A	7450.59	80	100.00

Component A Detail ---

Actual peak	Charge	Compound mass	Pred Peak (Avg)	Pred Peak (Fit)
743.80	10	7448.08	744.05	---
826.95	9	7451.62	826.84	---
930.50	8	7452.06	930.32	---

Figure 11: MS result of 24mer NarI-3AF, expected molecular weight: 7452.8, obtained molecular weight: 7451

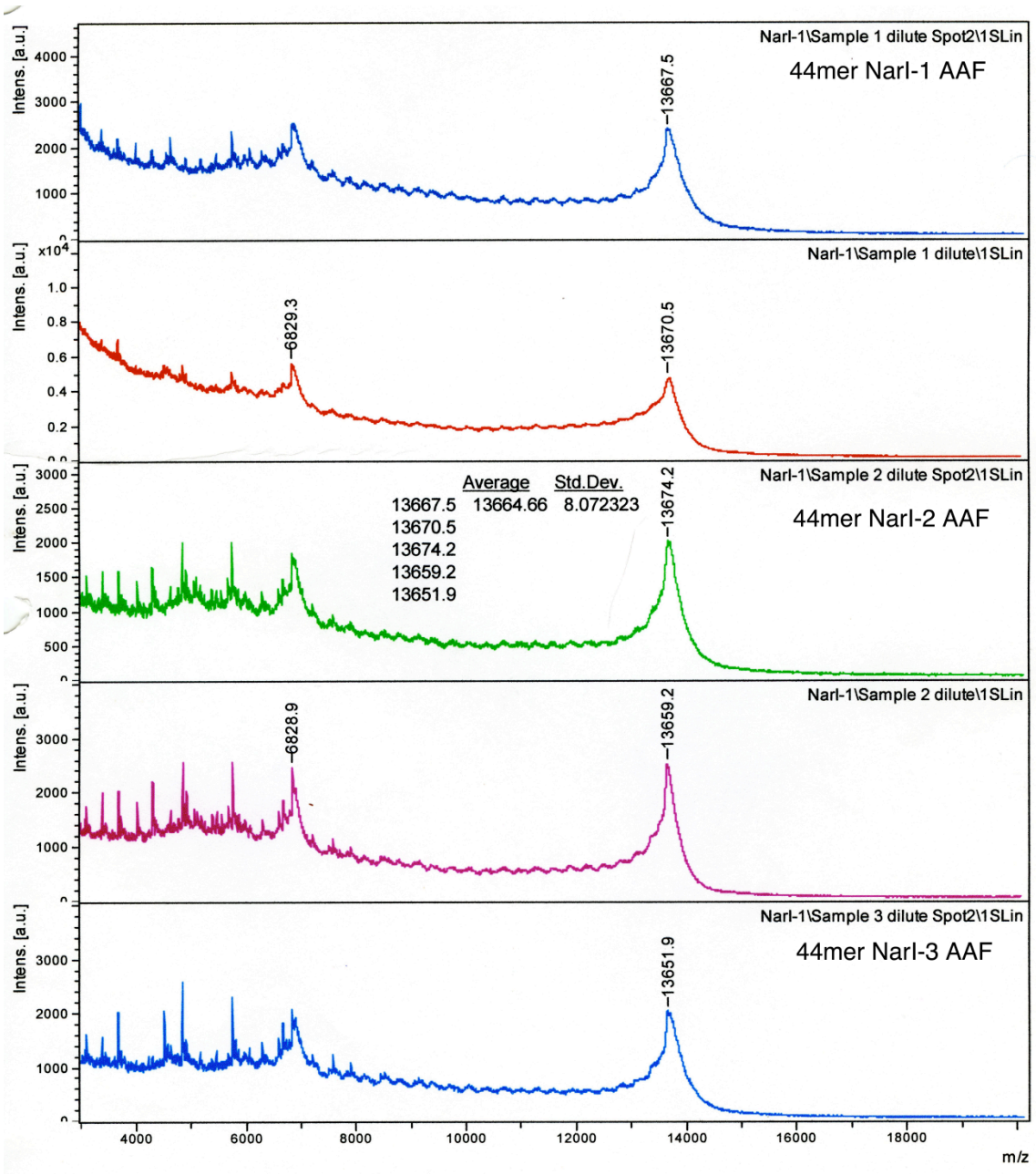


Figure 12: MS result of 44mer NarI-1, 2, and 3 AAF, expected molecular weight: 13659.1, obtained molecular weight (Average): 13664.66 from MALDI-TOF

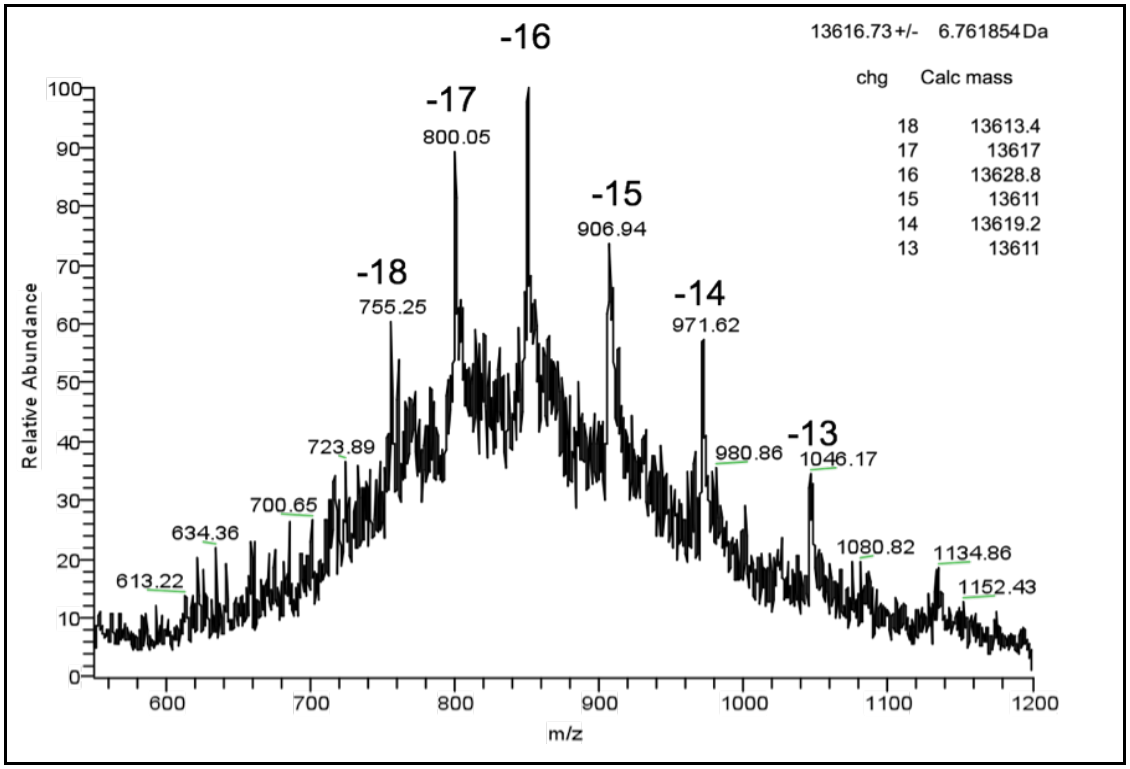


Figure 13: MS result of 44mer *Nar*I AF, expected molecular weight: 13617.1, obtained molecular weight (Average): 13616.73

CHAPTER 3:

MECHANISMS FOR REPAIR OF ARISTOLOCHIC-ACID DNA ADDUCTS

Adapted from a manuscript prepared by Yeo, J., Sidorenko, V.S., Attaluri, S., Bonala, R. R., Izaitseva, I., Johnson F., Iden, C. R., Grollman, A.P., and Schärer, O. D. (not published)

Abstract

Aristolochic acids I and II, AAI and AAII, are constituents of certain widely used herbal remedies. Both AAI and AAII are genotoxic and induce A→T transversions and cause urothelial cancer. Adducts of AAI with deoxyadenosine, dA-ALI, persist in human kidney cortex DNA for years after ingestion of AAI. The mutation spectra of the *p53* gene of aristolochic acid-associated urothelial cancers reveals a marked strand bias for mutations at dA residues located on the non-transcribed strand, suggesting that dA-AL adducts are selectively repaired by transcriptional-coupled nucleotide excision repair. We used a set of xeroderma pigmentosum proteins (XP) - or Cockayne syndrome proteins (CS)-deficient, complemented and normal human fibroblasts to investigate the toxic effect of AAII. Most XP-deficient cells were both more sensitive and had elevated amount of AA-DNA adducts in comparison to the control and XPC deficient cell lines. For CS-deficient cell lines we observed the same level of sensitivity to AAII as that for the control cells, with the adduct levels being higher. CSB-complemented cells with the adduct level similar to the normal cell line were more sensitive than the normal fibroblasts. To investigate AL-DNA adducts repair *in vitro* we have generated plasmids containing a single dA-ALII lesion in different sequence contexts and treated them with NER proficient (HeLa) and deficient (XPF) extracts with the following [a-^{32}P]-dCTP labelling. Together these studies and the binding assay with purified XPC-HR23B indicate that the repair efficiency of AL-DNA adduct depends on the sequence context and, unless placed in a mismatch, is highly resistant to global genomic repair, suggesting that a lack of destabilization by dA-ALII prevents binding by XPC-RAD23B and repair by NER.

Introduction

The General Background of Aristolochic Acid (AA) Toxicology

Aristolochic acids I and II (AAI and AAII) are the major phenanthrene metabolites of widely distributed *Aristolochia* spp. [70, 71], a plant whose extracts are still widely used in herbal medicine [71]. The mixture of aristolochic acids, AA, was considered as a potential antimicrobial drug and immune stimulator until the previous study showed that these compounds exhibited renal toxicity and carcinogenic potential in rodents [72].

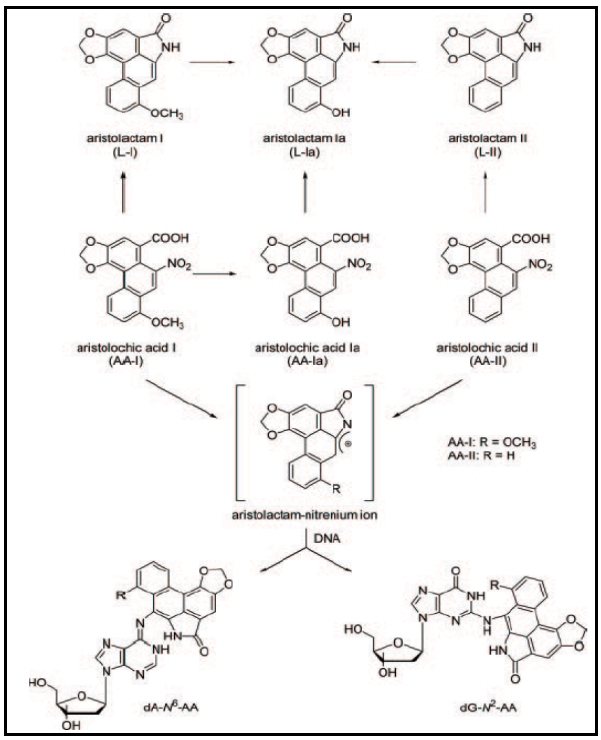
In the early 1990s over a hundred Belgian women were hospitalized with renal failure after being subjected to a slimming regiment containing *Aristolochia fangchi* instead of *Stefania tetrandra* by mistake. AAI and AAII constituents were found in extracts of the administered pills extract and aristolochic acid DNA adducts (AA-DNA) were identified in the kidney tissue samples of these patients confirming specific exposure to AA [71]. The outbreak was called Chinese herbs nephropathy (CHN) and later renamed aristolochic acid nephropathy, AAN, and was characterized by a severe tubulointerstitial nephritis, progressing to renal failure and to the urothelial tract malignancies in 50% of the cases. The use of *Aristolochia* species is banned by Federal Food and Drug Administration, but remedies containing AA are still widely available on the internet.

The endemic proportions of AAN has recently been realized when that devastating Balkan endemic nephropathy (BEN), characterized by chronic renal interstitial fibrosis with slow progression to terminal renal failure and the high incidence of unusual upper urothelial tract tumors was also caused by AA [73]. In this instance, the seeds of *Aristolochia clematitis* were comingled with wheat and AA compounds were ingested with bread [70] and the presence of AA-dA adducts was directly demonstrated in the DNA of kidney tissues of BEN patients. Mouse and cell culture studies, furthermore, confirmed that AA possesses nephrotoxic, mutagenic and carcinogenic properties [72, 74].

Aristolochic acids are activated by cellular cytochromes 1A1, 1A2 and some cytosolic reductases [75-78] to nitrenium metabolites that can react with proteins and

DNA (Scheme 1). AAI and AAII produce the following main adducts with DNA: 7-(deoxyadenosin-N^6-yl) aristolactam I (dA-ALI), 7-(deoxyguanosin-N^2-yl)aristolactam I (dG-ALI) and 7-(deoxyadenosin-N^6-yl)aristolactam II (dA-ALII) and 7-(deoxyguanosin-N^2-yl)aristolactam II (dG-ALII). These dG and dA adducts block DNA replication and have miscoding properties. Exposure to AA leads predominantly to A→T transversions, since the dA adducts are both more mutagenic and persistent than the dG adducts [79]. One of the most remarkable feature of the dA adducts of AA (dA-AL) is their extremely long persistence in the genome [73]; dA-AL has been detected in cells of exposed individuals years after the last AA ingestion, suggesting that at least a subset of these adducts is resistant to repair [73].

In humans nucleotide excision repair (NER) is mainly responsible for the removal of bulky DNA lesions, such as the ones formed by AA [4]. NER can be divided into two pathways, transcriptional coupled repair (TC-NER) and global genome repair (GG-NER). TCR is only active on transcribed strand of active genes and is initiated by an RNA polymerase stalled at the lesion [80]. By contrast, GG-NER monitors the enitre genome and is initiated by the damage sensor XPC-RAD23B [10, 16], in some cases with the help of a ubiquitin-ligase complex, containing the UV-damage binding complex UV-DDB [81]. For GG-NER, there is a strong correlation between the degree of thermodynamic destabilization induced in a duplex by a DNA lesion and the efficiency of NER [54]. In line with this observation, XPC-RAD23B has been found to bind lesion not by directly binding the lesion, but rather ssDNA opposite the lesion made more accessible by destabilization [11, 12, 82]. No such correlations have been established for TC-NER, and it is believed that bulkiest lesion will provide a block for a transcribing RNA polymerase II [83, 84].



Scheme 1: Metabolism of aristolochic acids and formation of AA-derived DNA adduct [86].

Analysis of the mutational profile of the *p53* tumor suppressor gene in BEN patients showed A→T transversions characteristic of AA exposure [85, 88]. Interestingly, these mutations were found exclusively in the non-transcribed strand, suggesting that dA-AL adducts are likely to be refractory to GG-NER, while being repaired by TC-NER on the actively transcribed strand. A similar strand bias was found in Hupki mice exposed to AAI [71]. Both the persistence of dA-AL adducts and a marked strand bias in A→T mutations suggest that dA- AL adducts are selectively repaired by TCR.

To verify the possible involvement of TC-NER and GG-NER in the repair of dA-AL DNA adducts we investigated the cytotoxicity and the genotoxicity of AAII in a series of cell lines with defects in one or both of these repair pathways. We further investigated the NER susceptibility and XPC-RAD23b binding affinity or dA-ALII lesions *in vitro*. Our results strongly suggest dA-AL DNA adducts are refractory to repair by GG-NER, because they fail to bind XPC-RAD23B, while they are efficiently processed by TC-NER.

RESULTS

In Vivo Cell Line Assay Results

Cell lines sensitivity to AAII

We used a series of fibroblast cell lines with mutations in key components of TCR-NER and/or GGR-NER to investigate the relative roles of these processes in the genotoxic and the cytotoxic response to AAII (Table 1). Cell lines deficient in one of the components of a pathway responsible for AA-DNA adducts repair are expected to be more sensitive to AA exposure compared to the control cell line. Cells lacking functional CSA or CSB are selectively deficient in TC-NER, those lacking XPC are deficient in GG-NER, whereas cells deficient in XPA, XPD, XPG, or XPF have deficiencies in both pathways. The following corrected human fibroblast cell lines were used: CSB^{+}, CSB^{mut} (E646Q an ATPase deficient mutant of CSB), XPA^{+} and XPD^{+}. The cell lines were treated with different doses of AA and for comparison with cisplatin and benz[o]pyrene and their viability was measured by an ATP luciferase-dependent proliferation assay.

Cell	Phenotype	Name	expected GGR status	expected TCR status
GM00637	wild-type	Control	+	+
Cs1an/p3.1	CSB	CSB	+	-
Cs1an/p3.1-CSBwt	CSB-corrected	CSB^{+}	+	+
Cs1anE7/p3.1-CSBE646Q	CSB-mutant corrected	CSB^{mut}	+	-
GM16094	CSA	CSA	+	-
GM15983	XPC	XPC	-	+
GM04312	XPA	XPA	-	-
GM15876	XPA corrected	XPA^{+}	+	+
GM08437	XPF	XPF	-	-
GM14930	XPG	XPG	-	-
GM08207	XPD	XPD	-	-
GM15877	XPD corrected	XPD^{+}	+	+

Table 1: Cell lines used in the study.

Cisplatin used as a control for the expected behavior of exposure of the cell lines to AA, as it has been shown that cells with a defect in TC-NER (CSA-, XPD-, XPG-, and XPA-deficient human fibroblasts) are hypersensitive to cisplatin in an MTT assay [86]. Our results obtained with the ATP assay were fully consistent with these observations (Supplementary Figure 1).

Among GGR/TCR deficient cell lines XPA was the most sensitive to AAII exposure. XPA complementation rescued the phenotype to some extent as compared to the control cells (Figure 1A). XPG and XPD cells were more affected than the control, with XPG being more sensitive than XPD. XPD^{+} cell line had the similar sensitivity as the normal human fibroblasts (Figure 1B and C). Only XPF among the GG-NER/TC-NER compromised cells had not displayed any hypersensitivity to the treatment (Figure 1D). According to the Coriell repository the XPA cell line has the unscheduled DNA synthesis (UDS) level less then 2% from normal. XPF cell line retained at least 10% of UDS. No information was available on XPD and XPG cells. Thus, varying UDS level between different cell lines may be implicated in the response to AAII. Neither GG-NER- nor TC-NER-deficient cells were hypersensitive to AAII exposure in comparison to normal fibroblasts, with CSA, CSB, XPC and the control cell line being equally affected (Figure 1E, F and G). Intriguingly, CSB complementation with the wild type protein, but not the E646Q mutant with ATPase activity deficiency, led to the increased cytotoxicity compared to control cells, suggesting that overexpression of active CSB influences the cellular response to AA (Figure 1F).

The deficiency only in TC-NER or GG-NER pathway was not enough to cause the hypersensitivity to AAII in compromised cells. Nevertheless the cells lacking the functional XPA, XPD or XPG components responded as predicted with the increased cytotoxicity levels, suggesting that AA-DNA adducts are repaired *in vivo* by NER.

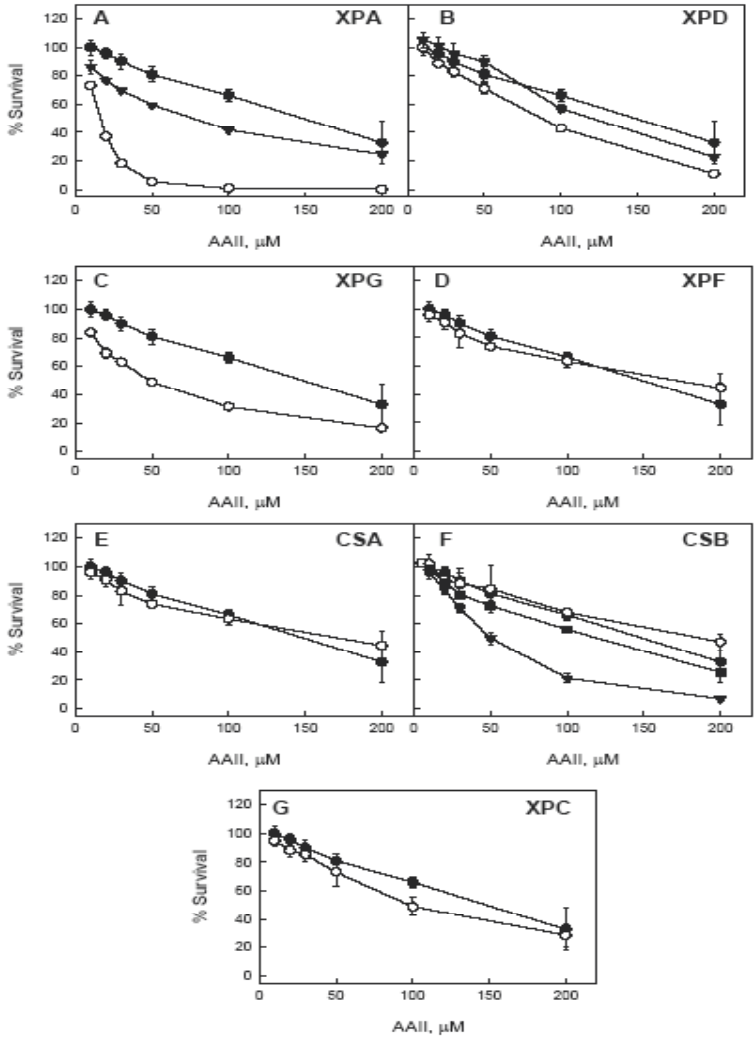


Figure 1: AAII cytotoxicity study. XP-, CS- deficient and control normal human fibroblasts cell lines were treated with 5-200 μM AAII for 72 hours with the following ATP level measurements. Filled circles (●) indicate the control normal human fibroblasts cell line, open circles (○) - the deficient cell lines corresponding to the panel names, filled triangles (▼) – complemented cell lines if available. On the panel F filled squares (■) indicate the CSB deficient cell line complemented with CSB ATPase mutant gene. The point for each AAII dose corresponds to the ATP ratio between treated and untreated cells and presents mean ± standard deviation for at least six wells in two independent experiments.

ALII-DNA adduct levels in XP- and CS-deficient cell lines

In order to identify the genotoxicity of AAII and find a possible correlation between the cytotoxicity and ALII-DNA adduct levels, the cell lines were treated with different AAII doses for 48 hrs and the levels of adducts were determined. The exposure time in these assays was chosen to avoid cell death and thus the depletion of cells with adducts from the plates. Following exposure, DNA was digested to single, nucleotides, labeled with γ-[^{32}P]-ATP and analyzed by gel-electrophoresis, which allows to determine

adduct. This post-labeling approach allows for the determination of adduct levels in many samples on one gel and with excellent sensitivity [86]. We first performed a time course exposure with the low 2 μM AAII dose for selected cell lines to check the linearity of the adduct formation (Supplementary Data, Figure 2). To estimate the amount of adducts, oligonucleotides with single dG-ALII or dA-ALII adduct were used. We have previously determined the mobility of these two adducts in a polyacrylamide gel [73]. In a mock treatment no bands corresponding to the dA- and dG-adduct were found on a gel (data not shown). Figure 2H shows the fragment of the acrylamide gel for the adduct analysis.

We detected 5-10 fold higher levels of dA-AL adducts compared to dG-AL in all cell lines after treatment with AA-II over a dose range of 20-100 uM (Figure 2). These adduct levels reflect the balance between rates of adduct formation and repair. It should be emphasized that the measurement of repair itself is complicated by intracellular accumulation of the chemical and different growth rates of the cell lines. After removal of AAII from the medium, washing and continued growth in AA-free medium, the cell lines continued to divide though at reduced rates. In spite of that, after incubation for 80 hours after exposure, the cells still had significant levels (>70% of initial adduct levels) of both dG- and dA-ALII adduct (data not shown). These results are consistent with our measurements of adduct levels in primary human urothelial cell cultures (unpublished data).

