TOLERANCE OF THYMINE GLYCOL AT THE BLUNT DNA DOUBLE STRAND BREAK FOR NONHOMOLOGUS END JOINING REPAIR AND INTERFERENCE BY BASE EXCISION REPAIR

Sri Lakshmi Chalasani
TOLERANCE OF THYMINE GLYCOL AT THE BLUNT DNA DOUBLE STRAND BREAK FOR NONHOMOLOGUS END JOINING REPAIR AND INTERFERENCE BY BASE EXCISION REPAIR

A thesis submitted in partial fulfillment of the requirements for the degree of Master of Science at Virginia Commonwealth University

By

Sri Lakshmi Chalasani
M.B.B.S, MD. NTR University of Health Sciences, 2012

Advisor: Lawrence F. Povirk, Professor, Department of Pharmacology and Toxicology

Virginia Commonwealth University

Richmond, Virginia

July, 2015
ACKNOWLEDGEMENT

I would like to give my heartfelt thanks to Dr. Lawrence Povirk who has made this possible for me by accepting me as his student. He is an excellent advisor providing me with constant support. Being an amateur myself in this field, I had much to learn. Dr. Povirk has had a lot of patience to train me all the while during my project. He has always been there constantly guiding me in my each and every step in the lab work. Dr. Povirk dedication to work is beyond compare. I have been immensely inspired by his knowledge and approach in science. He is a wonderful scientist and a great human being. The birthday song he played for me is which I can never forget in my life. I am more than happy to be under his guidance for these two years and looking forward for much more to learn from him in future.

My sincere thanks go to Dr. Konstantin Akopiants. He has been a great teacher explaining me everything in great detail. He has been a good advisor in my professional and personal matters. I really appreciate his very helpful nature. He is the person one can approach for anything and sure to get it done. I was treated as his kid with many small gifts all the time. I would say he is a valued asset. I would like to thank Dr. Matthew Hartman for his valuable advice and the great ideas regarding another project, which I could use during this project. Dr. Hartman’s instantaneous approach and persistence inspired me a lot.

I could ask for no one else but my colleagues Mohammed Almohaini, Ajinkya Kawalae and Duaa Bafail. I am really fortunate to have these guys who made feel at home. Mohammed’s and Ajinkya’s advice over my project were really valuable. They trained me things which I could never possibly figure out by myself. They always had my back in my sad and disappointing moments and made me strong all over again. I wish them all success in whatever they enter in future. I would like to thank Dr. Tong Zhou for maintaining the lab atmosphere healthy and conducive for experiments. I would like to thank Dr. Hartman lab members, especially David for teaching me chemistry and how to use instruments.

My gratitude goes to Dr. Akbar Ali for accepting me in to the department and guiding me all through the program. I appreciate his dedication to see that his graduate students are up to date and comfortable. I am also grateful to Dr. William Dewey for his valuable advice regarding my professional career and giving me an opportunity to be a part of this wonderful department and lab in future too.

I am indeed lucky to have my friends Ajinkya, Varsha who always have been a great support in my personal and professional life here. I am indebted to my friends Wisam Toma and Madhavi Puchalapally and their loving family members, who gave me home away from home.
and made my survival possible here. I would also like to thank Ms. Pamela O Haney, a wonderful human being for always being there, taking me on tours and teaching me the life here.

Finally, I would like to thank my family, my mom and dad for believing in me and supporting me in whatever decision I make. The sacrifices they made put me in this position. If not for them I would stand nowhere. May GOD give them health, peace and prosperity.
# TABLE OF CONTENTS

List of Figures........................................................................................................ vii

List of Abbreviations.................................................................................................. viii

Abstract.................................................................................................................. xi

I. Introduction........................................................................................................... 1

1.1 Radiotherapy...................................................................................................... 1

1.2 DNA structure and Radiation induced DNA damage ........................................... 2

1.3 Double strand breaks ......................................................................................... 3

1.4 Clustered DNA damage ...................................................................................... 4

1.5 Thymine Glycol .................................................................................................. 8

1.6 DNA Damage Response ....................................................................................... 11

1.7 Repair Pathway Selection.................................................................................. 12

1.8 Homologous recombination Pathway ............................................................... 13

1.9 Non-homologous End Joining (NHEJ) Repair ................................................... 14
1.10 NHEJ Pathway.................................................................15
1.10.1 Ku.............................................................................19
1.10.2 DNA-PKcs .................................................................22
1.10.3 XRCC4 and DNA Ligase IV ...........................................25
1.10.4 XLF .............................................................................28
1.11 Base Excision repair pathway ........................................31
1.12 Human EndoIII like protein (NTH) and bacterial Endonuclease III.................................34
1.13 Klenow Fragment .............................................................34

II. Materials and Methods ..........................................................36

2.1 Substrate.............................................................................36
2.2 Ligation of oligomeric duplexes to the Kasl/BstAPI cleaved pUC19.................................36
2.3 Electro elution of the modified substrate .........................................................................37
2.4 Bus cell extract preparation.............................................................................................38
2.5 End joining assays using cell extracts ............................................................................39
2.6 Polyacrylamide gel electrophoresis..................................................................................40
2.7 End joining reaction with ddTTP and Endonuclease .........................................................40
2.8 Klenow Enzymes treatment .........................................................................................40
2.9 End joining by the recombinant proteins..........................................................41

2.10 Statistics........................................................................................................41

III Results

3.1 T4 DNA ligase assay to test interference with the end joining by thymine glycol........45

3.2 Ability of NHEJ to ligate the substrate with Tg5 in Bus cell extracts in the presence and absence of XLF: ..............................................................48

3.3 Effect of Endonuclease and ddTTP on the end joining........................................51

3.4 Comparison of End Joining and Base Excision repair between Tg3 and Tg5 in cell extracts: .........................................................................................................................................................53

3.5 Assessment for fill in and ligation after BER......................................................55

3.6 Effect of Tg position on total, head-head and head-tail end joining....................58

3.7 Time dependent end joining and base excision repair of Tg3 and Tg5 plasmid..........60

3.8 Recombinant protein assay.................................................................................62

IV. Discussion..........................................................................................................64

V. Conclusions.........................................................................................................71

References.............................................................................................................72
LIST OF FIGURES

Fig 1.1: Schematic representation of Clustered damage in DNA by the ionizing radiation………7

Fig 1.2: Formation of thymine glycol………………………………………………………………………………….10

Fig: 1.3: Schematic overview of NHEJ…………………………………………………………………………………18

Fig 1.4: Schematic representation Ku heterodimer subunits 70 and 80………………………………………..21

Fig 1.5 Schematic representation of DNA PKcs…………………………………………………………………….24

Fig 1.6 Schematic representation of XRCC4………………………………………………………………………….27

Fig 1.7: Schematic representation of LigaseIV……………………………………………………………………….27

Fig 1.8: Structure of the human Cernunnos/XLF protein……………………………………………………………30

Fig. 1.9: Oxidative DNA base damage repair through BER-pathway………………………………………33

Fig 2.1: Preparation of the substrate……………………………………………………………………………………42

Fig 2.2: Assay Protocol……………………………………………………………………………………………………43

Fig 2.3: Base excision repair of the substrates……………………………………………………………………….44

Fig 3.1: T4 Ligase Assay……………………………………………………………………………………………………47

Fig 3.2: The ability of NHEJ to rejoin DSBs with Tg at the fifth position (Tg5)………………………50

Fig 3.3: The effect of the ddTTP and Endonuclease on the joined fragment (44bp)…………………52
Fig 3.4: Comparison of the repair between Tg3 and Tg5……………………………………54

Fig 3.5: Testing with the Klenow and Klenow exonuclease-……………………………………57

Fig 3.6: Effect of Tg position on total, head-head and head-tail end joining…………………………………………………………………………………………59

Fig 3.7: Time dependent repair kinetics of Tg3 and Tg5…………………………………………61

Fig 3.8: Repair by the recombinant NHEJ proteins Ku, X4L4 and/or XLF of Tg0, Tg1, Tg2, Tg3 and Tg5………………………………………………………………………………63