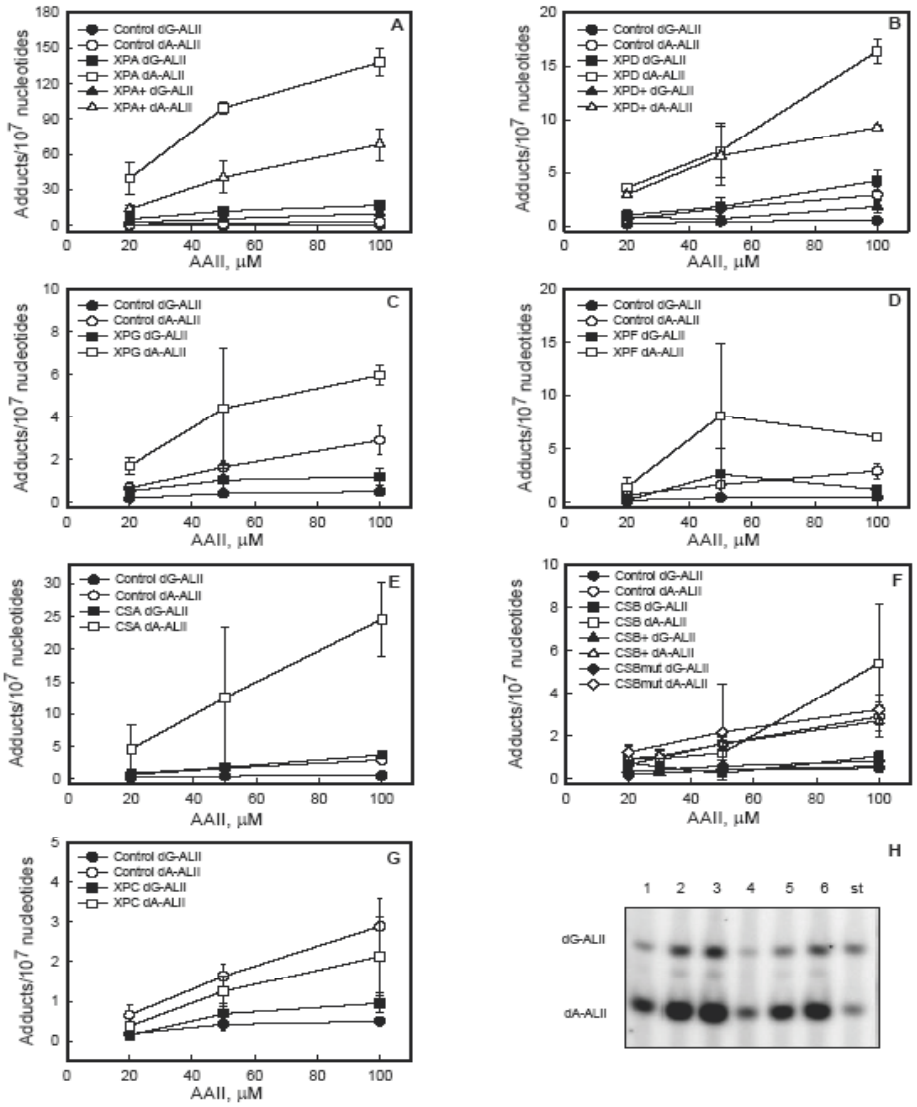


Figure 2. AAII genotoxicity study. XP-, CS- deficient, normal fibroblast and complemented cell lines if available were treated with 20, 50, 100 μM AAII for 48 hours. DNA was isolated and analyzed for adduct levels with ^{32}P-post-labelling assay. A-G – The dependence of dG- and dA-ALII adduct levels from the treatment AAII dose. Control indicates the control human fibroblasts cell line. XPA, XPD, XPG, XPF, CSA, CSB, CSBmut, XPC – deficient human fibroblasts cell lines. XPA$^+$, XPD$^+$, CSB$^+$ are corresponding complemented cells. All treatments for every cell line were done in triplicate and the results are presented as means ± standard deviations. H – The fragment of a 30% PEG after ^{32}P-labelling of DNA adduct nucleosides. 1-3 – XPA deficient cell line treated with 20, 50, 100 μM AAII for 48 hours; 4-6 – XPA complemented cell line treated the same way; 7 – The standard mixture of 24-mer oligonucleotides containing single dG-ALII or dA-ALII, the upper and lower band respectively. Each standard band corresponds to 1 adduct/10^6 nucleotides for 5 mg DNA.

The levels of dA-AA and dG-AA adducts in cells with deficiencies in the different NER genes decreased in the following order: XPA > XPD ≈ CSA > XPG ≈ XPF > CSB > XPC > WT control (Figure 2). CSB deficiency led to higher adduct levels in comparison to control and CSB$^+$ cell lines at the highest dose (Figure 2F) and the prolonged exposure time course (data not shown). Complementation of XPA and XPD deficiency

led to significant decreases in dG-AL and dA-AL levels (Figure 2A and B), while the adduct levels in XPC-deficient cells were comparable to control values, suggesting that GG-NER plays a minor role in adduct repair (Figure 2G). We did, however, observe a minor, but significant difference in dG-ALII, but not dA-ALII, adduct levels after exposure to 100 μM AAII between XPC cells and the control, suggesting that dG-ALII adducts are repaired by both TC-NER and GG-NER. The lack of the difference at the lower chemical doses might reflect the detection limitation of the post-labeling method where the poor signal intensity on a gel causes the higher standard deviation between the treatments. These studies, determining the levels of adducts *in vivo,* suggest that TC-NER is the main pathway for repairing AA adducts in DNA in particular for dA-ALII adducts in agreement with the mutation spectra observed in the p53 gene in patient with urothelial cancer caused by exposure to AA.

In Vitro Results

Preparation of ssDNA and New Plasmids for the Preparation of dA-AL substrates at the Codon and Splice Site sequences

In order to investigate a repair possibility and efficiency of dA-ALII adducts by GG-NER, several oligonucleotides containing a site-specific ALII were prepared [79] (Table 2) and introduced into plasmids for *in vitro* NER assays. Oligonucleotides containing dA-ALII at codon 209 (AC**A**GA) or at a splice site in Intron 6 (CT**A**GG) (see Table 2, [88]) were used for this study. The previously found sequences [88] in Table 1 and dA-AL adducts are remarkably persistent in human cells, indicating that they are refractory to repair. In more details, to introduce the oligonucleotides containing a defined dA-ALII into a plasmid, ssDNA restraining a complementary sequence to the new sequence (Table 2) is required as described in the published protocol [62]. To obtain ssDNA containing an useful sequence information (Table 2), p97 (pBluscript, BbsI NdeI in pBlkp II SK+, 3kb) was used as a template. A plasmid, p97, was doubly digested with BbsI to give sticky ends to accommodate new oligonucleotides containing the new sequences (Table 3). Gel extraction of doubly digested p97 was performed to have a pure linear DNA (Scheme 2A), which was ligated with annealed oligonucleotides to

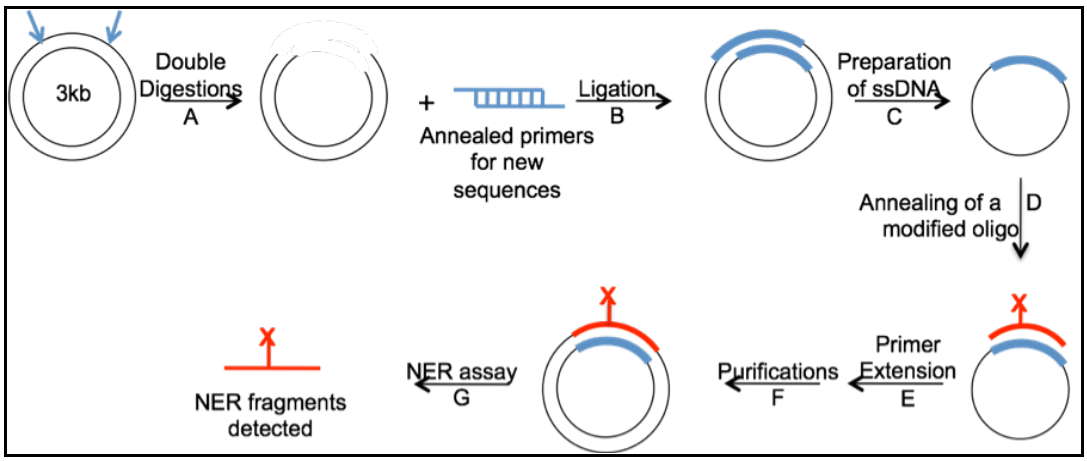
introduce the new sequences to be used to incorporate oligonucleotides containing aristolactam adducts (Scheme 2B). The new plasmids were sequenced (Figure 3) and used as a template to generate ssDNA with R408 helper phage (Scheme 2C). The purity confirmation of ssDNAs was verified on an agarose gel (Figure 3). The pure ssDNAs (Figure 4) were used for primer extension (Scheme 2E and Figure 5 in Scheme 3) to incorporate the 44mer oligonucleotides containing an aristolactam adducts prepared in the laboratory of Francis Johnson. After primer extension, the samples were loaded on an 1% agarose gel in Figure 5 to visualize the desired product (Scheme 3). As shown in lane 2 and 3, some dsDNAs was formed and the CsCl gradient purification

Table 2

Substrates	Sequences
NarI-AAF	5'-GGC**G(AAF)**CC-
NarI-AF	5'-GGC**G(AF)**CC-
C209	5'-ACAGA-
C209-ALII	5'-AC**A(ALII)**GA-
C209 with mismatch	5'-ACACA- -TCCCT-5'
C209-ALII with mismatch	5'-AC**A(ALII)**CA- -TCCCT-5'
SI6	5'-CTAGG-
SI6-ALII	5'-CT**A(ALII)**GG-
SI6 with mismatch	5'-CCACG- -GCCCC-5'
SI6-ALII with mismatch	5'-CC**A(ALII)**CG- -GCCCC-5'

Table 2. Sequences of prepared oligonucleotides used in this study. AAF (dG-AAF): *N*-(deoxyguanosin-8-yl)-2-acetylaminofluorene, AF (dG-AF): *N*-(deoxyguanosin-8-yl)-2-aminofluorene, ALII (dA-ALII): 7-(deoxyadenosin-N^6-yl) aristolactam II (R=H).

(Scheme 2F) was performed. Then dsDNA was purified, but there were some impurities such as nicked and linear DNAs above the cccDNA in lanes 1 and 2 (c 209) as well as lane 3 and 4 (SI6) of Figure 4. To remove these impurities, another purification step, sucrose gradient, was performed with the two samples, yielding good separation of the cccDNA and nicked or linear DNA (Figure 7 in scheme 3). The pure cccDNAs were collected and concentrated by Amicon Ultra from Millipore and used in *in vitro* NER assays.



Scheme 2: A scheme of preparation of ssDNA and plasmis for NER assays (A: double digestions, B: ligation with T4 DNA ligase, C: preparation of ssDNA with R408 helper phage, D: annealing, E and F: primer extension and purifications, G: NER assay)

Codon 209(C209):	5' d(CCCTAGTACGATGACAGAAACACTGC) 3'
Comp. Strand of Codon:	5' d(GCACGCAGTGTTTCTGTCATCGTACT) 3'
Splice Site, intron 6(SI6):	5' d(CCCTAGTACTCTCCTAGGTTGGCTCGC) 3'
Comp. Strand of Splice Site:	5' d(GCACGCGAGCCAACCTAGGAGAGTACT) 3'

Table 3: Sequences of inserted primers for C209 and SI6 (Cloning).

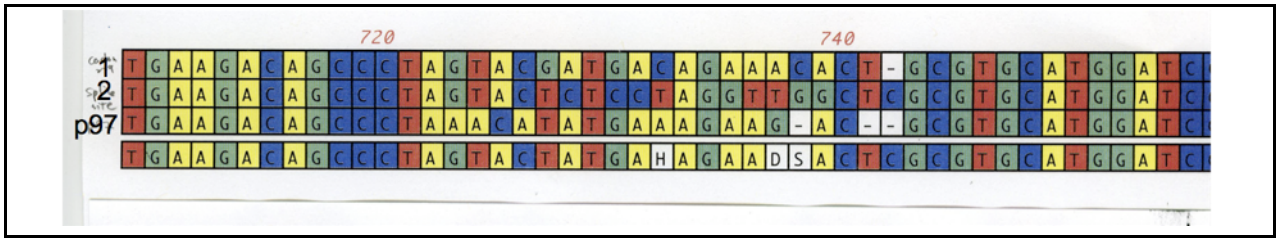


Figure 3: Sequencing data showing only inserted parts for newly prepared plasmids and p97 (an original plasmid), 1:C209-CCCTAGTACGATGACAGAAACACTGC, 2:SI6- CCCTAGTACTCTCCTAGGTTGGCTCGC.

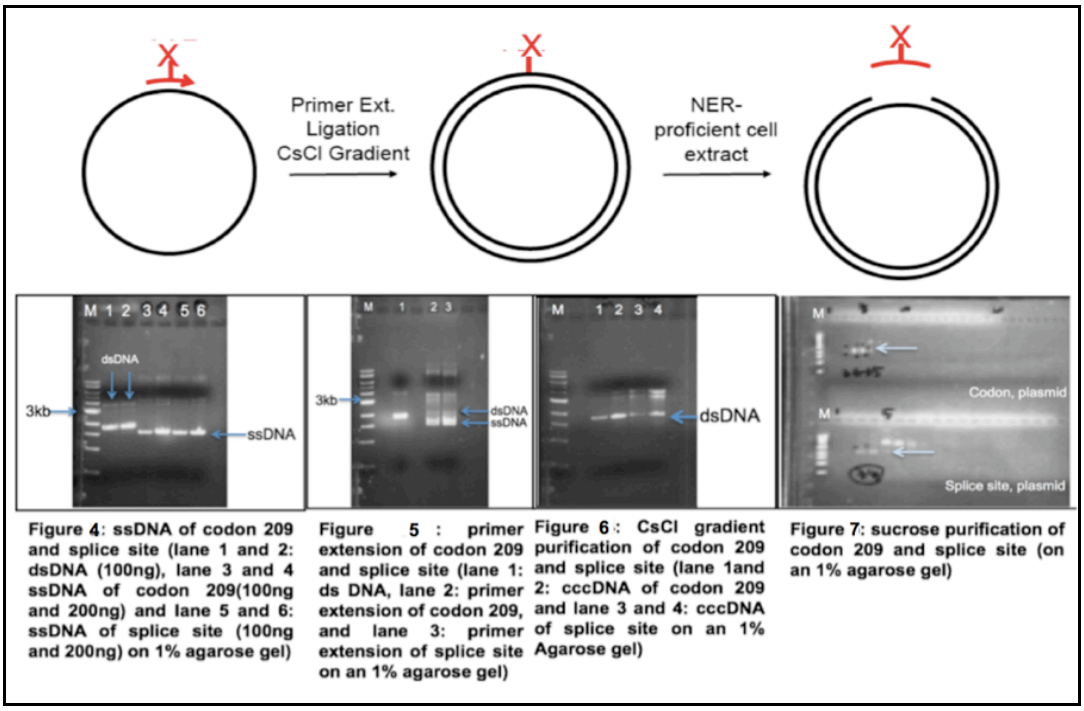


Figure 4: ssDNA of codon 209 and splice site (lane 1 and 2: dsDNA (100ng), lane 3 and 4 ssDNA of codon 209(100ng and 200ng) and lane 5 and 6: ssDNA of splice site (100ng and 200ng) on 1% agarose gel)

Figure 5 : primer extension of codon 209 and splice site (lane 1: ds DNA, lane 2: primer extension of codon 209, and lane 3: primer extension of splice site on an 1% agarose gel)

Figure 6 : CsCl gradient purification of codon 209 and splice site (lane 1and 2: cccDNA of codon 209 and lane 3 and 4: cccDNA of splice site on an 1% Agarose gel)

Figure 7: sucrose purification of codon 209 and splice site (on an 1% agarose gel)

Scheme 3: A scheme of preparation of ssDNAs and plasmids with figure 4-7 (X= Aristolactam, M=marker, ssDNA and dsDNA are marked on the gels.)

Preparation of New Plasmids for Preparation of ssDNA (C209 and SI6 with Mismatch)

In order to obtain ssDNA for new plasmids, allowing to place the dA-AL adducts in C209 and SI6 with mismatch to possibly facilitate XPC-RAD23B binding and NER, p97 (pBluscript, BbsI NdeI in pBlkp II SK+, 3kb) was used as a template. p97 was doubly digested with BbsI to give sticky ends to accommodate new oligonucleotides containing new desirable sequences (Table 2 and 4). The gel extraction of doubly digested p97 with BbsI was performed and the pure linear DNA (Scheme 2A) was ligated with the annealed oligonucleotides to introduce new sequences accommodating oligonucleotides containing aristolactam adducts in a mismatch (Scheme 2B). The new plasmids were sequenced (Figure 8) and used as a template to provide ssDNA with R408 helper phage (Scheme 2C), and the purity was verified on an agarose gel (Figure 9 in scheme 4). The ssDNA was used in primer extension reaction (Scheme 2E and Figure 10 in scheme 4) to introduce the oligonucleotides containing an dA- AL adducts, and samples were loaded on an 1% agarose gel to visualize the desire product, dsDNA

(Figure 10) and subjected to CsCl gradient purification. dsDNAs were purified but there were some impurities, nick and linear DNA (Figure 11), that were removed by sucrose gradient (Figure 12) and displayed good separations between cccDNA and nicked or linear DNA. The pure cccDNAs were collected and concentrated by Amicon Ultra from Millipore.

C209 with mismatch:	5' CCCTAGTACGATGAGGGAAACACTGC 3'
SI6 with mismatch:	5' CCCTAGTACTCTCCGGGGTTGGCTCGC 3'
Comp.Strand of C209 with mismatch:	5'GCACGCAGTGTTTCCCTCATCGTACT 3'
Comp.Strand of SI6 with mismatch:	5'GCACGCGAGCCAACCCCGGAGAGTACT 3'

Table 4: Sequences of inserted primers for C209 and SI6 with mismatch (Cloning).

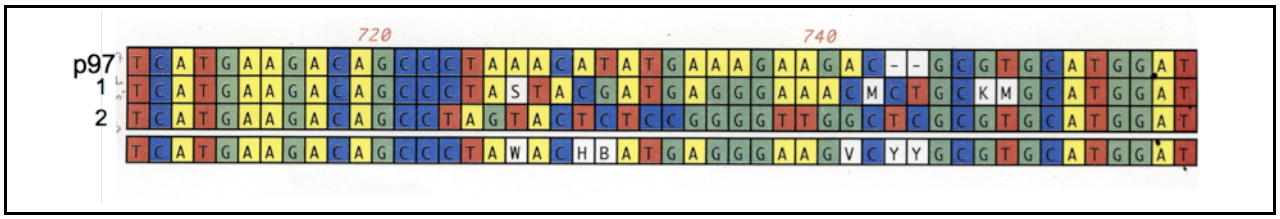


Figure 8: Sequencing data for newly prepared plasmids. (1: C209 with mismatch, CCCTAGTACG ATGAGGGAAACACTGC, 2: SI6 with mismatch: CCCTAGTACTCTCCGGGGTTGGCTCGC)

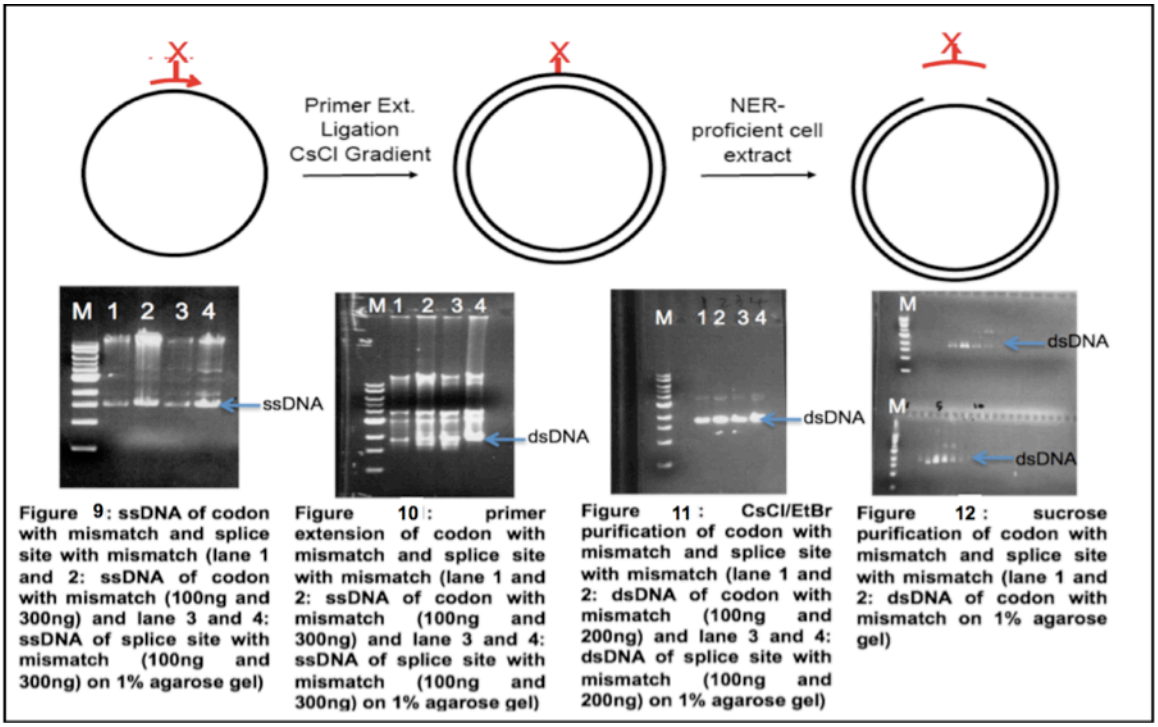


Figure 9: ssDNA of codon with mismatch and splice site with mismatch (lane 1 and 2: ssDNA of codon with mismatch (100ng and 300ng) and lane 3 and 4: ssDNA of splice site with mismatch (100ng and 300ng) on 1% agarose gel)

Figure 10: primer extension of codon with mismatch and splice site with mismatch (lane 1 and 2: ssDNA of codon with mismatch (100ng and 300ng) and lane 3 and 4: ssDNA of splice site with mismatch (100ng and 300ng) on 1% agarose gel)

Figure 11: CsCl/EtBr purification of codon with mismatch and splice site with mismatch (lane 1 and 2: dsDNA of codon with mismatch (100ng and 200ng) and lane 3 and 4: dsDNA of splice site with mismatch (100ng and 200ng) on 1% agarose gel)

Figure 12: sucrose purification of codon with mismatch and splice site with mismatch (lane 1 and 2: dsDNA of codon with mismatch on 1% agarose gel)

Scheme 4: A scheme of preparation of ssDNAs and plasmids with figure 9-12.

dA-ALII Adducts Are Refractory to Repair by GG-NER in Vitro

We next sought to investigate whether dA-ALII adducts are repaired by GG-NER in vitro or not. If they are refractory to NER, we would develop an understanding of why they are not repaired. We chose to investigate NER of dA-ALII in the context of hot spots for mutagenesis in the p53 of patients with upper urothelial cancer and likely sites with inefficient GG-NER [88]. Oligonucleotides containing dA-ALII at codon 209 (AC**A**GA) or at a splice site in Intron 6 (CT**A**GG) (see Table 2) were prepared [79] and introduced into plasmids for *in vitro* NER assays. These plasmids were incubated with an NER proficient HeLa cell extract and the products, expected to be about 25-30 nucleotides in length, were detected after annealing to complementary oligonucleotide with an overhang of 4G and a 'fill-in' labeling reaction with ^{32}P-dCTP [60, 62]. We compared the repair activity on the dA-ALII substrates to that of dG-AAF, a good NER substrate, and an inefficient NER substrate, dG-AF. Consistent with our previous observations (JYE and ODS, manuscript in preparation), dG-AAF was processed efficiently in the extracts, yielding the characteristic multi band pattern, while the dG-AF susbstrate was repaired with much lower efficiency (Figure 13, lanes 1 and 2). In contrast, no product was observed for the two dA-ALII containing plasmids (C209-ALII and SI6-ALII) (Figure 13, lane 3 and 5), indicating that dA-ALII adducts were not repaired by GG-NER in cell extracts.