LIST OF ABBREVIATIONS

AP apurinic/apyrimidinic
APE1 apurinic/apyrimidinidin endonuclease
ATM ataxia telangiectasia mutated
ATP adenosine triphosphate
ATR ATM and rad3-related
BER base excision repair
CtBP C-terminal binding protein
CtIP  CtBP-interacting protein
DNA  deoxyribonucleic acid
DNA - PK  DNA-dependent protein kinase
DNA - PKcs  DNA-dependent protein kinase catalytic subunit
dNTP  deoxynucleotides
ddTTP  dideoxythymidine
DSBs  double-strand breaks (in DNA)
EndoIII  Endonuclease III
Hr  hour(s)
HRR  homologous recombination repair
IR  ionizing radiation
MDC1  mediator of DNA damage checkpoint
μl  microliter
NHEJ  non-homologous end joining
PARP-1  poly(ADP-ribose) polymerase-1
SCID  severe combined immune-deficiency
sec  seconds
SDS  sodium dodecyl sulfate
SSBs  single-strand breaks (in DNA)
ssDNA  single-strand DNA
TBE  buffer solution mixture of Tris base, boric acid and EDTA
Tg  Thymine glycol
Tg1  Thymine glycol terminally located at double strand breaks
Tg2  Thymine glycol located at the second position of double strand breaks
Tg3  Thymine glycol located at the third position of double strand breaks
Tg5  Thymine glycol located at the fifth position of double strand breaks
V(D)J  variable, diversity, joining
XLF  XRCC4-like factor
X4L4  XRCC4-DNA ligase IV complex
XRCC1  X-ray cross-complementation factor 1
XRCC4  X-ray cross-complementation factor 4
ABSTRACT

TOLERANCE OF THYMINE GLYCOL AT THE BLUNT DNA DOUBLE STRAND BREAK FOR NHEJ REPAIR AND INTERFERENCE BY BASE EXCISION REPAIR

By Sri Lakshmi Chalasani MD.

A thesis submitted in partial fulfillment of the requirements for the degree of Master of Science at Virginia Commonwealth University

Advisor: Lawrence F. Povirk, Professor, Department of Pharmacology and Toxicology, Virginia Commonwealth University, Richmond, Virginia

Radiotherapy is the clinical application of the ionizing radiation to treat cancer. Ionizing radiation causes multiple modes of damage to the DNA damage such as SSBs, DSBs and modified bases such as thymine glycol. These lesions can exist as clusters in one or two helical turns of DNA. DNA double-strand breaks (DSBs) are extremely toxic to cells because they can lead to genomic rearrangements and even cell death. If base lesions accompany these DSBs, there will be a substantial hindrance for repair.
NHEJ is the primary DSB repair pathway in mammalian cells. HRR repairs single strand breaks (SSBs) or Double strand breaks (DSBs), during late S phase and G2 phase of the cell cycle, by using an undamaged copy of the DNA sequence, and is therefore largely error-free. The NHEJ pathway repairs DSBs without the requirement for sequence homology and can be error-free or error-prone, and is most active during G1 phase.

Thymine glycol (Tg) is the most common oxidation product of thymine. It is produced endogenously as a consequence of aerobic metabolism or via exogenous factors such as ionizing radiation (IR); it is one of the predominant types of base modifications produced by ionizing radiation. Due to clustering of radiation-induced damages, many DSBs are accompanied by damaged bases such as Tg at or near the DSB ends that may interfere with subsequent gap filling and ligation. The base excision repair pathway plays a major role in removal of thymine glycol from the damaged DNA strand. During NHEJ, after synapsis by Ku and DNAPKcs and processing of the DNA ends, XRCC4/Ligase IV complex ligates the DNA. This ligase activity is promoted by the interaction of XLF/XRCC4 filament with Ligase IV.

Linearized plasmids with Tg at the 5th -Tg5 positions from the broken end were subjected to a repair assay using XRCC4-like factor (XLF)-deficient cell extracts, with or without the addition of XLF and Endonuclease III and/or ddTTP and Klenow fragments. End joining of Tg5 was compared to plasmid with Tg at third position-Tg3 in extracts. In addition, the ability of purified NHEJ proteins Ku, DNAPKcs and XRCC4/Ligase IV, to repair the Tg1 (Tg at the end), Tg2 (Tg at the second position), Tg3 and Tg5 in the presence and absence of XLF was assessed.

The data indicated that the cell extract could ligate the Tg5 plasmids only in the presence of XLF. End joining of the Tg5 was less in comparison to Tg3 with base excision repair being more
active in Tg5 competing with the joining. Plasmids with Tg were treated with Endonuclease III and ddTTP to test whether the end joining occurred before or after Tg removal.

Endonuclease III and ddTTP treatment showed reduced intensity of the joined fragment suggesting that end joining occurred without removal of Tg in some cases. The extracts were not able to fill in the processed end by the BER though it has a 3′-OH which was filled in by the treatment with Klenow enzymes derived from E.coli DNA Polymerase I, following removal of extract proteins by proteolysis. When purified proteins were used to treat the plasmids, it is observed that there was increased efficiency of repair with increased distance of the Tg from the end may be due to less distortion of the ends as Tg is away from the end. While Tg1 and Tg2 required XLF presence for the repair, Tg3 and Tg5 could show a small amount of repair in its absence. XLF enhanced the repair of Tg3 and Tg5. While the repair by extract showed no repair in the same substrates without XLF, there was repair by purified proteins without XLF, suggesting there is some competition with the XRCC4/Ligase for the DSB by other proteins and XLF is required to overcome this.