We then wished to understand whether this absence of NER was due to a lack of distortion induced by dA-ALII lesions or not. Previous study has shown that some lesions, such as CPDs, are inefficiently repaired by NER, but that the NER activity can be dramatically increased by placing the lesion in the context of a mismatch [11, 47]. To test if this was also the case for dA-ALII, we placed dA-ALII in the C209 and SI6 sequences in the context of a CCC mismatch and measured whether this additional distortion helped to trigger NER activity. This was indeed the case (Figure 13, lanes 4 and 6) and the repair efficiency of the dA-ALII adducts in the context of a mismatch was similar to that of dG-AAF.

As is often observed during NER in vitro assays, the pattern of the product bands was different for the dG-AAF and dA-ALII adducts and we wished to confirm that all the

observed activity is indeed due to GG-NER. We, therefore, performed additional NER assay using extracts (XP-F) from cell lines with a deficiency in XPF. We have previously shown that these extracts are devoid of NER activity, but activity can be restored by adding recombinant purified ERCC1-XPF [21]. No NER-specific bands were observed for any of the substrates upon incubation with the XP-F extract alone (Figure 14, odd lanes), but addition of ERCC1-XPF restored robust NER activity to dG-AAF plasmid (Figure 14, lane 2) and the dA-ALII reactions where they were placed in a mismatch (Figure 14, lanes 8 and 12). These studies reveal that dA-ALII adducts are specifically repaired by NER only if they are placed in a context with additional distortion.

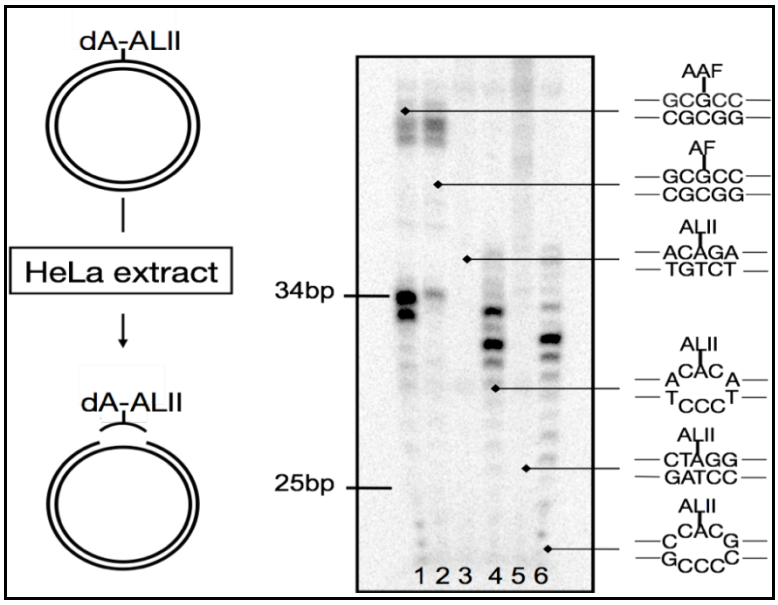


Figure 13: NER assay with Codon and Splice site sequences with ALII and with or without mismatch in the presence of HeLa cell extract. Plasmids containing site-specific dG-AAF, dG-AF or dA-ALII residues were incubated with HeLa cell extract. The 24mer to 32mer excision products containing dG-AAF, dG-AF or dA-ALII were detected by annealing to complementary oligonucleotides with a 5'-GpGpGpG overhang, which served as a template for end-labeling with [α-^{32}P] dCTP with sequenase. The reaction products were resolved on a 14% denaturing polyacrylamide gel. A low molecular weight ladder from NER was used as a size marker and the position of 25 and 34 nt bands are indicated. (lane 1: a plasmid of *Nar*I-3AAF, lane 2: a plasmid of *Nar*I-3AF, lane 3: a plasmid of codon with ALII, lane 4: a plasmid of codon with ALII on mismatch, lane 5:a plasmid of splice site with ALII, and lane 6:a plasmid of splice site with ALII on mismatch)

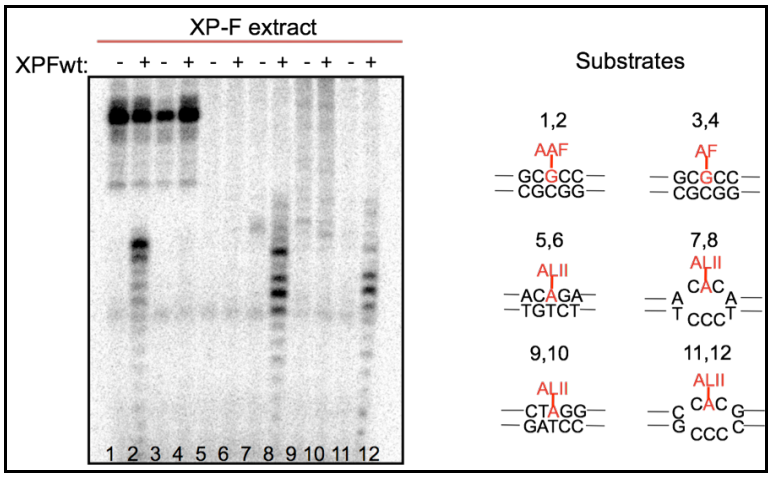


Figure 14: NER assay with Codon and Splice site sequences with ALII and with or without mismatch in presence of XP-F cell extract. Plasmids containing site-specific dG-AAF, dG-AF or dA-ALII residues were incubated with XP-F with/without XPFwt. The 24mer to 32mer excision products containing dG-AAF, dG-AF or dA-ALII were detected by annealing to complementary oligonucleotides with a 5'-GpGpGpG overhang, which served as a template for end-labeling with [α-^{32}P] dCTP with sequenase. The reaction products were resolved on a 14% denaturing polyacrylamide gel. (lane 1 and 2: a plasmid of *Nar*I-3AAF, lane 3 and 4: a plasmid of *Nar*I-3A, lane 5 and 6: a plasmid of codon with ALII, lane 7 and 8: a plasmid of codon with ALII on mismatch, lane 9 and 10 :a plasmid of splice site with ALII, and lane 11 and 12 :a plasmid of splice site with ALII on mismatch).

dA-ALII Adducts Do Not Exhibit Specific Binding to XPC-RAD23B in Duplex DNA

XPC-RAD23B is responsible for recognition of DNA lesions in GG-NER and for the subsequent recruitment of downstream NER factors [4]. It is generally believed that the thermodynamic stabilization induced in a DNA duplex influenced by parameters such as base-pairing disruption, bending, and flexibility is correlated with efficiency of XPC-RAD23B binding and overall NER efficiency [54, 55]. We investigated whether XPC-RAD23B binding affinity correlated to efficiency of GG-NER *in vitro* for dA-ALII adducts in the fully matched and mismatched duplexes used for NER experiments. 44mer oligonucleotides containing a dA-ALII residue in the same sequence context, which was used in the NER experiments (Table 2), were annealed to 5'-Cy5-labeled complementary strands and used in EMSA assays with XPC-RAD23B. These substrates were incubated with different concentrations of XPC/RAD23B in the presence of the three-fold excess of a competitor [11, 90] and bound and unbound fractions were separated on a 5% native PAGE gel.

Comparison of the binding affinities of C209 and C209-ALII revealed that the presence of dA-ALII had no effect on binding affinity of XPC-RAD23B (Figure 15, Table

5). However, when C209 and C209-ALII were placed opposite to a mismatch, the *Kd* values of both substrates were decreased to similar levels by about two fold, showing that the presence of dA-ALII did not affect the binding affinity of XPC-RAD23B. However, introducing a 3 base pair mismatch in the C209 and C209-ALII sequences led to a two-fold improvement of the binding affinity (Table 5) similar to the binding affinity of dG-AAF. Similar trends were observed in binding experiments with the SI6 sequence, with the exception that the C209-dA-ALII oligonucleotide had a binding affinity to XPC-RAD23B intermediate of that of C209 and the mismatch-containing sequences. Overall the binding affinity of XPC-RAD23B was proportional to the efficiency of NER, suggesting that dA-ALII adducts in DNA are refractory to GG-NER, because they fail to bind the damage sensor XPC-RAD23B with higher affinity than non-damaged DNA.

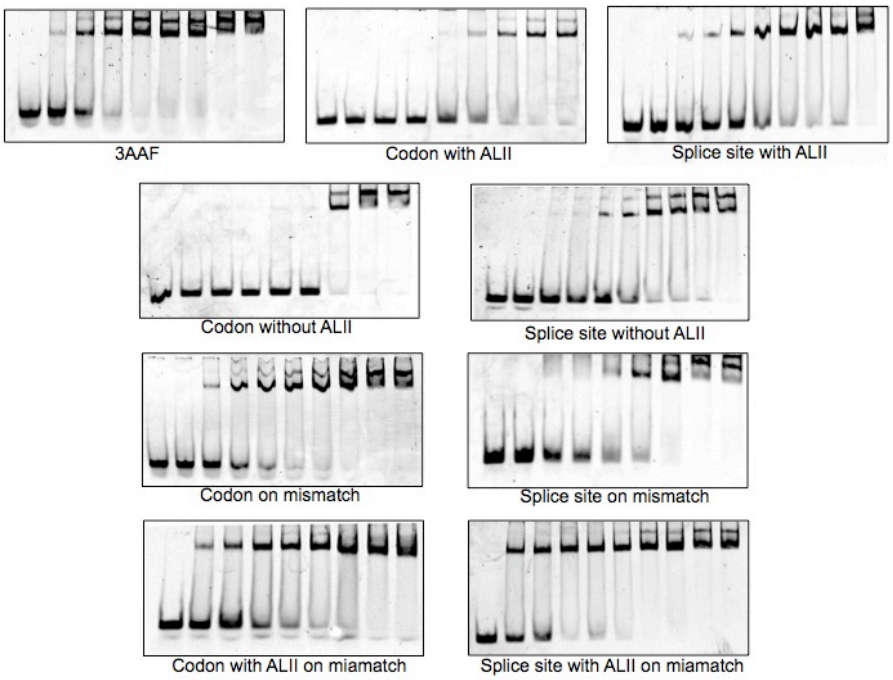
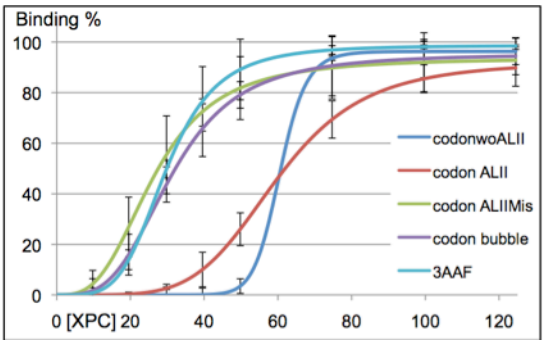
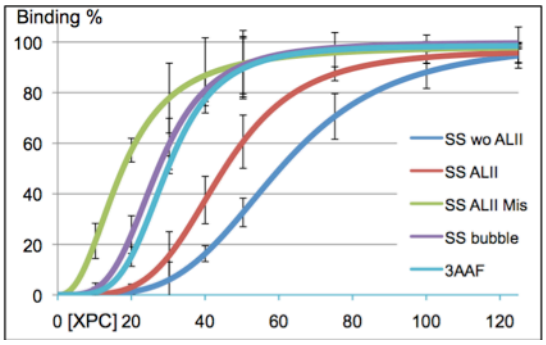


Figure 15: EMSA to measure binding affinity of XPC/RAD23B to a modified DNA. 44mer oligonucleotides containing a site-specific lesion were annealed to complementary oligonucleotides containing 5' Cy5. The annealed substrates were incubated with different concentrations of XPC/RAD23B (0-125nM) for 30min in proper binding buffer. The reactions were loaded on to a 5% native polyacrylamide gel. Shifted bands (DNA+XPC/RAD23B) and non-shifted bands (DNA) were quantified by image quantification program.





Substrate	3AAF	C209 without ALII	C209 with ALII	C209 with ALII on mismatch	C209 without ALII on Mismatch
K_d (nM)	30±3	61±2	60±4	26±11	31±6
Substrate	3 AAF	SI6 without ALII	SI6 with ALII	SI6 with ALII on mismatch	SI6 without ALII on Mismatch
K_d (nM)	30±3	60±4	45±5	18±8	27±4

Table 5: Sigmoid curves and K_d values from Sigmaplot software. DNA binding was analyzed using electrophoretic mobility shift assays. 44mer oligonucleotides containing a site-specific lesion were annealed to complementary oligonucleotides containing 5' Cy5. The annealed substrates were incubated with different concentrations (0-125nM) of XPC/RAD23B for 30min in binding buffer. The reactions were loaded on to a 5% native polyacrylamide gel. Shifted bands (DNA+XPC/RAD23B) and non-shifted bands (DNA) were quantified by image quantification program. K_d values were calculated from sigma plot program and graphs were modified by sigma plot program. Graphs show the quantification of three independent experiments. Standard deviations are indicated by error bars.

Discussion

Aristolochic acids (AA) are associated with AA nephropathy and urothelial cancer. Upon metabolic activation, AA reacts with DNA forming the mutagenic dA-aristolactam (AL) and dG-AL adducts. The previous study supported that mutations in the *TP53* gene in human urothelial cancers induced by AA are limited to the non-transcribed strand, suggesting that transcription coupled nucleotide excision repair (NER), not global genome NER, efficiently removed dA-AL adducts.

To survey the involvement of TC-NER to repair dA-AL adducts, a set of xeroderma pigmentosum proteins (XP)- or Cockayne syndrome proteins (CS)-deficient, complemented and normal human fibroblasts were used to investigate the toxic effect of AAII. We observed more sensitive and increase amount of AA-DNA adduct in most XP-

deficient cells than the control and XPC deficient cell lines. In case of CS-deficient cell lines, the same level of sensitivity to AAII as for the control cell was detected, but the AAII adduct levels were higher. For CSB-complemented cells, the adduct level was similar to the normal cell line and they were more sensitive than the normal fibroblasts. Results determining the levels of adduct *in vivo* support the contribution of TC-NER to repair AA adducts, particularly dA-ALII.

In addition, we have generated plasmids containing a site-specific dA-ALII lesion and have studied how these adducts are repaired by NER. From *in vitro* NER assay, we confirmed that dA-ALII adduct is not repaired by GG-NER unless it is placed in a mismatch. These results support that dA-ALII does not create enough distortion in DNA duplexes and the additional distortion created by a 3 base mismatch is able to restore NER activity. This finding is consistent with the structural studies in which NMR and thermodynamic studies have shown that dA-ALII causes minor destabilization on DNA duplex (Prof. Santos' data, not published). The binding affinity of XPC-RAD23B showed the proportional correlation to the repair efficiency. These overall results suggest that a lack of destabilization by dA-ALII prevents binding by XPC-RAD23B and repair by NER.

MATERIALS AND METHODS

Reagents and equipments

^{32}P-ATP (6000 Ci/mmol) was purchased from PerkinElmer (Boston, USA). C*is*-diamminedichloroplatinum(II), cisplatin, was obtained from Sigma Company (St. Louis, MO, USA), dissolved at 10 mM in DMSO and stored at -80°C. Aristolochic acid II with the purity of >97% was synthesized by Francis Johnson group (Stony Brook University, Stony Brook, NY), dissolved at 25 mM in DMSO and stored at -20°C. The concentration of AAII was verified by absorption in ethanol at 250 nm. Enzymes for post-labelling were obtained from Worthington (Newark, NJ, USA) and Sigma Company. T4 PNK, T4 polymerase, and T4 DNA ligase were from New England Biolabs. The sequenase enzyme (T7 DNA polymerase) was purchased from USB and [α-^{32}P] dCTP is from PerkinElmer. Complementary oligonucleotides for NER assay and 44mer oligonucleotides for EMSA were purchased from Integrated DNA Technologies (IDT).

All other chemicals were purchased from commercial sources and were of analytical grade.

Cell lines

Table below shows cell lines used for the cytotoxicity and genotoxicity assays. The XP- cell lines were obtained from the Coriell repository. Cs1an/pc3.1, Cs1an/p3.1-CSBwt and Cs1anE7/p3.1-CSBE646Q (CSB deficient and CSB corrected cell lines, CSB, CSB^{+}, CSB^{mut}, respectively) were kindly provided by Wilhelm Bohr (National Institute of Aging, Baltimore, MD). All cell lines were SV-40 transformed.

Cell	Phenotype	Name	expected GGR status	expected TCR status
GM00637	wild-type	Control	+	+
Cs1an/p3.1	CSB	CSB	+	-
Cs1an/p3.1-CSBwt	CSB-corrected	CSB^{+}	+	+
Cs1anE7/p3.1-CSBE646Q	CSB-mutant corrected	CSB^{mut}	+	-
GM16094	CSA	CSA	+	-
GM15983	XPC	XPC	-	+
GM04312	XPA	XPA	-	-
GM15876	XPA corrected	XPA^{+}	+	+
GM08437	XPF	XPF	-	-
GM14930	XPG	XPG	-	-
GM08207	XPD	XPD	-	-
GM15877	XPD corrected	XPD^{+}	+	+

Table 1: Cell lines used in the study.

Cell culture and chemical exposures

Cells were grown in 75 cm^2 flasks containing Dulbecco's modified minimum essential medium (high glucose DMEM) supplemented with fetal bovine serum (10%), penicillin (100 units/ml) and streptomycin (100mg/ml). The medium for Cs1an and XPD

complemented cell lines contained geneticin (0.4 mg/ml). Cultures were maintained at standard conditions in a humidified atmosphere containing 5% CO_2 at 37°C. Cultures were split in a 1:5 ratio every third to fourth day. Before the treatment the cells were passed in a 1:6 ratio and allowed to grow 3-4 days up to the confluence. 5 ml of the culture split was grown on 60 mm plates until the exposure and DNA isolation for the genotoxicity analysis, and 0.5 ml culture were maintained on 24-well plates prior to cytotoxicity assays.

Before the treatment the cultures were washed thoroughly with Dulbecco's phosphate buffered saline (PBS, without $CaCl_2$ and $MgCl_2$) and 5-200 mM AAII or 1-10 mM cisplatin diluted in DMEM without supplements were added to the cells with the following incubation at standard conditions. The Cs1an and XPD complemented cell lines were treated in the presence of geneticin (0.4 mg/ml). The cell lines were exposed to AAII or cisplatin for 48 hours for the adduct analysis and 72 hours to AAII for the cytotoxicity studies.

Cell viability assay

To measure cell viability, ATP concentrations were assayed using a FLASC kit from Sigma. After AAII exposure, the cells on 24-well plates were washed with PBS, and 400 ml lysis buffer (FL-SAR) was added to the each well with the following shaking at room temperature for 10 minutes. Luciferase (FL-AAM) was diluted 20Xfold in the FL-AAB buffer and the luminescence of 5 ml sample in 100 ml enzyme was measured on a luminometer (TD-20/20 *Luminometer*, Turner Designs). The ATP levels in treated cells for each chemical dose were adjusted to untreated control using the following ratio $ATP_{treated}/ATP_{untreated}$. The background readings of the buffer alone were no more the 0.1% readings of the samples. The measurements were done for at least two independent treatments in three different wells for each exposure. The results are presented as the dependence of adjusted ATP level from the AAII dose. Plots were analyzed using Sigma Plot v8.0 (SPSS Inc.).

^{32}P-post-labelling adduct analysis with PAGE electrophoresis

Analysis of DNA adduct levels is described [86, 89]. 15 fmol of each of the following 24-mer oligonucleotides were used as standards. In 5 mg DNA that corresponds to 1 adduct/10^6 nucleotides.

5'-TCT TCT TCT GTG CXC TCT TCT TCT-3' X = dA-ALII
5'-TCT TCT TCT GTX CAC TCT TCT TCT-3' X = dG-ALII

Oligonucleotides were synthesized by Francis Johnson group (Stony Brook University, Stony Brook, NY).

Briefly, DNA was isolated using the DNeasy Blood and Tissue Qiagen kit (Cat. N 69506, QIAGEN Sciences, MD, USA). 5-20 mg DNA were digested at 37°C overnight with the following nuclease P_1 (Sigma) treatment for 1 hour and adduct enriched by butanol extraction. Butanol extracted adduct nucleosides were dried down, ^{32}P-labelled with 20 mCu g^{32}P-ATP and 10 U of OptiKinase (Affymetrix Inc.). After drying down 5 ml of formamide-EDTA buffer with xylencyanol and bromphenol blue was added. The samples were loaded on 30% nondenaturating acrylamide gel prerun for 40 minutes at 1600V. The following running buffer was used, 200 mM boric acid, 2.5 mM EDTA, 100 mM Tris-base, pH 7.0. The gel was allowed to run for 4 hours at 1800V and then visualized by phosphorimaging. To estimate amount of adducts Image QuaNT v*5.2* (Molecular Dynamics) was used. The dependencies of adduct levels from AAII dose were plotted and analyzed using Sigma Plot v8.0 (SPSS Inc.).

XPC-RAD23B purification and HeLa whole-cell extract preparation

Polyhistidine-tagged RAD23B was expressed in E. coli BL21(DE3)LysS using the expression vector pET-24d and purified on nickel beads (Qiagen) as described. Polyhistidine XPC was expressed in Sf9 cells using the expression vector pFastBac1. The cells were lysed as described [57, 64] and S3 was combined with partially purified RAD23B. The correctly folded heterodimer was further purified through nickel beads (Qiagen), gel filteration (Pharmacia) and heparin (Amersham) columns. Expression and purification were performed by Adebanke Fagbemi. The HeLa whole-cell extracts was prepared as described [65].

Preparation of Oligonucleotides

5'-TCT TCT TCT GTG CXC TCT TCT TCT-3' X = dA-ALII
5'-TCT TCT TCT GTX CAC TCT TCT TCT-3' X = dG-ALII

Oligonucleotides were synthesized by Francis Johnson group (Stony Brook University, Stony Brook, NY, [79]).

The 44mer oligonucleotides containing dA-AL II were prepared as described [79] (Sequences below), purified by HPLC and characterized by ESI-MS using a Micromass Platform LC/MS. The 24mer/44mer oligonucleotides containing a site-specific dG-AAF/AF lesion were were prepared and purified as described [60].

The sequences of the oligonucleotides containing dA-ALII and dG-AAF/AF were as follows:

Substrates	Sequences (5'-3')
C209-ALII	d(AGACAGCCCTAGTACGATGACA(ALII)GAAACACTGCGTGCATGGATCC)
C209-ALII with mismatch	d(AGACAGCCCTAGTACGATGACA(ALII)CAAACACTGCGTGCATGGATCC)
SI6-ALII	d(AGACAGCCCTAGTACTCTCCTA(ALII)GGTTGGCTCGGTGCATGGATC)
SI6-ALII with mismatch	d(AGACAGCCCTAGTACTCTCCCA(ALII)CGTTGGCTCGCGTGCATGGATC)

Substrates	Sequences (5'-3')
24mer NarI-AAF	d(CTATTACCGGCG(AAF)CCACATGTCAGC)
24mer NarI-AF	d(CTATTACCGGCG(AF)CCACATGTCAGC)
44mer NarI-AAF	d(CCCTAGCTAGAGCTACGTAGCTATTACCGGCG(AAF)CCACATGTCAGC)
44mer NarI-AF	d(CCCTAGCTAGAGCTACGTAGCTATTACCGGCG(AF)CCACATGTCAGC)

Generation of plasmids for In Vitro NER Assay

The following sequences were cloned into pBluescript II SK+ were cloned into pBluescirpt II SK+ for *in vitro* NER assay.