In conclusion, cell extract was able to ligate the plasmid with Tg located at fifth position from the DSB but with lower efficiency compared to Tg3 plasmids. The base excision repair pathway is more functional if the modified base is far from the DSB. End joining by the purified proteins was proportional to the thymine glycol position from the DSB end. XLF was mandatory for repair of Tg1 and Tg2 by the purified proteins, and not for Tg3 and Tg5.
INTRODUCTION

1.1 Radiotherapy:

Clinical application of ionizing radiation, radiotherapy, is one crucial treatment option in cancer therapy apart from surgery and systemic therapy. More than 60% of the cancer patients receive radiation as a treatment modality today. Approximately, 45% of new cancer cases will receive radiotherapy (RT) and this proportion is increasing \cite{1}. Moreover, RT is responsible for 40% of cancer cure providing an excellent cost-effectiveness ratio \cite{2}. Radiotherapy can be used in various treatment settings as a definitive strategy or in multimodal settings as in adjuvant or neoadjuvant settings, with or without concomitant chemotherapy. Radiotherapy can rescue the surgical amputations, yielding better cosmesis and can also be used in palliative settings \cite{1, 3}. Radiotherapy can not only significantly prolong patient survival but can also improve the local and metastatic control rates of tumors.

The main subcellular target for the radiotherapy is DNA. The effect of radiotherapy can be a lethal damage, which is unrepairable and causes cell death irrevocably or sub lethal resulting in sub lesions which can be repaired by the repair response pathways in the cell under normal circumstances unless additional damage is added or interference with the pathways exist. Basically, three modalities of cell death occur, Mitotic\cite{4}, permanent arrest in G1\cite{5} or Apoptosis
Though the radiotherapy is directed to kill cancer cells, there is a probability of damage to the normal cells surrounding the tumor; radiotherapy by itself is pre-carcinogenic rendering normal cells to cancer cells and induces secondary cancers in primary tissue. Of all the lesions created in the DNA by radiotherapy, DSB are notorious for causing cell death\textsuperscript{[7-9]}.

Whenever there is cell damage with the radiation, as with the other modes of damage there is activation of robust repair pathways depending on the damage and the phase of the cell cycle. This activation of the repair pathways renders the tumor more radio resistant. Studies show that cells that are defective in the repair proteins are radiosensitive. Based on these data, attempts are being made to disrupt the various repair pathways using specific inhibitors and subsequent exposure to radiotherapy.

1.2 DNA structure and Radiation induced DNA damage:

Deoxyribonucleic acid (DNA) is a molecule that encodes the genetic instructions for the development and functioning of all known living organisms and many viruses. It exists as “double helix” formed by two biopolymer strands running antiparallel to each other. DNA is composed of multiple nucleotides of guanine (G), adenine (A), thymine (T), or cytosine (C) — with a monosaccharide sugar called deoxyribose and a phosphate group. The deoxyribose sugars are joined at both the 3′ -hydroxyl and 5′ -hydroxyl groups to phosphate groups in ester links, also known as "phosphodiester" bonds with an alternating sugar-phosphate sequence of the DNA back bone. Two polymers of the nucleotides run antiparallel to each other and held together by the hydrogen bonds. DNA sequences are highly conserved in evolution and its integrity is essential for the normal functioning of the cell. DNA is continuously under both physiological and pathological stress, expressed as DNA damage. Programmed DNA breaks are formed during
homologous recombination that occurs during meiosis \cite{10, 11}, during replication and during regulated genome rearrangements in lymphocytes [V (D) J] and germ cells \cite{12, 13, 14}. Mutagens such as oxidizing agents, alkylating agents, X-rays and also high-energy electromagnetic radiation such as ultraviolet light can produce DNA breaks. UV radiation can potentially cause DNA damage such as thymidine dimers \cite{15} and free radicals generated from hydrogen peroxide produce various forms of damage, including base modifications, especially of guanosine, and double-strand breaks \cite{16}. On an average about 150,000 bases undergo oxidative damage per day in a typical human cell \cite{17}. Ionizing radiation can produce lethality in cells by inducing DNA damage including single and double strand breaks, oxidation of purine and pyrimidine bases and crosslinks with proteins. Of all the oxidative lesions, double-strand breaks are the most dangerous and are difficult to repair. Attempted repair produces insertions and deletions, point mutations and chromosomal translocations \cite{18}. Two main pathways take part in the DNA double-strand breaks (DSBs) repair: homologous recombination and nonhomologous recombination. Either of the pathways have the propensity to be error prone. Loss of Heterozygosity (LOH) - loss of the only present allele of tumor suppressor genes can occur in HR and NHEJ is the primary cause of translocations \cite{19, 20} and dysfunctional telomeres fusion \cite{21}. These mutations ultimately result in cancer.