Insertions	Sequences
C209	5'-d(CCCTAGTACGATGACAGAAACACTGC)
Complementary strand of C209	5'-d(GCACGCAGTGTTTCTGTCATCGTACT)
C209 with mismatch	5'-d(CCCTAGTACGATGAGGGAAACACTGC)
Complementary strand of C209 with mismatch	5'-d(GCACGCAGTGTTTCCCTC ATCGTACT)
SI6	5'-d(CCCTAGTACTCTCCTAGGTTGGCTCGC)
Complementary strand of SI6	5'-d(GCACGCGAGCCAACCTAGGAGAGTACT)
SI6 with mismatch	5'-d(CCCTAGTACTCTCCGGGGTTGGCTCGC)
Complementary strand of SI6 with mismatch	5'-d(GCACGCGAGCCAACCCCGGAGAGTACT)

An aliquot of 100pmol of the 44mer oligonucleotides was 5'-phosphorylated by incubation with 20 U of T4 PNK enzyme and 2mM of ATP for 2h. After annealing with 31pmol of single-stranded modified pBluescript II SK^+ (described in [60]), further incubation with dNTPs, T4 DNA polymerase and T4 DNA ligase (described in [62]) yielded covalently closed circular DNA containing a single ALII adduct. The closed circular DNA was purified by cesium chloride/ethidium bromide density gradient centrifugation, by consecutive butanol extractions to remove the ethidium bromide and finally concentrated on a Centricon YM-30 (Millipore). In addition to the published purification procedure [62], the plasmid was repurified by sucrose gradient centrifugation. The collected fractions after centrifugation containing the closed circular plasmid were collected and concentrated on a Centricon YM-30 (Millipore). The purified plasmids containing a site-specific dA-ALII and dG-AAF/AF lesions were aliquoted and stored at -80°C. HeLa whole-cell extracts were prepared as described in [91], aliquoted and stored at -80°C. The *in vitro* NER assay was performed using an established protocol [62]. HeLa cell extracts (21mg/ml, 2 µl), or XP-F cell extract (12.5mg/ml, 3 µl), 2 µl of 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) and either purified

ERCC1-XPF protein (0.18mg/ml, 0.5 µL) or 400mM NaCl (final NaCl concentration was 70 mM) in a total volume of 10 µl were pre-incubated at 30°C for 10 min. dA-AL or dG-AAF/AF-containing plasmid DNA plasmid (1µL of 50 ng/µL) was added and the mixture incubated at 30 °C for 45 min. Samples were placed on ice, 0.5 µl of 1 µM of a complementary oligonucleotide (see below) was added and the mixtures heated at 95 °C for 5 min. The samples were allowed to cool down at room temperature for 15 min to allow the DNA to anneal. 1 µL of a Sequenase/[α-^{32}P]-dCTP mix (0.25 units of Sequenase and 2.5 µCi of [α-^{32}P]-dCTP per reaction) was added, the mixture incubated at 37 °C for 3 min, 1.2 µl of dNTP mix (100 µM of each dATP, dTTP, dGTP; 50 µM dCTP) added and the mixture incubated for another 12 min. The reactions were stopped by adding 12 µl of loading dye (80% formamide/10 mM EDTA) and heating at 95 °C for 5 min. The samples were run on a 14% sequencing gel (0.5x TBE) at 45 W for 2.5 hrs. The reactions products were visualized using a PhosphorImager (Typhoon 9400, Amersham Biosciences).

Complementary strands for NER assay	Sequences
NarI-AAF/AF	3'-phophorylated oligonucleotide: d(GGGGAGTGTTTCTGTCATCGTACTAGGGCTGT)
C209	3'-phophorylated oligonucleotide: d(GGGGAGTGTTTCTGTCATCGTACTAGGGCTGT)
C209 with mismatch	3'-phophorylated oligonucleotide: d(GGGGAGTGTTTGTGTCATCGTACTAGGGCTGT)
SI6	3'-phophorylated oligonucleotide: d(GGGGAGCCAACCTAGGAGAGTACTAGGGCTGT)
SI6 with mismatch	3'-phophorylated oligonucleotide: d(GGGGAGCCAACGTGGGAGAGTACTAGGGC

EMSA of XPC-RAD23B bound to dA-AL and dG-AAF/AF-containing oligonucleotides

Modified 44mer templates (4nM) were annealed to a 5' Cy5 labeled complementary strand (1.3nM) in the presence of 10mM Tris-HCl, pH8, containing 50mM NaCl and 10mM MgCl. The binding reactions were performed in 25mM Tris-HCl, pH7.5, containing, 0.1mg/mL BSA and 10% glycerol for 25 °C for 30min in a 15µL reaction

containing 4nM non-modified duplex 44mer competitor, increasing amounts of XPC-RAD23B (0-125nM), at a NaCl concentration of 40mM. The reaction mixtures were loaded onto a native 5% polyacrylamide gel pre-equilibrated with 0.5 X TBE buffer, and run at 4 °C for 50 min at 20 mA. Gels were scanned using a Typhoon 9400 imager. The intensities of the bands for free 44mer oligonucleotides and the complex formed with XPC-RAD23B were determined by Image Quant TL program from Amersham Biosciences. Each reaction was performed three times to determine the binding constant.

The eqilibrium dissociation constants (K_d) for the XPC-RAD23B bound to the 44mer duplexes containing ALII or AAF/AF were determined by an equation, $y=[L]^n / (Kd+[L]^n)$, (Y= % of protein-DNA complex and L: concentration (nM) of XPC (protein), Hill equation and n: Hill coefficient, describing cooperativity). *Kd* values are determined by a SigmaPlot program v10.0 by fitting the data to sigmoid curves. The sequences of the oligonucleotides used in the binding assays are given below.

Substrate	Sequences for oligonucleotides used in binding assays (5'-3')
C209	d(AGACAGCCCTAGTACGATGACAGAAACACTGCGTGCATGGATCC)
C209-ALII	d(AGACAGCCCTAGTACGATGACA(ALII)GAAACACTGCGTGCATGGATCC)
Comple. Strand for C209-ALII and C209	Cy5-d(GGATCCATGCACGCAGTGTTTCTGTCATCGTACTAGGGCTGTCT)
C209 with mismatch	d(AGACAGCCCTAGTACGATGACACAAACACTGCGTGCATGGATCC)
C209-ALII with mismatch	d(AGACAGCCCTAGTACGATGACA(ALII)CAAACACTGCGTGCATGGATCC)
Comple. Strand for C209-ALII and C209 with mismatch	Cy5-d(GGATCCATGCACGCAGTGTTTCCCTCATCGTACTAGGGCTGTCT)
SI6	d(AGACAGCCCTAGTACTCTCCTAGGTTGGCTCGGTGCATGGATC)
SI6-ALII	d(AGACAGCCCTAGTACTCTCCTA(ALII)GGTTGGCTCGGTGCATGGATC)
Comple. Strand for SI6 and SI6-ALII	Cy5-d(GATCCATGCACGCGAGCCAACCTAGGAGAGTACTAGGGCTGTCT)

SI6 with mismatch	d(AGACAGCCCTAGTACTCTCCCACGTTGGCTCGCGTGCATGGATC)
SI6-ALII with mismatch	d(AGACAGCCCTAGTACTCTCCCA(ALII)CGTTGGCTCGCGTGCATGGATC)
Comple. Strand for SI6 and SI6-ALII with mismatch	Cy5-d(GATCCATGCACGCGAGCCAACCCCGGAGAGTACTAGGGCTGTCT)

SUPPLEMENTARY DATA

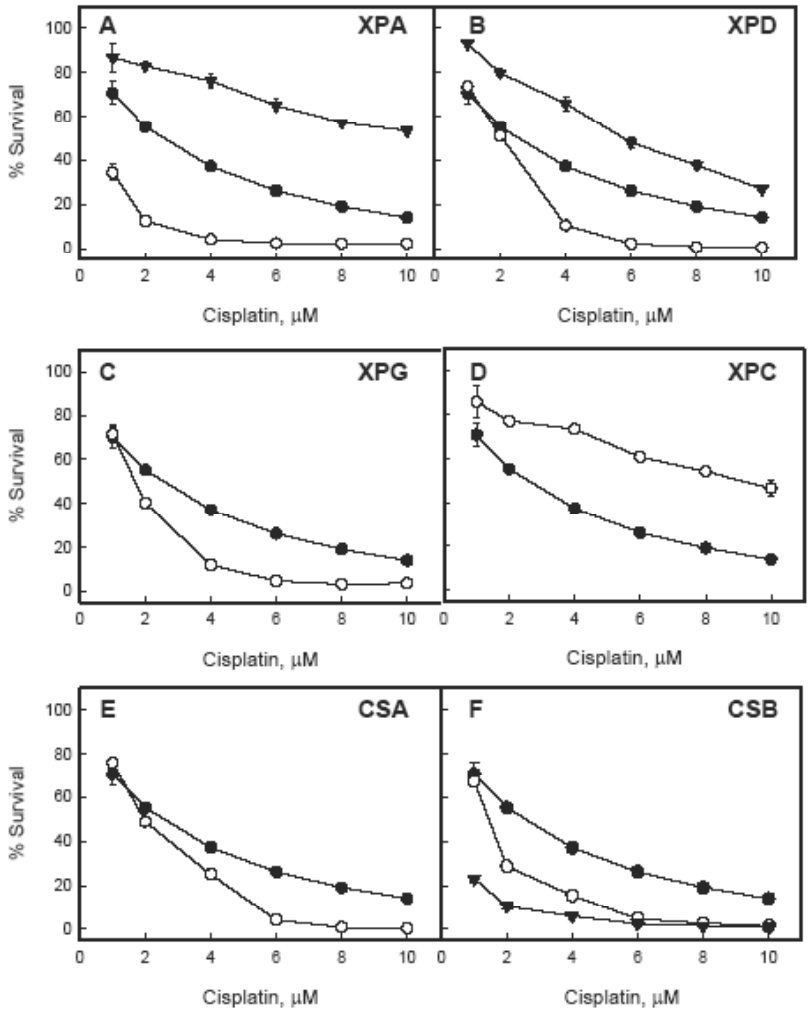


Figure 1: Cisplatin cytotoxicity study. XP-, CS- deficient and control normal human fibroblasts cell lines were treated with 1-10μM AAII for 48hours with the following ATP level measurements. Filled circles (•I indicate the control normal human fibroblasts cell line, open circles (o)- the deficient cell lines corresponding to the panel names, filled triangles – complemented cell lines if available. The point for each cisplatin does corresponding to the ATP ratio between treated and untreated cells and presents mean ± standard deviation for three wells.

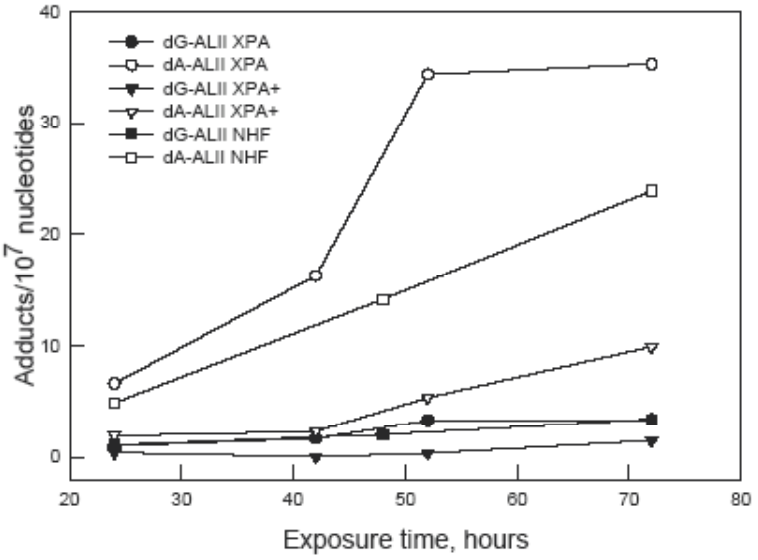


Figure 2: Adduct formation of dA or dG-ALII

CHAPTER 4:

A FLUORESCENT APPROACH TO MONITOR NER REACTIONS

ABSTRACT

The excision products of the NER pathway are typically visualized by radioactive labeling using a fill-in reaction with [^{32}P]-dCTP. Considering that a typical yield of an NER reaction is around 10%, it is important to have a method that not only detects NER products, but also unrepaired substrates or reaction intermediates. In this chapter, our efforts to monitor NER reactions by fluorescently labeling NER substrates and products are described. This approach will provide various advantages over presently used methods: The possibility to quantify reaction products and unreacted substrates, a streamlining of the process, eliminating several steps in the protocol, and a possibility to study key steps in NER. This chapter describes our studies using click chemistry to fluorescently label NER substrates containing 5-ethynylthymidine residues in vicinity of the lesion to monitor NER reactions.

INTRODUCTION

Background: Conventional NER Assay and New Ways to Monitor NER Reactions

Traditionally, NER reactions are monitored by radioactive labeling of the excised fragments (Figure 1). After incubation of a plasmid containing a lesion with an NER-proficient cell extract, the 24-32mer excision products containing a damage can be detected by annealing to a complementary oligonucleotide with a four G overhang and filling in the overhang with α-[^{32}P]-dCTP and a DNA polymerase (Figure 1). After the fill-in reaction, samples are loaded on a sequencing gel and reaction products are detected on a phosphorimager. There are several disadvantages associated with this method. First, the use of [^{32}P-dCTP] to detect the lesion-containing products is necessary, resulting in the inconvenience of working with radioactive materials [62]. Second, and more importantly, this assay is limited in that it only detects NER reaction products, since only the fragments that are excised from the plasmid, but unreacted starting materials or partially incised products are not detected. Considering that NER

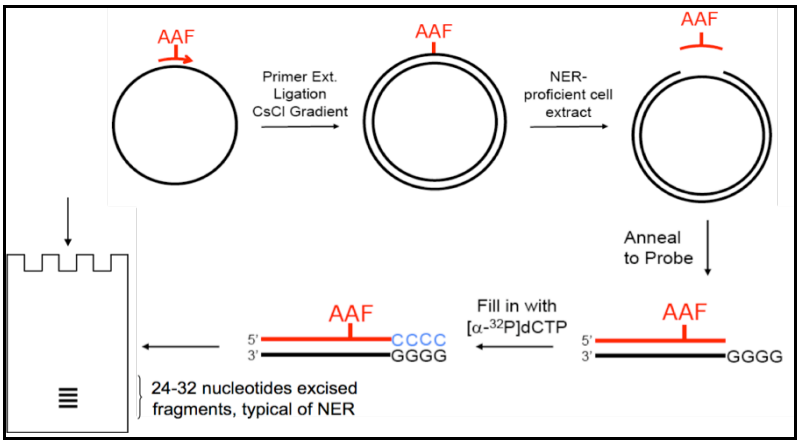


Figure 1: **A conventional NER assay to detect NER fragments.** A lesion containing plasmid was prepared by annealing of a 5' phosphorylated oligonucleotide containing a site-specific lesion to a single-strand plasmid. Primer extension was performed with T4 DNA polymerase and ligase and the product was purified by CsCl gradient purification. This plasmid containing a site-specific AAF was incubated with HeLa cell extract. The excised DNA fragments of 24-32 nucleotides in length were detected by annealing a complementary oligonucleotide containing a non-complementary 4G overhang and filling in with [α-^{32}P]dCTP. AAF: 2-acetylamino fluorene.

usually takes place with an efficiency of around 10% or less in a cell extract, this is a significant limitation. Furthermore, for the visualization of NER fragments exposure to a phosphorimager screen overnight, or even longer is necessary and the use of a

polymerase to label the products can result in the formation of non-specific bands from nicked plasmids (See the example in Figure 7 in chapter 2).

We propose to use click chemistry to monitor NER reactions by fluorescencently labeling the substrates and products. Advantages of this approach include overcoming the inconvenience of using radioactive materials, reducing exposure time of gels and the possibility to detect reaction products, substrates and intermediates. We will be employing two strategies to use click chemistry (Figure 2 and 3). In the first one, a substrate containing a 5-ethynylthymidine residue will be prepared and labeled with a fluorescent dye and used in NER assays in the labeled form and excision fragments directly monitored following the reaction (Figure 2).

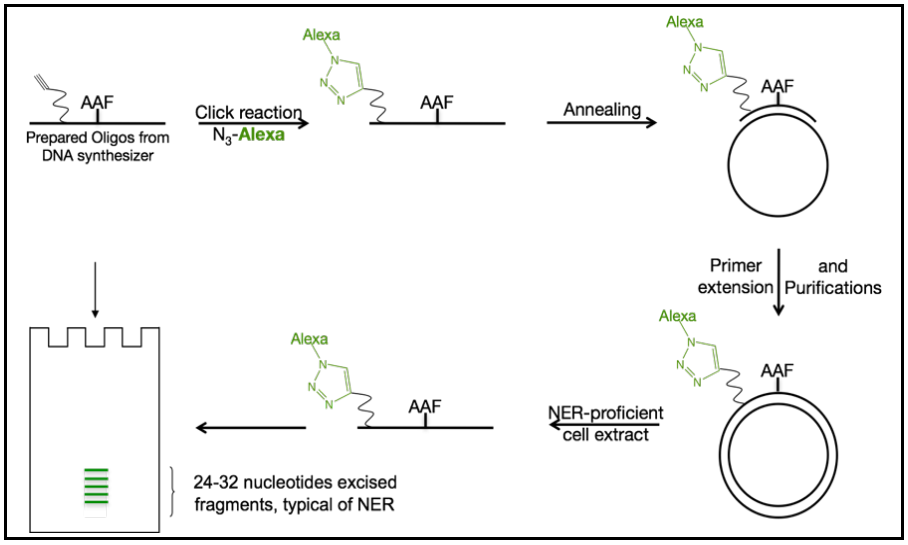


Figure 2: A scheme for the new way to monitor NER fragments – pre-labeling. An oligonucleotide containing a site-specific lesion and ethynylthymidine residue were prepared by DNA synthesizer. The prepared oligonucleotide was labeled with an azide derivative alexa 488 dye by a click reaction and then the labeled oligonucleotide was annealed to a ssDNA and then primer extension was performed with T4 DNA polymerase and ligase. The plasmid was purified over CsCl gradient purification. The plasmid was incubated with a HeLa cell extract and the excised DNA fragments of 24-32 nucleotides in length were detected by loading a reaction sample on a sequencing gel and visualization on a fluorescent imager. AAF: 2-acetyl amino fluorene.

In the second approach, fragments will be labeled after the NER reaction (Figure 3). Olignucleotides containing a lesion will be equipped with a 5-ethynylthymidine residue and incorporated into a plasmid. These plamsids will be incubated with a NER-proficient cell extract to release products containing the damage and an alkyne. These fragments will be fluorescently labeled by a click reaction following the reaction and loaded onto a sequencing gel.

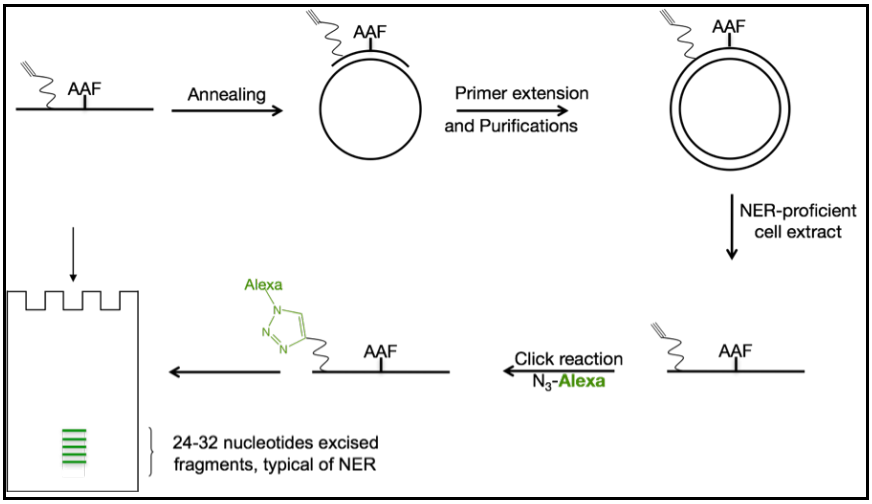


Figure 3: A scheme for the new way to monitor NER fragments – post labeling. An oligonucleotide containing a site-specific lesion and alkyne were prepared by DNA synthesizer. The prepared oligonucleotide was annealed to a ssDNA and then primer extension was performed with T4 DNA polymerase and ligase. CsCl gradient purification was achieved to have a pure plasmid. A purified plasmid containing a site-specific lesion was incubated with HeLa cell extract. The excised DNA fragments of 24-32 oligonucleotides were labeled with a alexa 488 dye by a click reaction. The labeled fragments were detected by loading a reaction sample on a sequencing gel and visualized by a fluorescent imager. AAF: 2-acetyl amino fluorene.

About 9 years ago, the copper-catalyzed azide-alkyne Huisgen type cycloaddition (CuAAC) reaction, termed click chemistry, was discovered [92, 93]. Since then it has been widely applied for a variety of purposes such as the labeling of biomolecules, nanotechnological applications, and medicinal chemistry [94-97] due to the high selectivity of the reaction between an alkyne and an azide. Importantly, azides and alkynes are not found in biomolecules making the click reaction highly bioorthogonal and specific [98]. The click reaction has been successfully used to label DNA to study DNA structure and recognition [99]. Therefore, click chemistry should provide a robust tool for the labeling of NER substrates (Figure 4). In general, a click reaction needs a copper catalyst with a proper chelating ligand to preserve the Cu^{I} oxidation state [100, 101]. It is known that copper (I) salts can lead to hydroxy radical production [102] that could result in the degradation of the DNA. In order to avoid this significant side reaction [110, 111], a number of reaction conditions have been developed for click reactions [100, 103, 104] that provide several options for preserving the integrity of DNA.

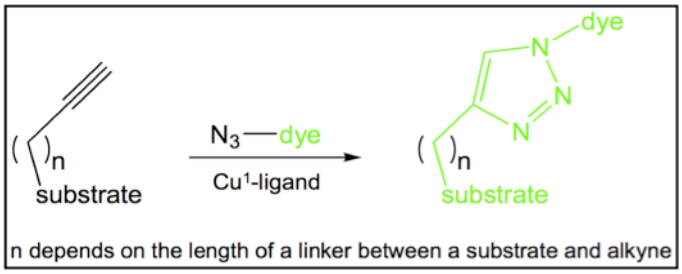


Figure 4: A scheme for a general click reaction. For general click reactions, Cu(I) or Cu (II) are used with different ligands.

Building blocks for the incorporation of substrates for a click reaction into DNA are commercially available from Glen Research and Berry & Associates. Among them, we decided to use 5-ethynyl-dU (EdU) and 5-octadiynyl-dU (OdU) (Figure 5), since they are both similar in structure to thymine and are not expected to interfere with the NER process as they do not distort the DNA [105, 106]. Of the two, 5-octadiynyl-dU has a longer linker to the alkyne, which allows for more flexibility and a more efficient click reaction [102].

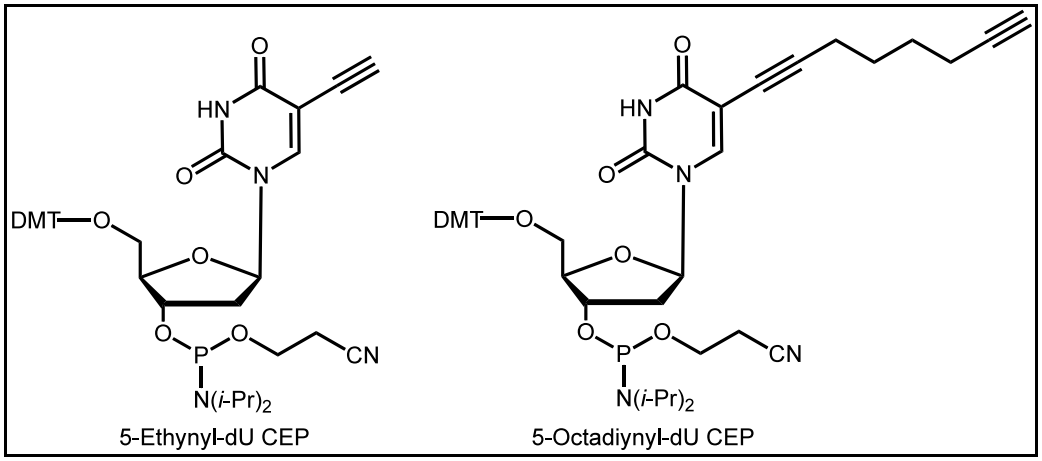


Figure 5: Structures of a 5-ethynyl-dU CEP and 5-octadiynyl-dU CEP. DMT: 4,4-dimethoxytrityl group, CEP: 2-cyano-*N,N*-diisopropylphosphoramidite.

RESULTS

Preparation of Lesion and Alkyne-containing Oligonucleotides

5-Ethynyl-2'-deoxyuridine (EdU) -2-cyanoethyl-*N*, *N*-diisopropylphosphoramidite (Figure 6) was incorporated into the position of a T residue in a *Nar*I sequence containing an AAF modification (See chapter I). Since EdU does not have significantly different structure from dT, we reasoned that it would not cause any distortion of the DNA duplex and would not interfere with the NER reaction (Figure 6). Four 24mer oligonucleotides were prepared containing EdU (Table1). Two substrates, 24mer EdU -1 and -2, do not contain an AAF lesion and are used as negative controls, while 24mer EdU AAF -1 and -2 are expected to trigger NER. All substrates were purified by standard reverse-phase HPLC and characterized by mass spectrometry (See supplementary data).

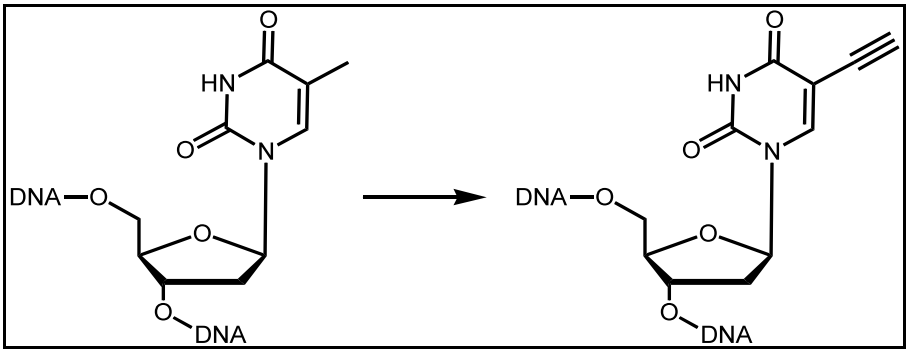


Figure 6: Structures of a thymine and 5-ethynyl-2'-deoxyuridine in oligonucleotides.

Name	Sequence
24mer EdU -1	5'-TAG CTA **U**TA CCG GCG CCA CAT GTC-3'
24mer EdU -2	5'-**U**AG CTA TTA CCG GCG CCA CAT GTC-3'
24mer EdU wAAF-1	5'-TAG CTA **U**TA CCG GC**X** CCA CAT GTC-3'
24mer EdU wAAF-2	5'-**U**AG CTA TTA CCG GC**X** CCA CAT GTC-3

Table 1: Sequences of 24mer oligonucleotides used in this study. These oligonucleoties were synthesized from expedite DNA synthesizer. X: 2-acetyl amino fluorene, U= 5-ethynyl-dU (EdU)

Click Reactions with Oligonucleotides Containing 5-Ethylthymidine

We tested several conditions for the click reaction to label these oligonucleotides (Table 1). When Cu(I) is used, chelating ligands such as bathophenanthroline and tris(benzyltriazolylmethyl)amine (TBTA) are generally used to stabilize Cu(I) oxidation state since Cu(I) is easily oxidized to Cu(II) [102-104]. Reactions using these ligands with Cu(I) provide high yields and in a short reaction time [99]. There is, however, a drawback in using Cu(I)-chelating ligands, since copper(I) is highly sensitive to oxygen, necessitating the use of an inert-atmosphere glove box [103, 104]. Other reaction conditions do not require the use of a glove box, in which Cu (II) is used with ascorbate or other water-soluble reducing agents that can reduce the oxidization state form Cu (II) to Cu (I) in situ [93, 104] and tris-(carboxyethyl)phophine hydrochloride (TCEP) is used to maintain Cu (I) oxidation state [100, 106].

Several reaction conditions with Cu (I) or Cu (II) were tested to label the oligonucleotides (Figure 7). In the first reaction, 24mer EdU No AAF-1 was incubated with CuBr and bathophenanthroline in Tris buffer (pH8) under rigorous exclusion of

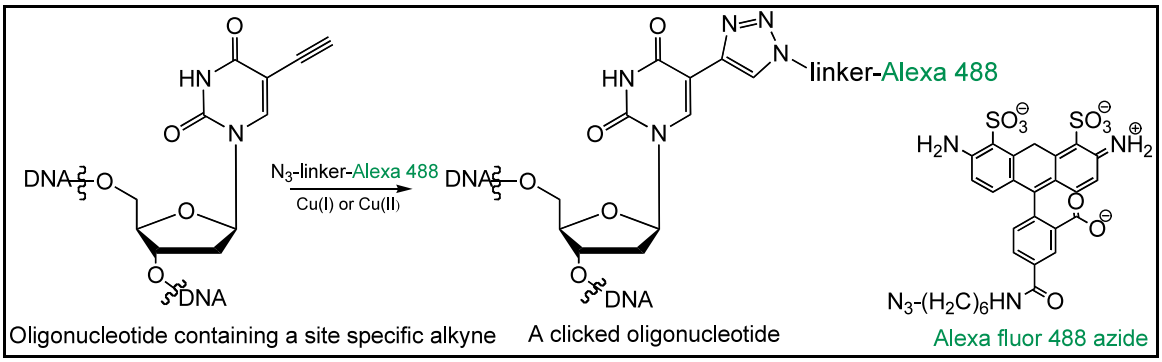


Figure 7: A click reaction of an oligo with Alexa488.

oxygen [107] in a glove box. After a 12 hr reaction, samples were heat-denatured and analyzed by 20% denaturing PAGE. Lane 1 in Figure 8A shows where the Alexa 488 dye was present on the gel, since the dye was used in a 10 fold (Lane 2 in Figure 8A) or 50 fold excess (Lane 3 in Figure 8A). However, no new bands were detected, indicating that the oligonucleotide did not undergo a click reaction. Therefore, we used different reaction conditions [100] with CuBr and TBTA to stabilize Cu (I) in Tris buffer (pH8) with

24mer EdU -1 (Lane 1 and 2 in Figure 8B) and 24mer EdU wAAF-1 (Lanes 3 and 4 in Figure 8B). Again, we could not detect any oligonucleotides coupled to the dye. A third set of reaction conditions used $CuSO_4$, TCEP, and TBTA in combination of a 4:3:1 mixture of H_2O: DMSO: BuOH [102]. Again none of new bands was detected under these conditions (Figure 8C). From these experiments, we concluded that the substrates with the alkyne connected directly to the 5' position of the base were not suitable for our purposes.

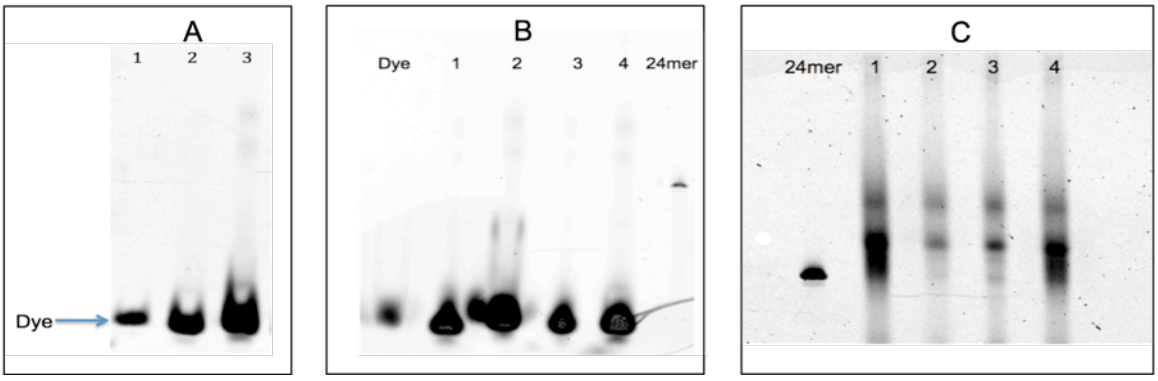


Figure 8A: Click reactions of EdU-containing oligonucleotides. The reaction condition 1 in method part was used for click reaction and 50fmol of each sample was loaded on lanes. (Lane 1: dye, Lane 2: 24mer EdU -1, Lane 3: 24mer EdU AAF -1 on a 20% denaturing gel)

Figure 8B: Click reactions of EdU-containing oligonucleotides. The reaction condition 2 in method part was used for Click reaction**.** (Lane 1: 50fmol and lane 2: 200fmol of 24mer EdU -1 and lane 3: 50fmol and 4: 200fmol of 24mer EdU AAF-1, A purchased 24mer oligonucleotide (from IDT) containing FAM (dye) was used as a marker. All samples were loaded on a 20% denaturing gel)

Figure 8C: Click reactions of EdU-containing oligonucleotides. The reaction condition 3 in method part was used for Click reaction and 50fmol of each sample was loaded on lanes. (Lane 1: 24mer with FAM, lane 1 and 4: without $CuSO_4$, lane 2 and 3: with $CuSO_4$, lane 1 and 2: 24mer EdU-1, reaction at 15°C and lane 3 and 4: 24mer EdU AAF-1, reaction at r.t., on a 20% Denaturing gel)

Preparation of Plasmids and NER Assay with Click Reaction Substrates with a Short Linker Alkyne

To introduce these oligonucleotides into plasmids, phosphorylated oligonucleotides with AAF or without AAF and EdU were annealed to ssDNA plasmid and primer extension was performed to generate the covalently closed circular (ccc) DNA as described in chapter 2 (Figure 9). After purification of the plasmid on a CsCl gradient, a pure cccDNA was obtained (Figure 9). The plasmid substrate was then used in a conventional *in vitro* NER assay using detection by radioactive labeling to determine if the alkyne interfered with the NER reaction. Three different substrates, 1, 2, and 3 in

Figure 10 were used to observe the influence of the presence of the EdU residue on the NER process. Identical incision patterns were observed for substrates containing only AAF and for substrate containing AAF and EdU (Figure 10, lanes 1 and 2), while no NER signal was observed for the plasmid containing only EdU (Figure 10, lane 3). Therefore, we conclude that an EdU residue does not interfere with the NER reaction.

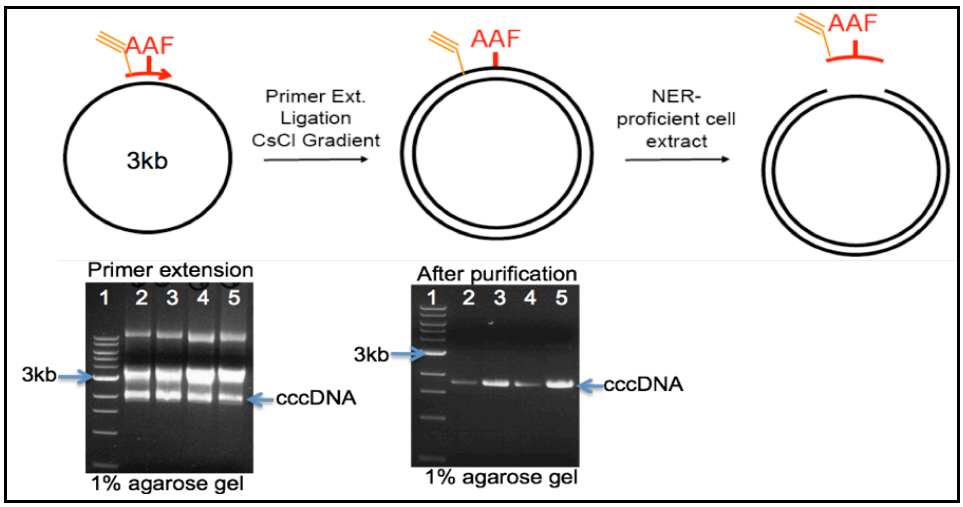


Figure 9: A scheme for primer extension. A plasmid for NER assay was prepared by annealing of a 5' phosphorylated oligonucleotide containing a site-specific lesion to a ssDNA. Primer extension was performed with T4 DNA polymerase and ligase and then CsCl gradient purification was achieved to have a pure plasmid. A purified plasmid containing a site-specific AAF was incubated with HeLa cell extract. (lane1: ladder, lane 2 and lane 3: 3µl and 5 µl of reaction solution containing 24mer EdU-1, lane 4 and 5: 3µl and 5 µl of reaction solution containing 24mer EdU wAAF-1 in primer extension gel) (lane1: ladder, lane 2 and 3: 70ng and 250ng of purified cccDNA containing 24mer EdU-1, lane 4 and 5: 70ng and 250ng of purified cccDNA containing 24mer EdU wAAF-1 in a gel after purification) Both gels were visualized with EtBr on 1% agarose gels.

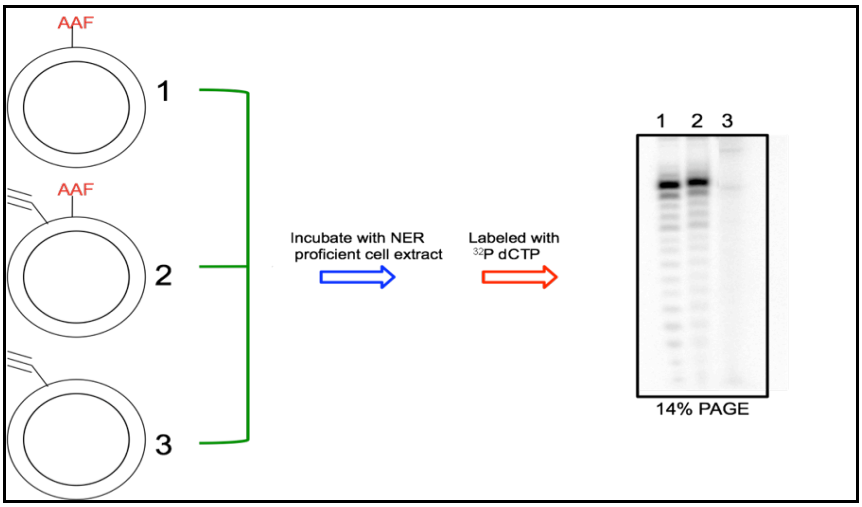


Figure 10: A NER assay with different substrates. A purified plasmid (50ng, 25fmol) was incubated with HeLa cell extract. The excised DNA fragments of 24-32 oligonucleotides were detected by annealing a complementary oligonucleotide containing a non-complementary 4G overhang and filling in with [α-^{32}P] dCTP. AAF: 2-acetyl amino fluorene. (1: a plasmid (6.3fmol) containing AAF, 2: a plasmid (6.3fmol) containing AAF and an alkyne, and 3: a plasmid (6.3fmol) containing an alkyne on a 14% denaturing sequencing gel.

Preparation of Oligonucleotides containing an Alkyne with a Longer Linker and Click Reaction conditions

A possible reason for the inability to label 5-ethynyl-2'-deoxyuridine with the fluorescent dye is that the short and rigid nature of the alkyne group and renders the click reaction inefficient [102]. Therefore, we decided to try the click reactions with the longer and more flexible alkyne in 5-octadiynyl-2'-deoxyuridine (OdU) [102, 108, 109] (Figure 11).

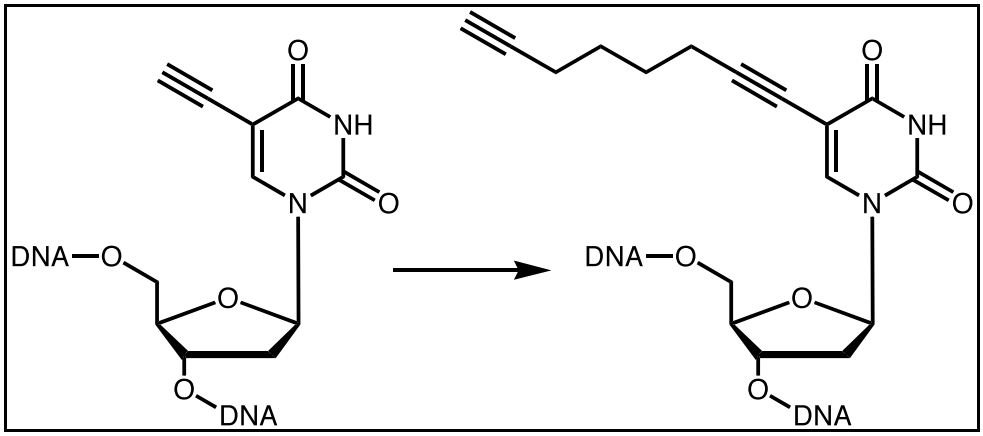


Figure 11: A change in substrates from 5-ethynyl-dU to 5-octadinyl-dU (OdU) for a click reaction.

Several oligonucleotides containing 5-octadiynyl-2'-deoxyuridine (Figure 11 and Table 2) were synthesized and purified by reverse-phase HPLC (See supplementary data). After first HPLC purification of 24mer OdU -1and 24mer OdU -3, these two oligonucleotides were tested in click reactions. As shown below, 24mer OdU -1 has only one OdU, but the 24mer OdU -3 has two OdU residues (Table 2). These two substrates were reacted with an azide-alexa 488 dye in the presence of CuBr, sulfonated bathophenanthroline ligand in Tris buffer (pH8) under inert atmosphere [107]. For a negative control, CuBr was omitted from the reactions. Reactions were stopped by adding EDTA and reactions were analyzed by 20% denaturing PAGE. The formation of two new bands was observed and their formation was dependent on the presence of CuBr indicative of a successful click reaction (Figure 12A, Lanes 1 and 4). The oligonucleotide with two OdU residues displayed a slower mobility after the click reaction (Figure 12A, compare lanes 1 and 3) consistent with the presence of one versus two alkyne groups.

Name	Sequence
24mer OdU -1	5'-TAG CTA **U**TA CCG GCG CCA CAT GTC-3'
24mer OdU -2	5'-**U**AG CTA TTA CCG GCG CCA CAT GTC-3'
24mer OdU -3	5'-**U**AG CTA TTA CCG GCG CCA CA**U** GTC-3'
24mer OdU -4	5'-**U**AG C**U**A TTA CCG GCG CCA CAT GTC-3'
24mer- OdU AAF -1	5'-TAG CTA **U**TA CCG GC**X** CCA CAT GTC-3'
24mer- OdU AAF -2	5'-**U**AG CTA TTA CCG GC**X** CCA CAT GTC-3'
24mer- OdU AAF -3	5'-**U**AG CTA TTA CCG GC**X** CCA CA**U** GTC-3'
24mer- OdU AAF -4	5'-**U**AG C**U**A TTA CCG GC**X** CCA CAT GTC-3'

Table 2: Sequences of the Oligonucleotides used in this click study. These oligonucleoties were synthesized from expedite DNA synthesizer. (X= AAF, U= Octadiynyl-dU).

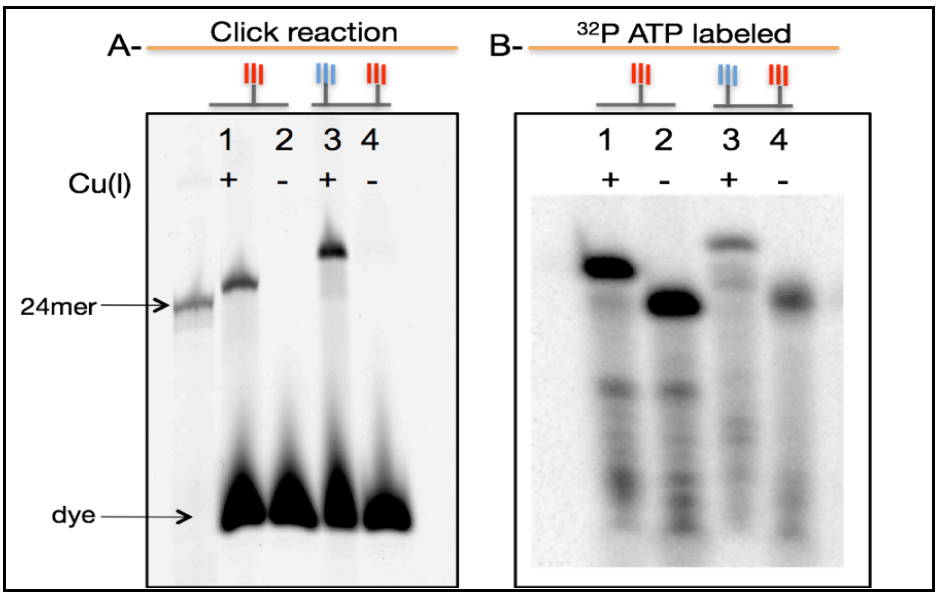


Figure 12: A click reaction with OdU-containing oligonucleotides. The reaction condition 1 in method part (CuBr/Sulfonated bathophenanthroline) was used. 24mer OdU-1 was used for a click reaction with CuBr (lane 1, 200fmol) /without CuBr (lane 2, 200fmol) and 24mer OdU-3 was used for a click reaction with CuBr (lane 3, 200fmol) /without CuBr (lane 4, 200fmol) in A. 24mer OdU-1 was used for a click reaction with CuBr (lane 1, 30fmol) /without CuBr (lane 2, 30fmol) and labeled with ^{32}P-ATP. 24mer OdU-3 was used for a click reaction with CuBr (lane 3, 30fmol) /without CuBr (lane 4, 30fmol) and labeled with ^{32}P-ATP in B.

To make sure that the new bands were the result of a click reaction with the oligonucletide, the same samples were 5'-labeled with γ-[^{32}P]-ATP and T4 polynucleotide kinase after the click reaction. Figure 12B shows faster migrating bands from the reactions in absence of CuBr (Figure 12B, Lanes 2 and 4) and slower

migrating bands from the reactions with CuBr (Figure 12B, Lanes 1 and 3), consistent with the addition of the fluorophore. Furthermore, no unreacted oligonucleotides were observed in lanes 1 and 3 (Figure 12B), indicating that the click reaction had gone to completion. Since the click reaction was successful with OdU-modified substrates, these were used for further studies.

Then, we tested the use of different ligands and reaction conditions to avoid the use of an inert-atmosphere glove box. 24mer- OdU AAF -2 was treated with pre-mixed CuBr and TBTA (1:2) solution [100] (Figure 13) and the formation of products was monitored in time course experiment and products analyzed by 20% denaturing PAGE. Under these conditions, long reaction times of about 10 hrs (Lane 4 in Figure 13) were needed for the reaction to go to completion. Additional reaction conditions were examined for more efficient labeling. Sodium ascorbate was used as a reducing agent for Cu (II) and

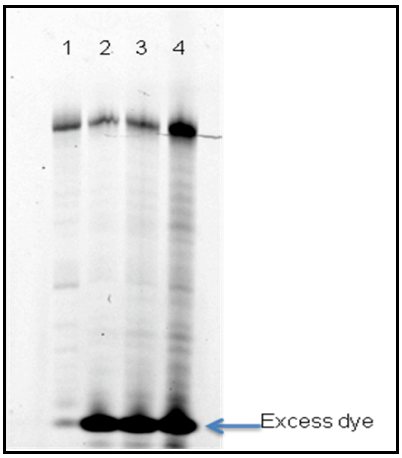


Figure 13: A click reaction with CuBr and TBTA. The reaction condition 2 in method part was used. 1: 100fmol of a clicked 24mer- OdU AAF -1, 2: 100fmol of a 24mer- OdU AAF -2 for 3hr click reaction, 3: 100fmol of a 24mer- OdU AAF -2 for 6hr click reaction, 4: 100fmol of a 24mer- OdU AAF -2 for 10hr click reaction on 20% denaturing PAGE.