1.3 Double strand breaks:

DNA-double strand breaks are considered to be the most deleterious lesions caused by ionizing radiation in case of a defective repair \cite{22, 23}. DSB can also be produced by radiomimetic drugs, inducing free radical damage of deoxyribose and topoisomerase-II inhibitors, by inhibiting the joining of breaks created by topoisomerase II. DSB can also be produced from DNA SSBs during replication, by specific nucleases during V(D)J and class switch recombination in
vertebrate lymphocytes, meiotic recombination in germ cells and retroviral integration \[^{24}\]. DSBs naturally occurring at chromosome ends expose telomeres resulting in critical shortening, leading to human cell aging \[^{25}\]. DSBs can be as simple as the occurrence of two single-strand breaks (SSBs) within approximately 6-10 base pairs caused by a single radiation track or complex in association with abasic sites (apurinic or apyrimidinic) or damaged bases (oxidized purines and pyrimidines) at the clustered damage sites \[^{26,27}\]. Both simple and complex DSB induced by ionizing radiation can have 3’ blocking ends, e.g. 3’-phosphate or 3’-phosphoglycolate moieties \[^{29}\] or possess single-stranded overhangs of variable length. Repair of these complex lesions is more difficult than the isolated lesions and sometimes they are irreparable, enhancing the chromosomal aberrations and carcinogenesis.

1.4 Clustered DNA damage:

Clusters of the DNA lesions are the signature mark of a single track of radiation consisting of two or more individual lesions within one or two helical turns of the DNA \[^{28-32}\]. As shown in Fig 1.1 the lesions can be a complex DSB with damaged base and/or abasic (AP) sites directly adjacent to the DSB ends \[^{33-35}\] or non-DSB clusters with the combination of two or more oxidized base, strand break or other DNA lesion that do not form a DSB. Closely associated base lesions, such as oxidized guanine \[8-oxo-7, 8- dihydroguanine (8-oxoG)] and thymine glycol \[5, 6- dihydroxy-5, 6-dihydrothymine or thymine glycol (Tg)] with other types of base lesions or apurinic/apyrimidinic (AP) sites exist in clusters \[^{36-38}\]. In contrast, endogenous processes produce individual lesions which are more homogenously distributed \[^{39-41}\]. The damage caused by the radiation can be a direct deposition of energy or indirectly via formation of reactive chemical species in the surrounding of DNA \[^{42,43}\]. Radiolysis of water in the vicinity of the DNA is primarily attributed to the production of hydroxyl radicals \((\text{OH}^-)\). All four types of
bases are equally sensitive to the attack by the OH⁻, as the energy deposition is proportional to mass \[^{[42]}\]. The formation of these clusters depends on the quality of the radiation and the presence/absence of oxygen. The number of individual DNA lesions per unit of absorbed dose is about the same for low- and high-Linear Energy Transfer radiations \[^{[44]}\]. The distribution of these lesions is within a smaller number of sites (segments of DNA), implying a higher level of cluster complexity, i.e., the average number of lesions per cluster tends to increase with increasing LET. Under both aerobic and anaerobic conditions the complexity of the clusters increases with the dose of LET as predicted by the measure of SSB to DSB ratio. There will be a decrease in the cluster complexity as the oxygen levels go down, thus hypoxic condition contributing to radio-resistance in tumor therapy \[^{[45]}\]. The repair of the non-DSB cluster lesions is attempted by Base Excision Repair pathway proteins APE1 or PARP1, as evidenced from recent studies using primary human fibroblasts \[^{[46]}\] or cell extracts \[^{[47]}\]. The type of lesion and the inter-lesion distance and the orientation of nearby lesions (e.g., bi-stranded or tandem lesions) are the rate limitation to the process. There will be a reduction or inhibition of the processing enzymes such as the glycosylases and nucleases. This slow processing of the lesions leads to the persistence of these lesions until they hit a replication fork leading to the formation of chromosome breaks, mutations and misrepair \[^{[48, 49]}\]. There will be formation of inadvertent DSB from these clustered lesions during the repair or replication \[^{[50]}\]. Human AP-endonuclease-1 (hAPE1) is also capable of producing a DSB by cleaving the two opposite stranded AP sites. DSB can also result from the formation of nick in the intact strand, during processing of the apyrimidinic site on the opposite strand. There is a hierarchy of the repair of the lesions in case of three or more complex lesions in order to avoid the formation of DSB \[^{[53]}\]. There is a severe impairment of the excision of a base lesion such as 8-oxo-guanine, opposing an AP site or strand break until the AP site/strand break
is repaired, though it occurs slowly due to the opposed damaged base \[^{52}\]. DSBs clustered DNA lesions either resulting from the direct action of the radiation or as a result of processing of non-DSB clusters are repaired by the NHEJ pathways to produce less complex clusters \[^{53}\]. When the DSB is due to replication fork hitting the clustered SSB lesion, the repair is by a homologous recombination \[^{54-56}\].
Fig 1.1: Schematic representation of Clustered damage in DNA by the ionizing radiation