(Tris (3-hydroxypropyl triazolylmethyl)amine THPTA) was used to accelerate the decomposition of peroxide produced by copper [104]. In addition, aminoguanidine was added to prevent some side reactions from dehydrated Na ascorbate. In a time course experiment we found no difference in the fluorescent intensities from 1 hr reaction and 3 hr reactions (Figure 14). To further characterize the reaction products, we purified the products by spin column or dialysis to remove excess dye and copper. Buffer exchanged by centricon (centrifugal column, molecular weight cut off (MWCO)=3,000) resulted in the efficient removal of excess dye (Figure 15, compare lanes 1 and 2), but

this method yielded only a low sample recovery (~50%) and required long centrifugation times (~2days). Therefore, we needed to find an alternative way to purify the reaction

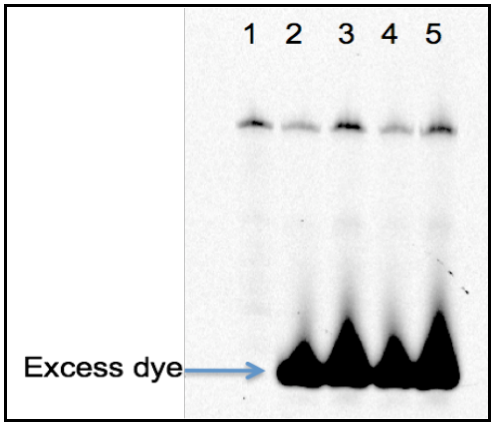


Figure 14: A click reaction with $CuSO_4$ and TBTA. 1: 50fmol of a clicked 24mer- OdU AAF -2, 2: 50fmol of a clicked 24mer- OdU AAF -2 for 1hr reaction, 3: 100fmol of a clicked 24mer- OdU AAF -2 for 1hr reaction, 4: 50fmol of a clicked 24mer- OdU AAF -2 for 3hr reaction, 5: 100fmol of a clicked 24mer- OdU AAF -7 for 3hr reaction on 20% denaturing PAGE.

products. Purification of the product over a spin column (molecular weight cut off (MWCO) = 10bp) (Figure 15, lanes 3 and 4) helped to remove the excess of dye. Purification by dialysis (MWCO = 3500) (Figure 15, lanes 5 and 6) resulted only in partial removal of the dye. To determine the purification efficiency of spin column, the crude sample after a click reaction was fractionated by HPLC (in Figure 16A) and the peaks 1-3 in 16A were collected and identified with MALDI-MS in the crude mixture. Peak 1 in Figure 16A could not be identified by MS, but peak 2 corresponded to a desired product (MW: 8280±4 corresponding to 24mer-Alexa 488 (8284) + Na^+(23), see Supplementary Data). Analysis of peak 3 revealed that it corresponded to the Alexa 488 dye. HPLC analysis of the sample after purification by spin column showed two peaks (Figure 16B), with Peak 2 corresponding to the product by MS analysis (See in Supplementary Data), while no Alexa dye was found, confirming that the use of spin column was suitable for purification of the product. The use of the spin column in this experiment provided a higher sample recovery (~ 70%) and shorter time (less than 30 min.) than the buffer exchanged method. Based on these findings, we decided to use purification over a spin column to purify the products of the click reaction in our experiments.

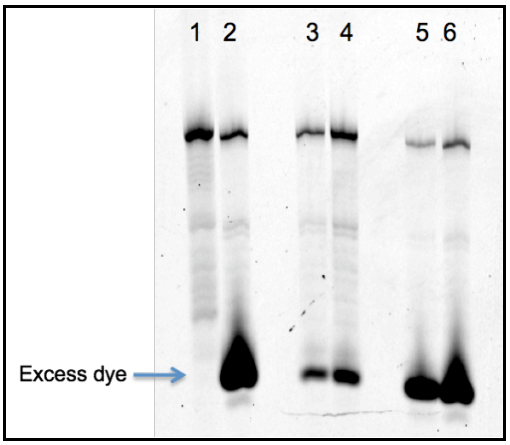


Figure 15: Different purification conditions after a click reaction with $CuSO_4$ and TBTA. 1: 100fmol of a clicked 24mer- OdU AAF -2 (buffer exchanged), 2: 100fmol of a clicked 24mer- OdU AAF -2 without any purification, 3: 100fmol of a clicked 24mer- OdU AAF -2 after applying spin column, 4: 200fmol of a clicked 24mer- OdU AAF -2 after applying spin column, 5: 100fmol of a clicked 24mer- OdU AAF -2 after O/N dialysis, 6: 200fmol of a clicked 24mer- OdU AAF -2 after O/N dialysis on 20% denaturing PAGE.

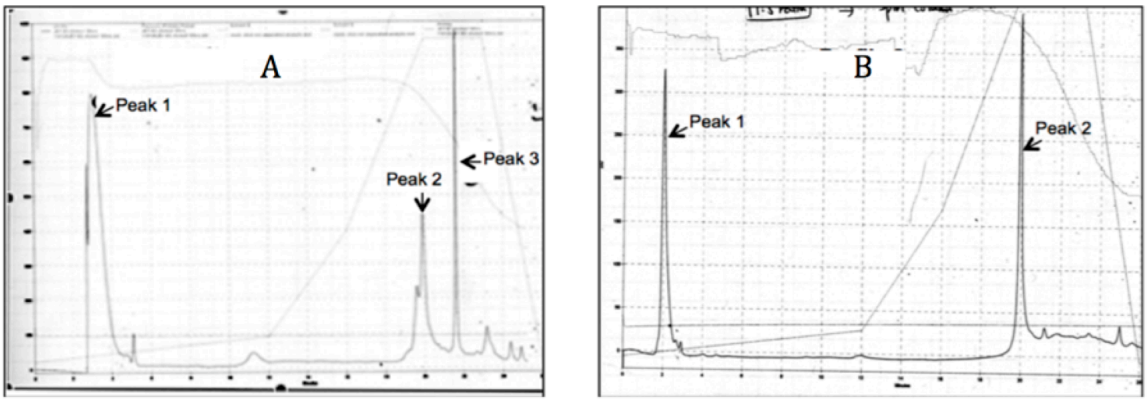


Figure 16: Reverse-phase HPLC Purification after a click reaction with $CuSO_4$ and TBTA. A: HPLC purification after a click reaction and B: HPLC purification after applying spin column of a click reaction - 500pmol of a sample (a clicked 24mer- OdU AAF -2) was injected in to HPLC.

NER Studies with OdU-labeled substrates

We then used OdU-containing dG-AAF substrates in the NER assay. Plasmids containing a dG-AAF adduct and either one or two OdU residues were prepared by primer extension and CsCl gradient purification, yielding cccDNA NER substrates containing one or two alkynes (Figure 17). Once the plasmids were prepared, a digestion with restriction enzymes was performed to verify that the fluorescently labeled fragments could be detected. These fragments would also represent unreacted plasmid after the NER reaction, allowing for the visualization of unreacted substrates of the

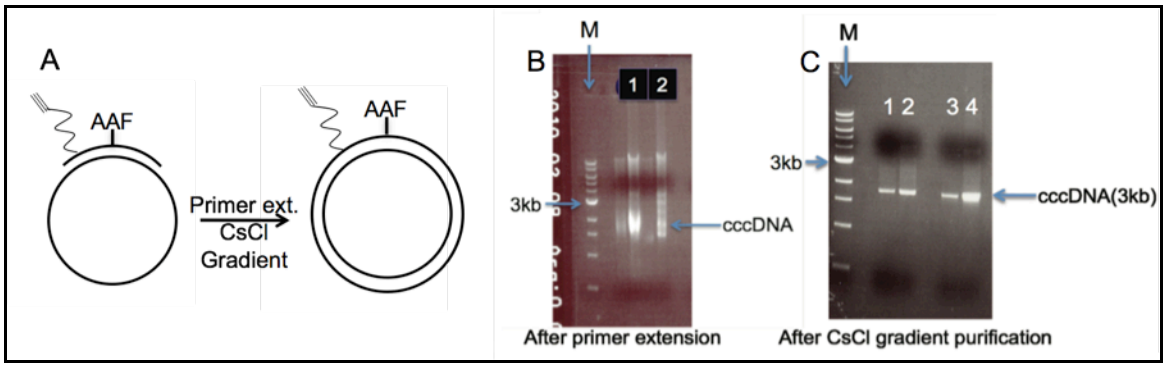


Figure 17: Primer extension and CsCl gradient purification. In gel B, lane 1: 150fmol of primer extension reaction for 24 mer OdU -1, lane 2: 150fmol of primer extension reaction for 24mer OdU AAF-1, In gel C, lanes 1 (100ng) and 2 (200ng): after CsCl purification of primer extension reaction with 24mer OdU-1, lanes 3 (100ng) and 4 (200ng): after CsCl purification of primer extension reaction with 24 mer OdU-1. The gels were stained with EtBr on 1% agarose gels. M: 1Kb ladder from NEB. The position of a 3kb is indicated.

NER reaction. The OdU-containing plasmid was digested with EcoRI and BamHI to yield a 67bp fragment, which was subjected to a click reaction with Alexa 488 in presence of CuBr and bathophenanthroline under rigorous exclusion of oxygen. Products were analyzed by a denaturing 15% PAGE gel. In this reaction, a band just above 60bp was present (Band 1 in Figure 18), at the proper position for the restriction fragment. However, bands for two bigger fragments (Bands 2 and 3 in Figure 18) were also observed and the origin of these bands was unclear.

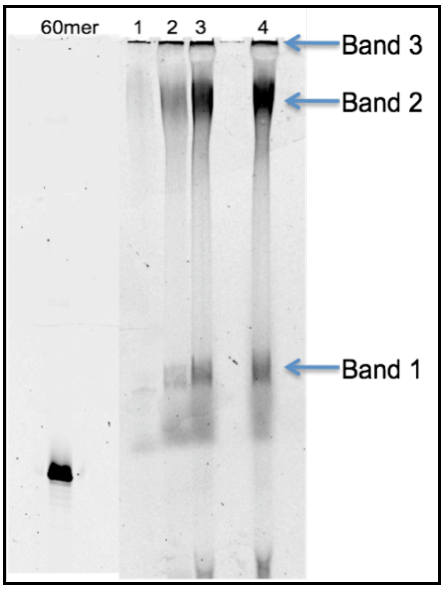


Figure 18: A click reaction after double digestions. Plasmids were incubated with BamHI and EcoRI. A digested pladmid containing a 24mer OdU AAF -1 was loaded on 15% PAGE. (Lane 1: 15fmol, Lane 2: 50fmol, Lane 3: 100fmol, and Lane 4: 200fmol) A fluorescently labeled 60mer with Cy5 (purchased from IDT company) was used as a marker.

In vitro NER Assay with Fluorescently labeled Substrates containing an Alkyne with a Longer Linker

Another way to monitor the NER reaction is to labele substrates by the click reaction prior to incorporation into the plasmids. To produce fluorescently labeled substrates, 24mer OdU -1 and 24mer OdU AAF -1 were used for a click reaction and the product purified by spin column and HPLC. The purified oligonucleotides were used to generate plasmids by primer extension and CsCl gradient and sucrose gradient purification (Figure 19). After purification of plasmids, a simple double digestion with the two restriction enzymes BamHI and EcoRI was performed to ensure that the substrates could be detected by fluorescence (Figure 20). To compare the intensities of the signal

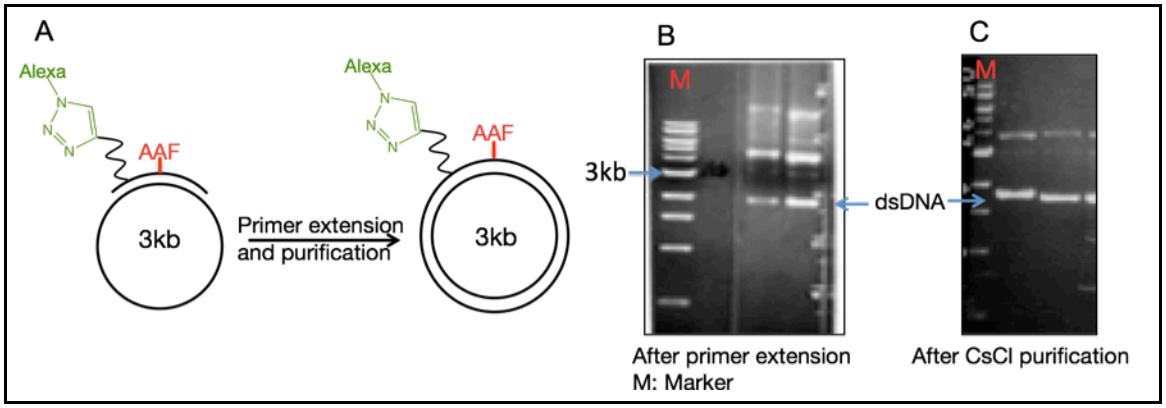


Figure 19: Primer extension and purification with pre-labeled 24mer OdU AAF -1. 100ng of samples was loaded after primer extension (B) and purification (C) on 1% agarose gel and stained with EtBr. M: 1Kb ladder from NEB. The position of a 3kb is indicated. The position of dsDNA (3Kb) was indicated on gels.

after the digestion, 10 fmol of a purified, fluorescently labeled 24mer OdU AAF -1 was loaded on the gel. The digested fragments could readily be detected by fluorescence and already the lowest amount (12.5 fmol) loaded on the gel, yielded a robust signal (Figure 20). These finding suggested to suitability of these substrates for NER assays. We, then, examined the effect of the Alexa488-OdU conjugated substrate on the NER reaction. We first aimed to detect the reaction products using radioactive detection to compare the signals observed for dG-AAF, dG-AAF/Alexa488-OdU and Alexa488-OdU containing plasmids following incubation with a HeLa whole cell extract. The plasmid containing only the dG-AAF lesion was efficiently repaired by NER displaying the typical NER signals after carrying out the fill-in reaction to label the product (Figure 21, lane 1).

By contrast no signal was observed for the reaction with the plasmids that contained the Alexa488-OdU modification only (Figure 21, lane 3). Incubation of the dG-AAF/Alexa488-OdU substrate also led to the formation of NER-specific bands, but interestingly, their position was shifted to higher molecular weight bands compared to the dG-AAF substrate (Figure 21, Lane 2). We believe that the presence of OdU-alexa 488 residue on a plasmid could change the way the NER pre-incision complex is assembled, since the XPD helicase, which scans the DNA to verify the presence of the lesion, is likely to stall the OdU-Alexa 488 residue before reaching the dG-AAF reisude (Figure 22). This leads to the positioning of the ERCC1-XPF nuclease that makes the incision 5' to the lesion more 5' to the lesion and the formation of a larger product under our assay conditions.

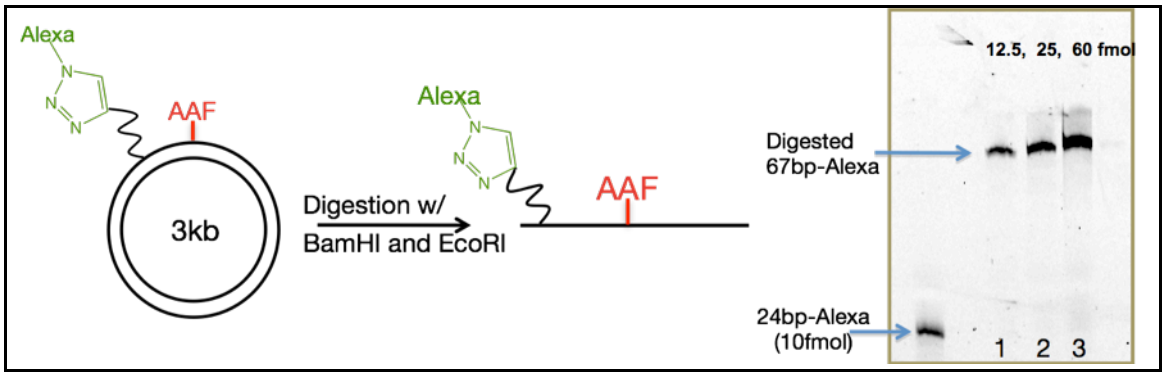


Figure 20: Double digestion of a plasmid containing a pre-labeled 24mer OdU AAF -1. For double digestion, BamHI and EcoRI were used to give a 67bp fragment. Lane 1: 12.5 fmol, Lane 2: 25 fmol, Lane 3: 60 fmol of reaction samples on 15% denaturing PAGE.

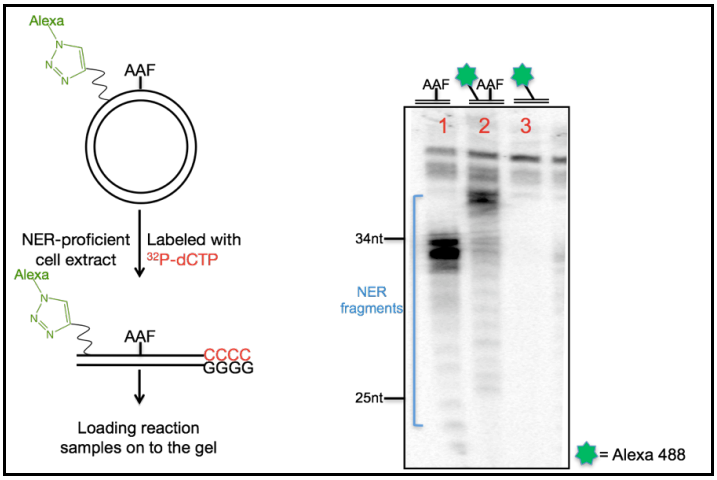


Figure 21: The conventional NER assay with clicked substrates (in this reaction, a prelabeled 24mer OdU AAF-1 containing plasmid was used). Lane 1: a plasmid with AAF (a positive control), lane 2: a clicked plasmid with AAF, lane 3: a clicked plasmid without AAF on a 14% sequencing gel. A labeled low molecular weight DNA ladder (NEB) was used as a marker. The positions of 25nt and 34nt are indicated.

Since the repair of dG-AAF/OdU-Alexa488 substrate was confirmed using the conventional NER assay, we set out to monitor NER using direct fluorescent detection. This plasmid was incubated with HeLa cell extract and the reaction products analyzed on a 12% sequencing gel and any products visualized on a Typhoon imaging system (Figure 23). A signal at the expected size was indeed observed (Figure 23, lanes 2-4), although the intensities were rather weak. The NER reaction is expected to proceed with a yield of around 10%, thus the fluorescent signals on the gel are expected to

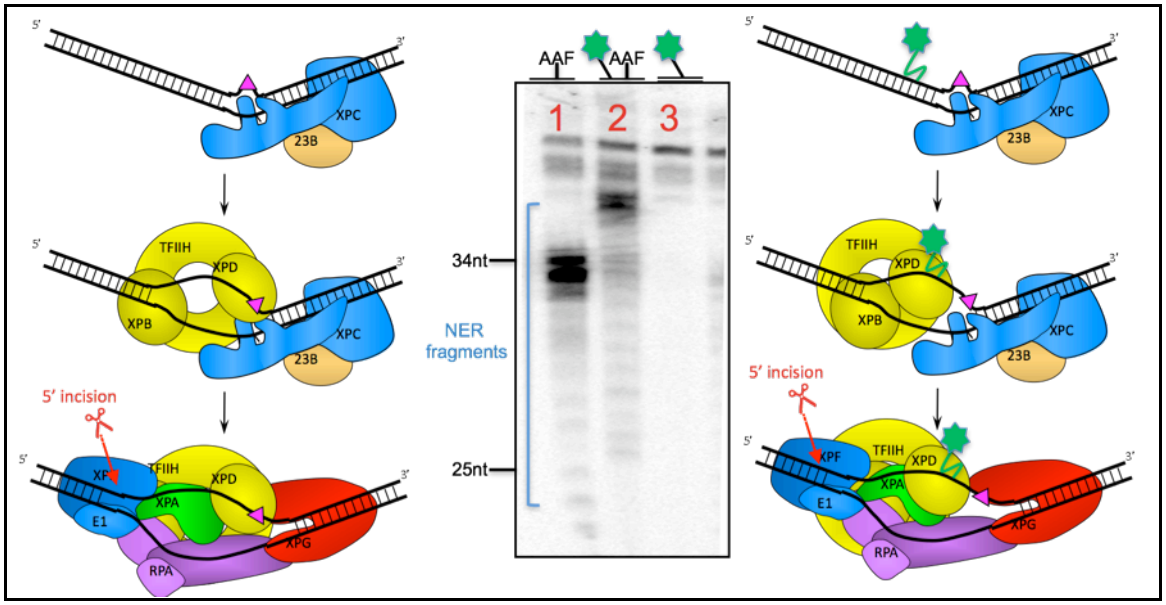


Figure 22: Localization of pre-incision complexes on a damaged DNA without/with Alexa 488. Pinkish triangle: a damage, Blue+orange: XPC-RAD23B, Yellow: TFIIH, Green: XPA, Purple: RPA, Green star: Alexa 488.

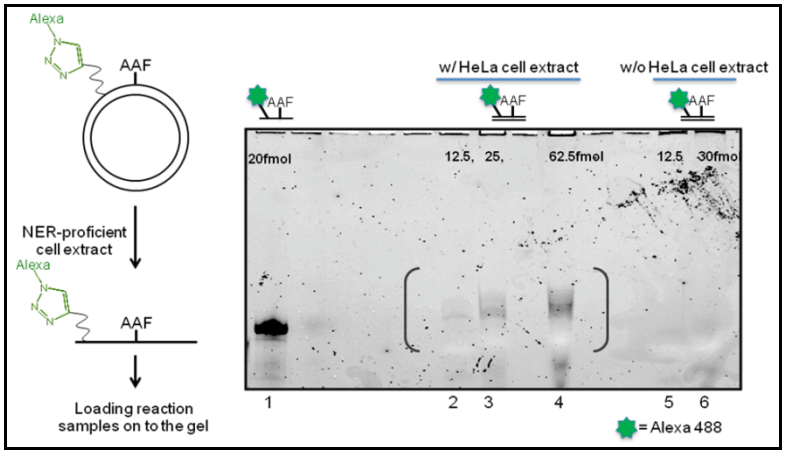


Figure 23: NER assay with clicked substrates (in this reaction, a prelabeled 24mer OdU AAF-1 containing plasmid was used). Lane 1: 20fmol of a clicked 24mer with Alexa 488, lane 2: 12.5fmol of a clicked plasmid with AAF with HeLa cell incubation, lane 3: 25fmol of a clicked plasmid with AAF with HeLa cell incubation, lane 4: 62.5fmol of a clicked plasmid with HeLa cell incubation, lane 5: 12.5fmol of a clicked plasmid without HeLa cell extract, lane 6: 30fmol of a clicked plasmid without HeLa cell extract on a 20% denaturing PAGE.

be about 1.2 fmol (lane 2), 2.5 fmol (lane 3) and 6.5 fmol (lane 4). However, the observed signal was specific, as no signal was observed in the absence of HeLa cell extract (Figures 23, lanes 5 and 6). These studies show that NER can be monitored by direct detection by fluorescent, although the intensity of the signals still needs to be improved.

DISCUSSION

Conventional ways (Figure 1) of monitoring NER require the inconvenience of using radioactive materials and is limited to the detection of only products, but not substrates or reaction intermediates. To overcome these disadvantages, we developed an approach that uses click chemistry to introduce fluorophores on NER substrates that allow the monitoring of NER reactions by fluorescence detection. We incorporated nucleotide residues containing alkyne groups that could be fluorescently labeled by reaction with an azide-conjugated fluorescent dye. The oligonucleotides were incorporated into plasmids and used in NER assays. Substrates containing the NER lesion dG-AAF and alkynes were repaired by NER, while Substrates containing only an alkyne group or an alkyne linked to alexa dye were not repaired by GG-NER. In addition, our experiments (Figure 16) show that the NER reaction can be monitored by fluorescence in substrates containing fluorescent labels, albeit only with low sensitivity. Present work is aimed at improving the reaction parameters to enhance the sensitivity of detection and make the approach practical for routine use in monitoring NER reaction.