Alexandros et al., 2013
1.5 Thymine Glycol:

Thymine glycol is the most common damaged base formed due to oxidation of thymine or deamination and oxidation 5-methyl-cytosine by ionizing radiation \[57\], cancer therapy or endogenous aerobic metabolism \[58\]. Radiation generates oxidants that can react with thymine leading to the formation of thymine glycol, thymine peroxide, thymine hydro peroxide, and other oxidized forms of the base. Further reaction of these products yields urea. Approximately 10%-20% of the radiation or radiotherapy induced damage results in thymine base oxidation and fragmentation \[59\]. Thymine glycol has four diastereomers of, i.e. (5R, 6S), (5R, 6R), (5S, 6R) and (5S, 6S), but exists as either the 5R cis–trans pair [(5R, 6S) and (5R, 6R)] or the 5S cis–trans pair [(5S, 6R) and (5S, 6S)] in solution, due to epimerization at the C6 position \[60\]. Oxidation of thymidine or thymidine in the oligonucleotides preferentially yields (5R, 6S)-thymine glycol \[61, 62\]. Depending on the base that is oxidized, the counter base can be either A or G. Thymine glycol induces both structural and functional perturbations of the DNA. Free radical addition to the C (5)C (6) double bond of pyrimidines leads to a class of nonplanar ring-saturated lesions, unable to form hydrogen bonds with the opposite base taking conformations which are unstable in the solution \[62,63\]. Due to the interaction of the Tg with the polymerases, structural changes can be induced at both initiation and termination points resulting in the alteration of the relative amounts of termination bands of DNA synthesis \[64\]. Thymine glycol is a weak mutagen and can primarily block replication \[65-76\]. Studies using a template containing thymine glycol for primer extension studies with several DNA polymerases have shown that the extension occurs up to the site of Tg with adenine insertion and not beyond \[66-68\]. Thymine glycol is capable of blocking both the repair and replicative polymerases in vitro and can be lethal in vivo \[71, 72\]. However
there are several contexts in which thymine glycol is bypassed by certain DNA polymerases in vitro \cite{67} and in vivo \cite{65}. As Tg, like unmodified thymine, pairs with adenine it is a poor premutagenic lesion \cite{64}. If it is a simple thymine glycol lesion, it is removed via single nucleotide patch base excision repair pathway \cite{77}, but, if it accompanies a DSB a more robust cell response to this damage may be required.

Repair of the lesion is sequence \cite{64, 70} and counter-base dependent. The counter base has the propensity to either block the replication or promote misinsertions during translesional synthesis. Damaged base repair enzymes, Endonuclease III \cite{78} and its human orthologue hNHT1 via their bifunctional activity i.e., N-glycosylation and AP-Endonuclease activity, can release thymine glycol, producing a compatible DNA 3′ end. hNTH1 functions as the initial step of base excision repair(BER) to repair the hydrated, reduced, or oxidized, from the DNA backbone\cite{79}. 