MATERIALS AND METHODS

Reagents and equipments

Chemicals and solvents were purchased from Fluka-Sigma-Aldrich. Reagents for DNA synthesis were purchased from Glen Research. The “ultra-mild” phosphoramidites were available from Glen Research and the 1000 Å “Q-columns” from Biosearch

Technologies. 5-ethynyl-2'-deoxyuridine 2-cyanoethyl-*N*, *N*-diisopropylphosphoramidite (CEP) and 5-octadiynyl-2'-deoxyuridine CEP were purchased from Barry & Associates company. A 24mer oligonucleotide labeled with FAM (Fluorescein amidite) and a 60mer oligonucleotide labeled with Cy5 (Cyanine) were purchased from Integrated DNA Technologies (IDT). DNA syntheses were performed on a PerSeptive Biosystems Expedite 8909 DNA synthesizer. HPLC purifications were performed on a JASCO system equipped with a Phenomenex Clarity Oligo-RP Semi-prep column: C18, 5µm, 50×10.00 mm. The C18-SepPak cartridges were from Waters. Alexa fluor 488 azide was purchased from invitrogen. 1Kb ladder, T4 PNK, T4 polymerase, and T4 DNA ligase were from New England Biolabs (NEB). The sequenase enzyme (T7 DNA polymerase) was purchased from USB and [α-^{32}P] dCTP from PerkinElmer.

The 24-mer oligonucleotides were analyzed by Waters micromass platform LCZ or by Voyager-DE STR (MALDI-TOF) (Applied Biosystems).

Preparation of Oligonucleotides containing AAF and alkynes

DNA synthesizer: 24-mers with alkynes were synthesized based on modified 'ultra-mild' DNA synthesis according to a published procedure [60].

Reverse-phase HPLC purification: All HPLC elutions were performed at a flow rate of 4ml /min, with following gradient: linear 5-10% B over 10min, linear 10-25% B until 15min, linear 25-95% B until 19min. isocratic 95% until 22min, linear 95-5% B until 25 min, isocratic 5% until 27min; buffer A: 0.1M TEAA (pH7); buffer B: CH_3CN. The peak of the DMT on oligonucleotide-eluting between 17-19 min-was collected (See Supplementary data), concentrated and treated with a 80% acetic acid for 40 min at room temperature to remove the 5'-DMTr group. After concentration, the oligonucleotide was redissolved in 1ml of 1M TEAA (pH 7) and repurified on HPLC. The major peak-eluting between 12 and 13min- was collected, concentrated redissolved in 0.1M TEAA (pH 7), desalted on a C18-SepPak catridge and concentrated. The desalted samples were redissolved in 200-300 µl of milli-Q water and the concentration was measured on

a NanoDrop 1000 (Thermo Scientific). The characterization of samples was performed by ESI or MALDI mass spectrometry (Supplementary data).

The conditions for click reactions

1) CuBr and Bathophenanthroline condition: an oligonucleotide containing an alkyne (40nM) was incubated with Alexa 488-azide (400nM) in Tris buffer (100mM, pH 8) involving bathophenanthroline (3mM) and CuBr (1mM) in 25µl for 12 h at room temperature with rigorous exclusion of dioxygen. After incubation, 10mM EDTA was added to quench the reaction. This click reaction was followed by a published procedure [107]. Reactions were resolved by denaturing polyacrylamide gel electrophoresis and gels were visualized using a Typhoon 9400 imager.

2) CuBr and TBTA (Tris-(benzyltriazolylmethyl)amine) condition: an oligonucleotide containing an alkyne (50µM) was incubated with Alexa 488-azide (250µM) in DMSO : t-BuOH (3:1) involving 500µM of CuBr and 1mM of TBTA in 20µl for up to 10h at room temperature with vigorous shaking. After incubation, 10mM EDTA was added to quench the reaction. This click reaction was followed by a published procedure [100]. Reactions were resolved by denaturing polyacrylamide gel electrophoresis and gels were visualized using a Typhoon 9400 imager.

3) $CuSO_4$ and ligands condition: an oligonucleotide containing an alkyne (6µM) was dissolved with 124.1 µl of 100mM Hepes (pH 7.5). Alexa 488 dye (30µM) was added and then premixed CuSO4 (120µM) and THPTA (600µM) was added to the reaction. Aminoguanidine hydrochloride (6µM) and then freshly prepared Na ascorbate (6µM) were added to the reaction to give 167µl final volume. This reaction was incubated at room temperature with vigorous shaking for 1hr. After incubation, 10mM EDTA was added to quench the reaction. This click reaction was followed by a published procedure [104]. Reactions were resolved by denaturing polyacrylamide gel electrophoresis and gels were directly visualized using a Typhoon 9400 imager.

Preparation of single strand DNA for plasmid preparation (NarI sequence, p98 plasmid)

2ml inoculation was performed with single colony of p98 (*NarI* MOCK in pBlkp II SK+) in presence of appropriate antibiotics for 6h. After inoculation, 0.5ml R408 helper phage was added to 2ml inoculation and incubated for 10min at room temperature. After incubation, this 2ml culture was transferred to 600ml LB with an appropriate antibiotic for over night. Harvested 600ml culture was centrifuge at 4000 rpm for 40min and the supernatant was transferred to the other centrifuge tube and this step was repeated. Clarified supernatant was treated by adding 200µl of 5mg/ml RNase A and 50µl DNase I (10U/µl) and incubated at 37°C for 1h 30min. After incubation, the supernatant was treated 20% PEG (8000); 2.5M NH_4OAc and incubated at 4°C for over night. After overnight incubation, the incubated inoculation was centrifuged at 4000rpm for 40min at 4°C and then the pellet was dissolved in 10mL of 10mM tris-HCl (pH8). The dissolved pellet was extracted with Phenol/chloroform and centrifuged at 4000xg for 10min and then repeat until no more white precipitant is visible at interface. Additional extraction was performed with 10ml chloroform and centrifuge at 4000xg for10min. The water layer was added 3M NaOAc (pH4.8) (0.1 volume of sample) and 100% cold EtOH (2.2 or 2.5 eq. volume of sample) and mix well. The mixed sample was incubated at -80°C for overnight. The overnight incubation was centrifuged at 4000 rpm for 40min. at 4°C. The precipitated pellet was washed with 70% EtOH (0.5 volume of sample) at 4000rpm, for 10min at 4°C. The EtOH was decanted and the pellet was dried in the air. The dried pellet was resuspended with 10mM tris-HCl (pH8) and then concentration was check and purity was checke by an 1% agarose gel.

Preparation of plasmids (primer extension) and Conventional In vitro NER assay

100 pmol of 24mer oligonucleotides for a click reaction were 5'-phosphorylated by incubation with 20 units of T4 PNK enzyme and 2mM of ATP at 37 °C for 2h. After annealing with 30pmol of single-stranded p98 (NarI), further incubation with dNTPs, T4 DNA polymerase and T4 DNA ligase yielded to covalently closed circular DNA containing a single AAF (or AF) adduct. The closed circular DNA was purified by cesium

chloride/ethidium bromide density gradient centrifugation, by consecutive butanol extractions to remove the ethidium bromide and finally concentrated on a Centricon YM-30 (Millipore). In addition to further purification, the collected covalently closed circular DNA was purified by sucrose gradient centrifugation to remove the rest of ethidium brimide. The plasmids were aliquoted and stored at -80 °C. The *in vitro* NER assay was performed as described [62]. HeLa (2 µl) cell extract, 2 µl of 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) (final NaCl concentration was 70 mM) in a total volume of 10 µl were pre-incubated at 30°C for 10 min. One µL of a covalently-closed circular DNA plasmid (50 ng) was added before incubating the mixture at 30 °C for 45 min. After placing the samples on ice, 0.5 µl of 1 µM of a 3'-phophorylated oligonucleotide: d(GGGGCATGTGGCGCCGGTAATAGCTACGTAGCTC)was added and the mixtures heated at 95 °C for 5 min. The samples were allowed to cool down at room temperature for 15 min to allow the DNA to anneal. One µL of a Sequenase/[α-^{32}P]-dCTP mix (0.25 units of Sequenase and 2.5 µCi of [α-^{32}P]-dCTP per reaction) was added before incubating at 37 °C for 3 min, 1.2 µl of dNTP mix (100 µM of each dATP, dTTP, dGTP; 50 µM dCTP) was added and the mixture incubated for another 12 min. The reactions were stopped by adding 12 µl of loading dye (80% formamide/10 mM EDTA) and heating at 95 °C for 5 min. The samples were run on a 14% sequencing gel (0.5x TBE) at 45 W for 2.5 hrs. The reactions products were visualized using a PhosphorImager (Typhoon 9400, Amersham Biosciences).

In vitro NER assay with clicked substrates

The *in vitro* NER assay was performed as described [62]. HeLa (2 µl) cell extract, 2 µl of 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) (final NaCl concentration was 70 mM) in a total volume of 10 µl were pre-incubated at 30°C for 10 min. One µL of a covalently-closed circular DNA plasmid (50 ng) was added before

incubating the mixture at 30 °C for 45 min. After incubation, the reaction was stopped by adding 12 µl of 1X loading buffer containing 80% formamide/10mM EDTA and heating at 95 °C for 5 min. The samples were run on a 14% sequencing gel (0.5x TBE) at 45 W for 2.5 hrs. The reactions products were visualized using fluorescent scan (Typhoon 9400, Amersham Biosciences).

The double digestion condition with 24mer oligonucleotides with an alkyne and a click reaction

1µg of a plasmid (3kb) was incubated with 20unit of EcoRI and 20unit of BamHI to give 67bp in presence of 1×Bovine serum albumin (BSA) and 1×New England Buffer (NEB) in a 10µl reaction at 37°C for 2hr. After double digestion, the click reaction was carried out (described as above). After incubation, the reaction was stopped by 1X loading buffer containing 80% formamide/10mM EDTA and heating at 95 °C for 5 min. The samples were run on a 15% denaturing PAGE (0.5x TBE). The reactions products were visualized using fluorescent scan (Typhoon 9400, Amersham Biosciences).

SUPPLEMENTARY DATA

General HPLC profiles

During the oligonucleotide for a click reaction purification process.

1) HPLC profile, purification of the 'DMTr-ON' oligonucleotides; MeOH/iPr_2NH deprotection

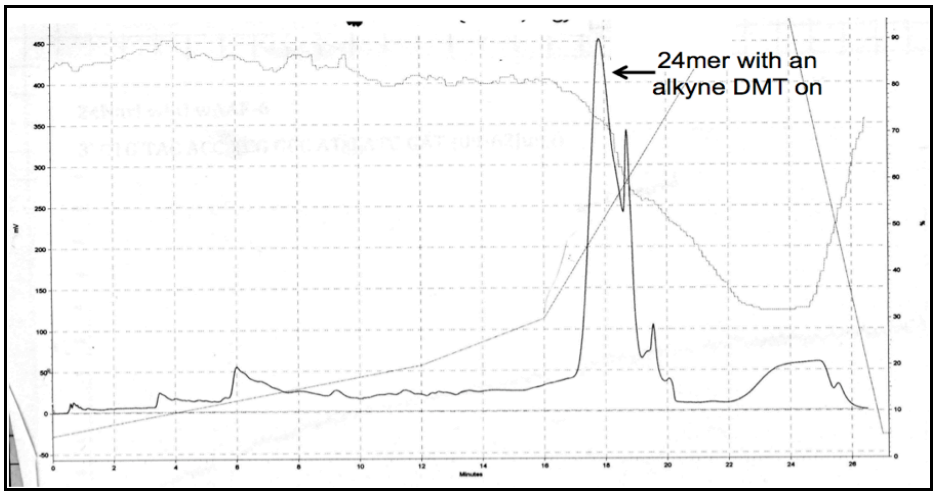


Figure 1: HPLC profile for oligonucleotide containing an alikyne and DMT. Gradient: linear 5-10% B over 10min, linear 10-25% B until 15min, linear 25-95% B until 19min. isocratic 95% until 22min, linear 95-5% B until 25 min, isocratic 5% until 27min; buffer A: 0.1M TEAA (pH=7); buffer B: CH_3CN.

2) Acetic acid cleavage; purification of the 'DMTr-OFF' oligonucleotides

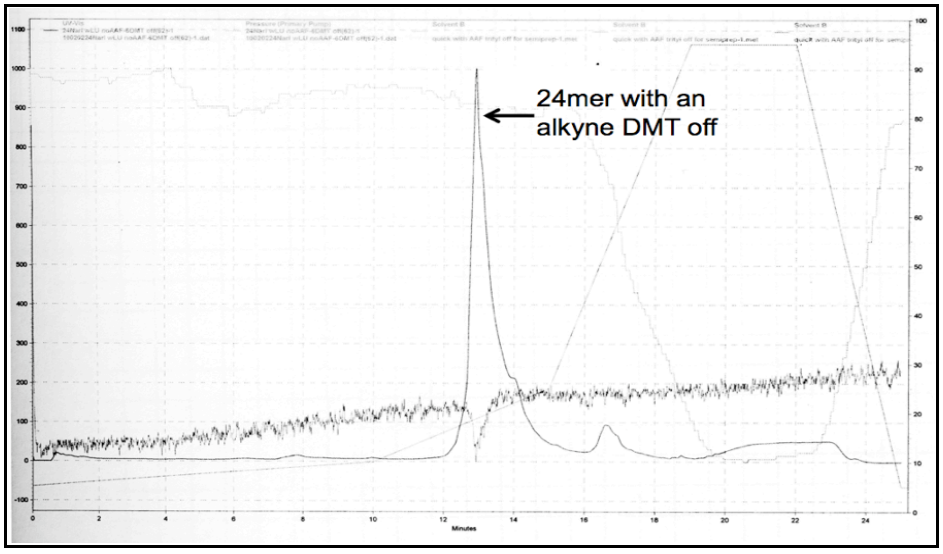


Figure 2: HPLC profile for oligonucleotide containing an alikyne (DMT off). Gradient: linear 5-10% B over 10min, linear 10-25% B until 15min, linear 25-95% B until 19min. isocratic 95% until 22min, linear 95-5% B until 25 min, isocratic 5% until 27min; buffer A: 0.1M TEAA (pH=7); buffer B: CH_3CN.

Data from Mass spectrometry

1) 24mer-AAF OdU-6,
   Expected molecular weight: 7601.96, Obtained molecular weight: 7604.06

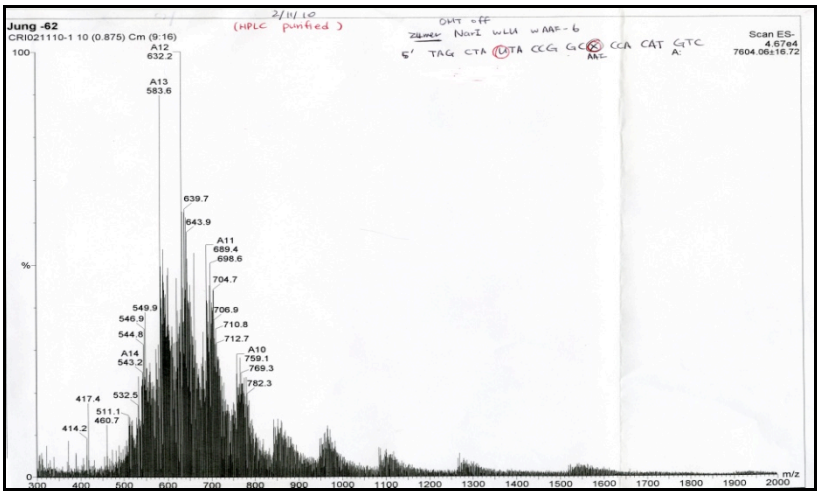


Figure 3: Mass spectrometry (MS) result of 24mer-AAF OdU-6 from ESI, negative mode.

2) 24mer OdU No AAF-1,
Expected molecular weight: 7383.96, Obtained molecular weight: 7382.02

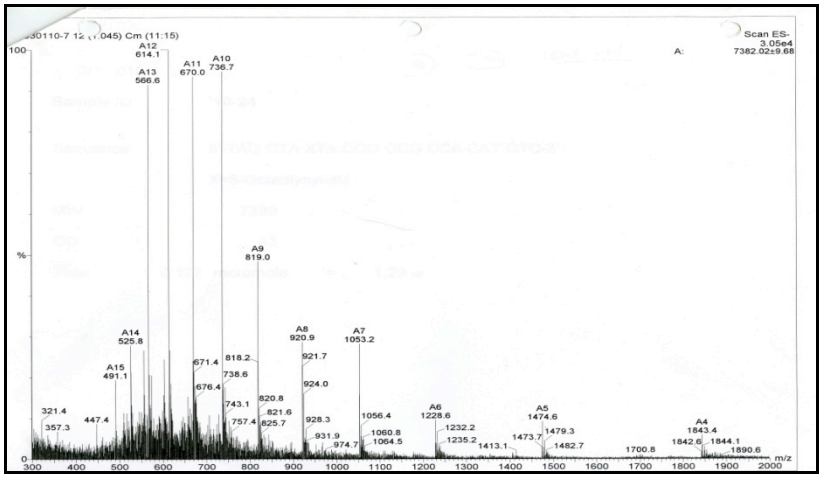


Figure 4: MS result of 24mer OdU No AAF-1 from ESI, negative mode.

3) Alexa 488-peak 3 in figure 12A,
Expected molecular weight: 657.04, Obtained molecular weight: 657.1

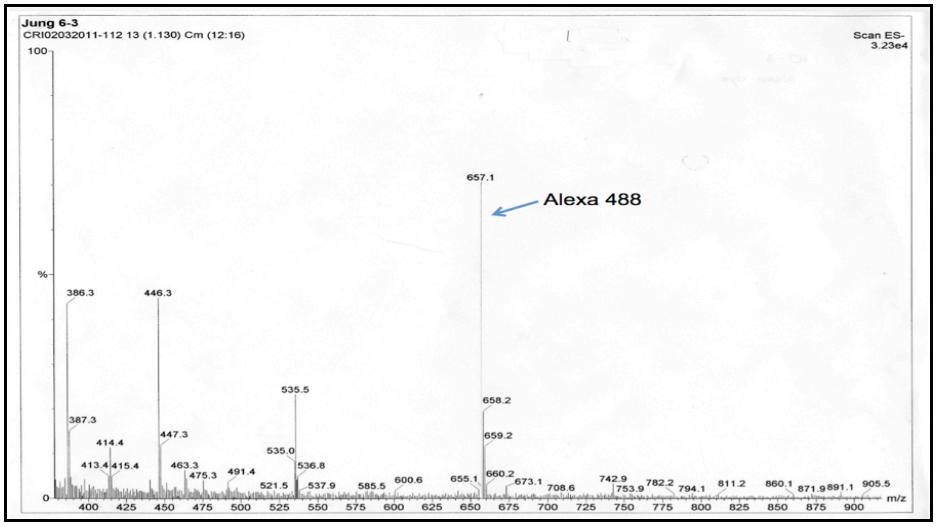


Figure 5: MS result of Alexa 488 from HPLC purification, negative mode.

4) 24mer-AAF OdU-7+Alexa 488 clicked-peak 2 in figure 12A and 12B,
Expected molecular weight: 8261, Obtained molecular weight: 8280.02(+Na)

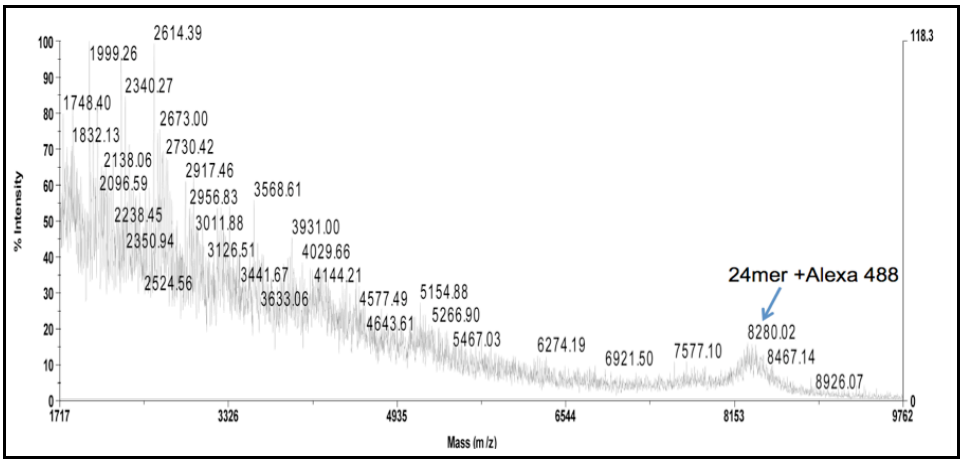


Figure 6: Mass spectrum (MALDI) of a Clicked 24 with Alexa 488 from HPLC purification.

REFERENCES

[1] E. C. Friedberg, G. C. Walker, W. Siede, *DNA Repair and Mutagenesis*, American Society for Microbiology, Washington, **1995.**

[2] B. B. Zhou, S. J. Elledge, The DNA damage response: putting checkpoints in perspective, *Nature,* **2000**, *408*, 433.

[3] de Laat, W.L., Jaspers, N.G. & Hoeijmakers, J.H., Molecular mechanism of nucleotide excision repair, *Genes Dev.,* **1999**, *13*, 768-785.

[4] Gillet, L. C.; Schärer, O. D., Molecular Mechanisms of Mammalian Global Genome Nucleotide Excision Repair, *Chem. Res.,* **2006**, *106*, 253-276.

[5] Svejstrup, J. Q., Mechanisms of transcription-coupled DNA repair, *Nat. Rev. Mol. Cell Biol.,* **2002**, 3, 21-29.

[6] Lehmann, A. R., DNA repair-deficient diseases, xeroderma pigmentosum, Cockayne syndrome and trichothiodystrophy, *Biochimie*, **2003**, 85, 1101-1111.

[7] Schärer, O.D., Chemistry and biology of DNA repair, *Angew. Chem. Int. Ed.* **2003**, *42*, 2946-74.

[8] Xeroderma Pigmentosum. Genetic People Website 2011; available from http://geneticpeople.com

[9] Children of the moon; available from http://ds9.ssl.berkeley.edu/lws_gems/4/uve.htm

[10] Sugasawa, K. Ng, J. M. Y., Masutani C., Iwai S., Spek, P. J., Eker, A. P. M., Hanaoka, F., Bootsma D. and Hoeijmakers, J. H. J., Xeroderma pigmentosum group C protein complex is the initiator of global genome nucleotide excision repair, *Mol. Cell*, **1998**, *2*, 223-32.

[11] Sugasawa, K., Okamoto, T., Shimizu, Y., Masutani, C., Iwai, S., and Hanaoka, F., A multistep damage recognition mechanism for global genomic nucleotide excision repair, *Genes Dev.*, **2001**, *15*, 507-21.

[12] Min, J. H. and Pavletich, N. P., Recognition of DNA damage by the Rad4 nucleotide excision repair protein, *Nature*, **2007**, *449*, 570-5.

[13] Yokoi, M., Masutani, C., Maekawa, T., Sugasawa, K., Ohkuma, Y. and Hanaoka, F., The xeroderma pigmentosum group C protein complex XPC-HR23B plays an important role in the recruitment of transcription factor IIH to damaged DNA, *J. Biol. Chem.*, **2000**, *275*, 9870-5.

[14] Evans, E., Moggs, J. G., Hwang, J. R., Egly, J. M., and Wood, R. D., Mechanism of open complex and dual incision formation by human nucleotide excision repair factors, *Embo J.*, **1997,** *16*, 6559-73.

[15] Wakasugi, M.; Sancar, A., Assembly, subunit composition, and footprint of human DNA repair excision nuclease, *Proc Natl Acad Sci U S A*, **1998**, *95*, 6669-74.

[16] Volker, M., Moné, M.J., Karmakar, P., van Hoffen, A., Schul, W., Vermeulen, W., Hoeijmakers, J.H., van Driel, R., van Zeeland, A.A., Mullenders, L.H., Sequential assembly of the nucleotide excision repair factors in vivo, *Mol. Cell*, **2001**, *8*, 213-24.

[17] Riedl, T., Hanaoka, F., and Egly, J., The comings and goings of nucleotide excision repair factors on damaged DNA, *Embo J.*, **2003**, *22*, 5293-303.

[18] Matsunaga, T., Mu, D., Park, C.H., Reardon, J.T., Sancar, A., Human DNA repair excision nuclease. Analysis of the roles of the subunits involved in dual incisions by using anti-XPG and anti-ERCC1 antibodies, *J. Biol. Chem.*, **1995**, *270*, 20862-9.