Fig 1.2: Formation of thymine glycol: Reactive oxygen species (ROS), such as peroxide and hydroxyl radicals, are generated as byproducts during normal oxidative metabolism and ionizing radiation modify DNA bases. A common product of thymine oxidation is thymine glycol. Figure copied from Aller et al. 2007.
1.6 DNA Damage Response:

Sophisticated cellular networks that constantly monitor genome integrity are collectively termed the DNA damage response (DDR). DNA DSB’s are deleterious as there is no complementary strand to be used as a template for repair. Repair of these breaks is the function of two major pathways: Homologous recombination and Non-Homologous end joining repair. These repair pathways are complementary and function under different circumstances. As homologous repair requires a sister chromatid for template, it occurs in S and G2 phase of the cell cycle \cite{80, 81}. NHEJ is functional throughout the cell cycle and does not require a template. This makes NHEJ more error prone to insertions and deletions \cite{82}. When classical NHEJ is impeded because of a missing or mutated component, alternative end-joining pathways operate. These alternative end-joining pathways utilize the factors such as the MRE11–RAD50–NBS1 (MRN) complex, poly(ADP-ribose) polymerase-1 (PARP-1), XRCC1, and DNA Ligase I or III that are involved in SSB repair (SSBR) or HR and there is reliance on terminal micro-homologies for the joining reactions \cite{83}.

In response to DNA damage three distinct functionalities are activated: (1) Damage detection, (2) control of cell cycle and transcriptional programs of damage response, and (3) mechanisms for catalyzing the repair of the lesion \cite{84}.

Cell cycle progression and DNA repair are tightly regulated via cell cycle checkpoints due to activation of orchestrated cell signaling pathways \cite{85}. On detection of DNA damage, these pathways stall the cell cycle progression at crucial stages such as before or during DNA replication (G1/S and intra-S checkpoints) and before cell division (G2/M checkpoint), thereby preventing duplication and segregation of damaged DNA. MRN sensor complex detects DSB’s
and activates and recruits apical DDR kinase ATM (ataxia telangiectasia mutated) \[86\] while RPA recruits the ATR (ATM and rad3-related) kinase via its partner protein, ATRIP (ATR-interacting partner) at the regions of ssDNA exposed by DSB processing or ssDNA at stalled replication forks \[87\]. Mediator proteins for ATM such as BRCA1, MDC1 (mediator of DNA damage checkpoint) and 53BP1 (p53-binding protein 1) activate the effector kinases Chk1 and Chk2, spreading the signal throughout the nucleus stalling the progression of cell cycle. Reports show that ATM inhibit both RNA polymerase I- and II-dependent transcription locally at sites of DNA breaks in human cells\[88, 89\] and recruits DNA repair proteins.

1.7 Repair Pathway Selection:

The MRN complex plays a pivotal role in directing a DSB down the appropriate repair pathway, via its DNA end-processing activities \[90, 91\]. In DDR, MRN complex responds very early and acts as a DSB’s sensor \[92\]. In addition, it functions as a co-activator of cell cycle checkpoint signaling and a DSB repair effector in both the HR and NHEJ pathways \[93-97\]. Resection at DSBs plays a central role in determining the result of the competition between HR and NHEJ. While NHEJ is functional at any phase of cell cycle \[98\], HR is active during S and G2 phase, suppressing NHEJ during these phases. The factor which directs MRN complex mediated pathway selection is CtIP. CtIP concentrations are very low in G0/G1 phase, attributed to the proteosomal degradation during this phase. Upon progression to S/G2, CtIP protein levels are upregulated due to alleviation of this degradation \[99\]. CtIP is phosphorylated at Ser327 position in a CDK-dependent fashion, promoting complex formation between CtIP, MRN and BRCA1 in S/G2 \[100\]. This binding of CtIP to MRN complex, confers it with the exonuclease activity, and thus triggers 5’→ 3’ resection via ExoI and 3’→5’ resection via Mre11 activity \[100\]. Following this initial processing, other nucleases and helicases produce further resection of DSB generating
extensive 3’ ssDNA overhangs on each side of the break \cite{101-103}. The ssDNA thus generated is an excellent substrate for the HR-specific ssDNA-binding factors RPA and Rad51 \cite{91}, and a poor one for NHEJ-specific double strand DNA termini binding factor Ku70/80 \cite{104}. The phosphorylation state and low concentration of MRN-associated CtIP during G1 prevent the 5’ → 3’ resection, prevents the formation of ssDNA and gives an opportunity for the NHEJ proteins to process the repair.