[19] Sijbers, A. M., de Laat, W.L., Ariza, R.R., Biggerstaff, M., Wei, Y.F., Moggs, J.G., Carter, K.C., Shell, B.K., Evans, E., de Jong, M.C., Rademakers, S., de Rooij, J., Jaspers, N.G., Hoeijmakers, J.H., Wood, R.D., Xeroderma pigmentosum group F caused by a defect in a structure-specific DNA repair endonuclease, *Cell*, **1996**, *86*, 811-22.

[20] Tapias, A., Ordered conformational changes in damaged DNA induced by

nucleotide excision repair factors, *J. Biol. Chem*., **2004**, *279*, 19074-83.

[21] Lidija, S., Fagbemi, A. F., Enzlin, J. H., Gourdin, A. M., Wijgers, N., Dunand-Sauthier, I., Giglia-Mari, G., Clarkson, S. G., Vermeulen, W. and Schärer, O. D., Coordination of dual incision and repair synthesis in human nucleotide excision repair, *EMBO*, **2009**, *28*, 1111-1120.

[22] Hanawalt, P.C., Repair replication in the bacterial genome. In Genetical Aspects of Radiosensitivity: *Mechanisms of Repair*, **1966**, *24*, 97–104.

[23] Hanawalt, P.C., Haynes, R.H., The repair of DNA. *Sci. Am*., **1967**, *216*: 36–43.

[24] Bohr, V. A., Smith, C.A., Okumoto, D.S., and Hanawalt, P.C., DNA repair in an active gene: removal of pyrimidine dimers from the DHFR gene of CHO cells is much more efficient than in the genome overall, *Cell*, **1985**, *40*, 359-369.

[25] Gool, A.J.V., Citterio, E., Rademakers, S., Os, R.V., Vermeulen, W., Constantinou, A., Egly, J., Bootsma, D., and Hoeijmakers, J.H.J., The cockayne syndrome B protein, involved in transcription-coupled DNA repair, resides in an RNA polymerases II-containing complex, *EMBO*, **1997**, *16,* 5955-5965.

[26] Bergmann, E., and Egly, J., Trichothiodystrophy, a transcription syndrome, *Trend in Genetics*, **2001**, *17*, 297-286.

[27] Mellon, I., Spivak, G., and Hanawalt, P.C., Selective removal of transcription-blocking DNA damage from the transcribed strand of the mammalian DHFR gene, *Cell*, **1987**, *51*, 241-249.

[28] Fousteri, M., and Mullenders, L.H.F., Transcription-coupled nucleotide excision repair in mammalian cells: molecular mechanisms and biological effects, *Cell Research*, **2008**, *18*, 73-84.

[29] Boer, J. D. and Hoeijmakers, J.H.J., Nucleotide Excision repair and human syndromes, Carcinovenesis, **2000**, *21*, 453-460.

[30] Gilchrest, B. A., Photodamage, *Blackwell Science Inc*., Cambridge, MA, **1995**.

[31] Sarasin, A., An overview of the mechanisms of mutagenesis and carcinogensis, *Mut. Res*., **2003**, *544*, 99-106.

[32] McAteer, K., Jing, Y., Kao, J., Taylor, J.S., Kennedy, M.A*.,* Solution-state structure

of a DNA dodecamer duplex containing a *cis-syn* thymidine cyclobutane dimmer, the major UV photoproduct of DNA, *J. Mol. Biol*., **1998**, *282*, 1013-1032.

[33] Lee, J., Park, C.J., Shin, J.S., Ikegami, T., Akutsu, H., Choi, B.S., NMR structure of the DNA decamer duplex containing double T*G mismatches of cis-syn cyclobutane pyrimidine dimmer: implications for DNA damage recognition by the XPC-hHR23B complex, *Nucleic Acids Res*., **2004**, *32*, 2474-2481.

[34] Kim, J., Choi, B., The solution structure of DNA duplex-decamer containing the (6-4) photoproduct of thymidylyl (3'-5')thymidine by NMR and relaxation matrix refinement, *Eur. J. Biochem*., **1995**, *228*, 849-854.

[35] Szymkowski, D. E., Lawrence, C. W., and Wood, R. D., Repair by human cell extracts of single (6-4) and cyclobutane thymine-thymine photoproducts in DNA, *Proc. Natl. Acad. Sci. USA*, **1993**, *90*, 9823-9827.

[36] Jamieson, E. R. and Lippard, S. J., Structure, recognition, and processing of cisplatin-DNA adducts, Chem. Rev. **1999**, *99*, 2467-2498.

[37] Harvey, R. G., Polycyclic aromatic hydrocarbons: chemistry and carcinogenicity; *University press: Cambridge*, **1991**.

[38] Geacintov, N. E., Cosman, M., Hingerty, B.E., Amin, S., Broyde, S., Patel, D.J., NMR solution structures of stereoisomeric covalent polycyclic aromatic carcinogen-DNA adducts: principles, patterns, and diversity, *Chem. Res. Toxcol.***1997**, *10*, 112-146.

[39] Gillet, L. C.; Schärer, O. D., Preparation of C8-amine and acetylamine adducts of 2'-deoxyguanosine suitably protected for DNA synthesis, *Org. Lett*., **2002**, *4*, 4205-4208.

[40] Patel, D. J., Mao, B., Gu, Z., Hingerty, B.E., Gorin, A., Basu, A.K., Broyde, S., Nuclear magnetic resonance solution structures of covalent aromatic amine-DNA adducts and their mutagenic relevance, *Chem. Res. Toxicol*., **1998**, *11*, 391-407.

[41] Heflich, R. H.; Neft, R. E., Genetic toxicity of 2-acetylaminofluorene, 2-aminofluorene and some of their metabolites and model metabolites, *Mutat. Res*., **1994**, *318*, 73-114.

[42] Hoffmann, G. R.; Fuchs, R. P., Mechanisms of frameshift mutations: insight from aromatic amines, *Chem. Res. Toxicol*., **1997**, *10*, 347-59.

[43] Belguise-Valladier,; P. Fuchs, R. P., N-2-aminofluorene and N-2 acetylaminofluorene adducts: the local sequence context of an adduct and its chemical structure determine its replication properties, *J. Mol. Biol*., **1995**, *249*, 903-13.

[44] Doisy, R.; Tang, M. S., Effect of aminofluorene and (acetylamino)fluorene adducts on the DNA replication mediated by Escherichia coli polymerases I (Klenow fragment) and III, *Biochemistry*, **1995**, *34*, 4358-68.

[45] Shibutani, S., Suzuki, N., Grollman, A.P., Mutagenic specificity of (acetylamino) fluorene-derived DNA adducts in mammalian cells, *Biochemistry*, **1998**, *37*, 12034.

[46] Gunz, D., Hess, M.T., Naegeli, H., Recognition of DNA adducts by human nucleotide excision repair, evidence for a thermodynamic probing mechanism, *J. Mol. Biol*., **1996**, *271*, 25089-25098.

[47] Hess M.T., Schwitter, U., Petretta, M., Giese, B., and Naegeli H. , Bipartite substrate discrimination by human nucleotide excision repair, *Proc. Natl. Acad. Sci. USA*, **1997**, *94*, 6664-6669.

[48] Hess, M.T., Schwitter, U., Petretta, M., Giese, B., Naegeli, H., Site-specific DNA substrates for human excision repair: comparison between deoxyribose and base adducts, *Chem. Biol*. **1996**, *3*, 121-128.

[49] Grunberger, D., Nelson, J. H., Cantor, C. R., and Weinstein, I. B., Coding and conformational properties of oligonucleotides modified with the carcinogen N-2-acetylaminofluorene, *Proc. Natl. Acad. Sci. U.S.A*. **1970**, *66*, 488-494.

[50] Fuchs, R. P. P., Lefevre, J.F., Pouyet, J., Daune, M.P., Comparative orientation of the fluorene residue in native DNA modified by N-acetoxy-N-2-acetylaminofluorene and two 7-halogen derivatives, *Biochemistry*, **1976**, *15*, 3347-3351.

[51] Hess, M.T., Gunz, D., Luneva, N., Geacintov, N.E., Naegeli, H., Base pair conformation-dependent excision of benzo[α]pyrene diol expoxide-guanine adducts by human nucleotide excision repair enzymes, *Mol. Cell. Biol*., **1997**, *17*, 7069-7076.

[52] Buterin, T., Hess, M.T., Luneva, N., Geacintov, N.E., Amin, S., Kroth, H., Seidel, A., Naegeli, H., Unrepaired fjord region polycyclic aromatic hydrocarbon-DNA adducts in ras codon 61 mutational hot spots, *Cancer research*, **2000**, *60*, 1849-1856.

[53] Yan, S., Wu, M., Buterin, T., Naegeli, H., Geacintov, N.E., Broyde, S., Role of base sequence context in conformational equilibria and nucleotide excision repair of

benzo[α]pyrene diol expoxide-adenine adducts, *Biochemistry*, **2003**, *42*, 2339-2354.

[54] Geacintov, N. E., Broyde, S., Buterin, T., Naegeli, H., Wu, M., Yan, S., Patel, DJ., Thermodynamic and structural factors in the removal of bulky DNA adducts by the nucleotide excision repair machinery, *Biopolymers*, **2002**, *65*, 202-210.

[55] Isaacs, R.J. and Spielmann, H.P., A model for initial DNA lesion recognition by NER and MMR based on local conformational flexibility, *DNA rep*., **2004**, *3*, 455-464.

[56] Mathieu, N., Kaczmarek, N., and Naegeli, H., Strand- and site-specific DNA lesion demarcation by the xeroderma pigmentosum group D helicase, *Proc. Natl. Acad. Sci. U.S.A*. **2010**, *107*, 17545-17550.

[57] Sugasawa, K**.,** Akagi J.**,** Nishi, R.**,** Iwai, S. and 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*, **2009**, *36*, 642-653.

[58] Zhou, L., Rajabzadeh, M., Traficante, D. D., and Cho, B. P., Conformational heterogeneity of arylamine-modified DNA: ^{19}F NMR evidence, *J. Am. Chem. Soc.,* **1997**, *119*, 5384-5389.

[59] Schut, H. A.; Snyderwine, E. G., DNA adducts of heterocyclic amine food mutagens: implications for mutagenesis and carcinogenesis, *Carcinogenesis*, **1999**, *20*, 353-68.

[60] Gillet, L. C., Alzeer, J. and Schärer, O.D., Site-specific incorporation of N-(deoxyguanosin-8-yl)-2- acetylaminofluorene (dG-AAF) into oligonucleotides using modified 'ultra-mild' DNA synthesis, *Nucleic Acids Res*., **2005**, *33*, 1961-9.

[61] Stöhrer, G., Osband, J.A. and Alvarado-Urbina, G., Site-Specific modification of the lactose operator with acetylaminofluorene. *Nucleic Acids Res*. **1983**, *11*, 5093-5101.

[62] Shivji, M. K. K., Moggs, J.G., Kuraoka, I. and Wood, R. D., Dual-incision assays for nucleotide excision repair using DNA with a lesion at a specific site, *Method Mol. Biol.,* **1999**, *113*, 373-92.

[63] Sarges, R., Hank, R.F., Blake, J.F., Bordner, J., Bussolotti, D.L., Hargrove, D.M., Treadway, J.L., Gibbs, E.M., Glucose transport-enhancing and hypoglycemic activity of 2-methyl-2-phenoxy-3-phenylpropanoic acids, *J. Med. Chem*. **1996**, *39*, 4783-4803.

[64] Batty, D., Rapic'-Otrin, V., Levine, A.S., Wood, R.D., Stable binding of human XPC complex to irradiated DNA confers strong discrimination for damaged sites, *J. Mol. Biol*., **2000**, *300*, 275-390.

[65] Wood, R.D., Robins, P. and Lindahl, T. Complementation of the xeroderma pigmentosum DNA repair defect in cell-free extracts. *Cell*, **1988**, 53, 97-106.

[66] Shibutani, S., Gentles, R., Johnson, F., Grollman, A.P., Isolation and characterization of oligodeoxynucleotides containing dG-N^2-AAF oxidation products of dG-C8-AF, *Carcinogenesis*, **1991**, *12*, 813-818.

[67] Jen Bohon and Carlos R. de los Santos, Effect of 6-thioguanine on the stability of duplex DNA, *Nucleic Acids Research*, **2005**, *33*, 2880-2886.

[68] Zaliznyak, T., Bonala, R., Johnson, F., de Los Santos, C., Structure and stability of duplex DNA containing the dG-(N2)-AAF lesion: a bulky adduct that persists in cellular DNA, *Chem. Res. Toxicol*., **2006**,*19*, 745-752.

[69] Allawi, H. T. and Jr. SantaLucia, J., Thermodynamics and NMR of internal GT Mismatches in DNA, Biochemistry, **1997**, *36*, 10581-10594.

[70] Hranjec, T., Kovaè, A., Kos, J., Mao, W., Chen, J. J., Grollman, A. P., Jelakoviæ, B., Endemic nephropathy: the case for chronic poisoning by aristolochia, *Croat. Med. J.*, **2005**, *46*, 116-125.

[71] Arlt, V. M., Stlborova, M. and Schmelser, H.H., Aristolochic acid as a problem human cancer hazard in herbal remedies: a review, *Mutagenesis*, **2002**, *17*, 265-277.

[72] Mengs, U., Tumour induction in mice following exposure to aristolochic acid, *Arch. Toxicol.,* **1988**, 61, 504-505.

[73] Grollman, A. P., Shibutani, S., Moriya, M., Miller, F., Wu, L., Moll, U., Suzuki, N., Fernandes, A., Rosenquist, T., Medverec, Z., Jakovina, K., Brdar, B., Slade, N., Turesky, R.J., Goodenough, A.K., Rieger, R., Vukelić, M., Jelaković, B.. Aristolochic acid and the etiology of endemic (Balkan) nephropathy, *Proc. Natl Acad. Sci*., **2007**, *104*, 12129-12134.

[74] Kohara, A**.,** Suzuki, T., Honma, M., Ohwada, T. and Hayashi, M.**,** Mutagenicity of Arisotolochic acid in the *Lambda1lac*Z transgenic mouse (Muta ™ Mouse), *Mutat. Res*., **2002,** *515*, 63-72.

[75] Schmeiser, H.H., Pool, B.L., and Wiessler, M., Identification and mutagenicity of metabolites of aristolochic acid formed by rat liver. *Carcinogenesis,* **1986**, *7***,** 59–63.

[76] Schmeiser, H. H., Schoepe, K.B., Wiessler, M., DNA adduct formation of aristolochic acid I and II in vitro and in vivo, *Carcinogenesis*, **1988**, *9*, 297-303.

[77] Krumbiegel, G., Hallensleben, J., Mennicke, W.H., Rittmann, N., Roth, H.J., Studies on the metabolism of aristolochic acids I and II, *Xenobiotica*, **1987**, *17*, 981-991.

[78] Stiborová, M., Hájek, M., Frei, E., Schmeiser, H.H., Carcinogenic and nephrotoxic alkaloids aristolochic acids upon activation by NADPH:cytochrome P450 reductase form adducts found in DNA of patients with Chinese herbs nephropathy, *Carcinogenesis*, **2001**, *9*, 297-303.

[79] Attaluri, S., Bonala, R. R., Yang, I., Lukin, M. A., Wen, Y., Grollman, A.P., Moriya, M., Iden, C.R., and Johnson, F., DNA adducts of aristolochic acid II: total synthesis and site-specific mutagenesis studies in mammalian cells, *Nucleic Acids Res.* **2010**, *38*, 339-352.

[80] Hanawalt, P. C. and Spivak, G., Transcription-coupled DNA repair: two decades of progress and surprises, *Nat. Rev. Mol. Cell Biol*., **2008**, *9*, 958-970.

[81] Scrima, A.**,** Koníčková, R.**,** Czyzewski, B. K.**,** Kawasaki, Y., et al., Structural Basis of UV DNA-Damage Recognition by the DDB1–DDB2 Complex, *Cell*, **2008**, *135*, 1213-1223.

[82] Buterin, T., Meyer, C., Giese, B., Naegeli, H., DNA quality control by conformational readout on the undamaged strand of the double helix, Chem. Biol., **2005**, *12*, 913-922.

[83] Brueckner, F., Hennecke, U., Carell, T., Cramer, P., CPD damage recognition by transcribing RNA polymerase II, *Science*, **2007**, *315*, 859-862.

[84] Damsma, G. E., Alt A., Brueckner, F., Carell, T., Cramer P.*,* Mechanism of transcriptional stalling at cisplatin-damaged DNA, *Nat. Struct. Mol. Biol*., **2007**, *14*, 1127 – 1133.

[85] Martin, K. R., Jokinen, M.P., Heneycutt, H.P., Quinn, A., Kari, F. W. , Barrett, J. C. and French, J.E., Tumor Profile of Novel p53 Heterozygous Tg.AC (v-Ha-ras) Bitransgenic Mice Treated with Benzo(a)pyrene and Fed Dietary N-acetyl-L-cysteine (NAC), *Toxicol. Sci*., **2004**, *81*, 293-301.

[86] Shibutani, S., Dong, H., Suzuki, N., Ueda, S., Miller, F., and Grollman, A. P., Selective Toxicity of Aristolochic Acids I and II, *Drug Metab. Dispos*., **2007**, *37*, 1217-1222.

[87] Furuta, T., Ueda, T., Aune, G., Sarasin, A., Kraemer, K.H., Pommier, Y., Transcription-coupled nucleotide excision repair as a determinant of cisplatin sensitivity of human cells, *Cancer Res*., **2002**, *62*, 4899-4902.

[88] Moriya, M., Slade, N., Brdar, B., Medverec, Z., Tomic, K., Jelaković, B., Wu, L., Truong, S., Fernandes, A., Grollman, A.P., TP53 mutational signature for aristolochic acid: an environmental carcinogen, *Int. J. Cancer*, **2011**, DOI:10.1002/ijc.26077

[89] Bakhiy, N., Arlt, V.M., Bahn, A., Burckardt, G., Phillips, D.H., Glatt, H., Molecular evidence for an involvement of organic anion transporters (OATs) in aristolochic acid nephropathy, *Toxicology*, **2009**, *264*, 74-79.

[90] Trego, K. S. and Turchi, J. J., Pre-steady-state binding of damaged DNA by XPC-hHR23B reveals a kinetic mechanism for damage discrimination, *Biochemistry*, **2006**, *45*, 1961-1969.

[91] Biggerstaff, M. and Wood, R. D., Assay for nucleotide excision repair protein activity using fractionated cell extracts and UV-damaged plasmid DNA, *Methods Mol. Biol.,* **1999**, *113*, 357-372.

[92] Tornøe, C. W., Christensen C., and Meldal M., Peptidotriazoles on Solid Phase: [1,2,3]-Triazoles by Regiospecific Copper(I)-Catalyzed 1,3-Dipolar Cycloadditions of Terminal Alkynes to Azides, *J. Org. Chem.* **2002**, *67*, 3057-3064.

[93] Rostovtsrv, V. V., Green, L. G., Fokin, V. V., Sharpless, K. B., A Stepwise Huisgen Cycloaddition Process: Copper (I)-Catalyzed Regioselective "Ligation" of Azides and Terminal Alkynes, *Angew. Chem. Int. Ed*., **2002**, *41*, 2596-2599.

[94] Best, M. D., Click Chemistry and Bioorthogonal Reactions: Unprecedented Selectivity in the Labeling of Biological Molecules, *Biochemistry*, **2009**, *48*, 6571-6584.

[95] Binder, W. H. and Sachsenhofer, R., The present review focuses on the azide-alkyne 'click' reaction catalyzed by Cu^{1}-species, *Macroml, Rapid Commun.* **2008**, *29*, 952-981.

[96] Moorhouse, A. D., Moses, J. E., Click Chemistry and Medicinal Chemistry: A Case of "Cyclo-Addiction"**,** *ChemMedChem*., **2008**, *3*, 715-723.

[97] Kolb, H. C. and Sharpless, K. B., The growing impact of click chemistry on drug discovery, *Drug discovery tody*, **2003**, *8*, 1128-1137.

[98] Agard, N. J., Baskin, J. M., Prescher, J. A., Lo, A. and Bertozzi, C. R., A Comparative Study of Bioorthogonal Reactions with Azides, *ACS Chemical Biology*, **2006**, *10*, 644-648.

[99] Seela, F., Sirivolu, V.R., Chittepu, P., Modification of DNA with octadiynyl side chains: synthesis, base pairing, and formation of fluorescent coumarin dye conjugates of four nucleobases by the alkyne-azide "click" reaction, *Bioconjug. Chem*., **2008**, *19*, 211-224.

[100] Chan, T. R., Hilgraf, R., Sharpless, K. B., and Fokin, V. V., Polytriazoles as Copper (I)-Stabilizing Ligands in Catalysis, *Org. Lett*., **2004**, *6*, 2853-2855.

[101] Rodionov, V. O., Presolski, S. I., Díaz, D. D., Fokin, V. V., and Finn, M. G., Ligand-Accelerated Cu-Catalyzed Azide−Alkyne Cycloaddition: A Mechanistic Report, *J. Am. Chem. Soc*., **2007**, *129*, 12696-12712.

[102]Gierlich, J., Burley, G. A., Gramlich, P. M. E., Hammond, D. M., and Carell, T., Click chemistry as a reliable method for the high-density postsynthetic functionalization of alkyne-modified DNA, *Org. Lett*. **2006**, *8*, 3639-3642.

[103] Wang, Q., Chan, T. R., Hilgraf, R., Fokin, V. V., Sharpless, K. B., and Finn, M. G., Bioconjugation by Copper (I)-Catalyzed Azide-Alkyne [3+2] Cycloaddition, *J. Am. Chem. Soc*., **2003**, *125*, 3192-3193.

[104] Hong, V., Presolski, S. I., Ma, C., and Finn, M. G., Analysis and Optimization of Copper-Catalyzed Azide-Alkyne Cycloaddition for Bioconjugation, *Angew. Chem. Int. Ed*., **2009**, *48*, 9879-9883.

[105] Seela, F., and Sirivolu, V. R., Nucleosides and Oligonucleotides with Diynyl Side Chains: Base Pairing and Functionalization of 2'-Deoxyuridine Derivatives by the Copper (I)-Catalyzed Alkyne Azide Click Cycloaddition, *Helv. Chim. Acta,* **2007**, *46*, 4184-4187.

[106] Ming, X., Leonard, P., Heindl, D., and Seela, F., Azide-alkyne "click" reaction performed on oligonucleotides with the universal nucleoside 7-ocatadinyl-7-deaza-2'-deoxyinosine, *Nucleic Acids Symposium Series,* **2008**, 52, 471-472.

[107] Gupta, S. S., Kuzelka, J., Singh, P., Lewis, W. G., Manchester, M., and Finn, M. G., Accelerated bioorthogonal conjugation: A practical method for the ligation of diverse functional molecules to a polyvalent virus scaffold, *Bioconjugate Chem.*, **2005**, *16*, 1572-1579.

[108] Seela, F. and Sirivolu, V. R., DNA Containing Side Chains with Terminal Triple Bonds: Base-Pair Stability and Functionalization of Alkynylated Pyrimidines and 7-Deazapurines, *Chem. Biodiversity*, **2006**, *3*, 509-514.

[109] Hammond, D. M., Manetto, A., Gierlich, J., Azov, V. A., Gramlich, P. M. E., Burley, G. A., Maul, M., and Carell, T., DNA Photography: An Ultrasensitive DNA-Detection Method Based on Photographic Techniques, *Angew. Chem. Int. Ed*., **2007**, *46*, 4184-4187.

[110] Haldsrud, Renate and KrØkje, Åse, Induction of DNA Double-Strand Breaks in the H4IIE Cell Line Exposed to Environmentally Relevant Concentrations of Copper,

Cadmium, and Zinc, Singly and in Combinations, *Journal of Toxicology and Environmental Health*, **2009**, Part A, *72*, 155-163.

[111] Lloyd, D. R. and Phillips, D. H.**,** Oxidative DNA damage mediated by copper (II), iron(II) and nickel(II) fenton reactions: evidence for site-specific mechanisms in the formation of double-strand breaks, 8-hydroxydeoxyguanosine and putative intrastrand cross-links, *Mutation Research,* **1999**, *424*, 23-36.