**1.8 Homologous recombination:**

Homologous recombination is the DNA repair process which utilizes a homologous sequence, a homologous chromosome or a sister chromatid for accurate DNA repair. This is a highly conservative process. As discussed above, the initial sensor and responder to localize at the DSB is the MRN complex \cite{105} to form nuclear foci. MRN complex incites the further steps in HR mechanism by recruiting a phosphatidylinositol-3-related kinase, ATM, which phosphorylates histone H2AX at Ser139 position forming γH2AX. MDC1, a mediator protein facilitates the binding and phosphorylation between ATM and H2AX. By marking one or more megabases of DNA surrounding the break, γH2AX serves as a molecular beacon signaling the presence of damage \cite{116}. γH2AX forms as a scaffold protein for the accumulation of other proteins in the repair pathway. Formation of NBS1/MRE11/Rad50, 53BP1, and BRCA1 foci is also regulated by MDC1. The RING-finger ubiquitin ligase, RNF8, at its FHA domains interacts with phosphorylated MDC1 \cite{106-108}. RNF8 further ubiquitinylates H2AX and facilitates the accumulation of BRCA1 at the site of damage \cite{109,110}. BRCA1-C complex, BRCA1-CtIP, is formed by the direct interaction of BRCA-1 and phosphorylated CtIP, and is recruited by MRN complex \cite{111}. Cyclin dependent kinases phosphorylation of CtIP on Ser327 is essential for physical interactions between CtIP and BRCA1, which occurs only during G2 phase of cycle.
CtIP enrichment at the DSB site is MRN and ATM dependent. Upon binding to the MRN complex, CtIP catalyzes Mre11 endonuclease activity, to introduce nicks on the DNA duplexes as far as 300 nt from DSB ends. The nick thus produced triggers Exo1-dependent 5′-3′ resection, while the Mre11 exonuclease mediates 3′-5′ degradation of the nicked strand toward the break releasing any other proteins such as Ku(NHEJ factor) bound to the DSB, thus blocking NHEJ. This extensive DSB end resection generates single strand DNA (ssDNA) overhangs. ssDNA binding complex RPA (replication protein A) coats these ssDNA overhangs initially, later substituted by RAD51, forming a nucleoprotein filament. Other recombinant proteins such as RAD52/RAD54 stabilize this interaction. This RAD51 nucleofilament, together with various other HR factors, mediates strand invasion near a homologous sequence in sister chromatid or homologous chromosome. Once the strand is copied and filled in by DNA polymerases and ligated by Ligase I, DNA helicase and resolvase enzymes mediate the cleavage and resolution of HR intermediates.

1.9 Nonhomologous End Joining:

NHEJ, as the name suggests, involves direct joining of the DNA ends with minimal processing of the ends. DSB repair in mammals appears to be primarily by NHEJ. This mechanism is constitutively active throughout the cell cycle, although it is partially suppressed by certain HR factors in S/G2.

The three main steps of NHEJ are: (I) DNA end recognition and assembly and stabilization of the NHEJ complex at the DNA double strand break; (II) Bridging of the DNA ends and promotion of end stability; (III) DNA end processing; and (IV) Ligation of the broken ends and dissolution of the NHEJ complex.
XRCC5, XRCC6 and XRCC7 are the genes that encode some of the core NHEJ proteins that form DNA-dependent protein kinase (DNA-PK). XRCC5 and XRCC6 encode the 80 and 70 kDa subunits of the Ku70/80 heterodimer (the DNA-binding subunit of DNA-PK), and XRCC7 encodes the 470 kDa DNA-catalytic subunit of protein kinase DNA-PKcs\textsuperscript{[117]}. In addition to the above factors, a heterodimer XRCC4/Ligase IV is required for ligation of the strands\textsuperscript{[118]} and Cerrunos/XLF as an activator of ligase activity\textsuperscript{[119]}. Various end processing factors such as Artemis, TdT, WRN, PNK, hTDP1, FEN1 are required to clean the ends of the DNA of any factors that hinders the direct joining and DNA polymerase λ for filling the gaps created in the process. Proteins such as PARP-3 and APLF accelerate the NHEJ by retaining the XRCC4/Ligase-IV complex at the site of break\textsuperscript{[120]}.

1.10 NHEJ Pathway:

DSB associated with V (D)J recombination, irradiation or chemotherapy damage are primarily recognized by the MRN complex, which by further signaling via ATM, phosphorylates and activates H2AX to γH2AX. γH2AX recruits BRCA-1 or 53bp1, one of the deciding factors for the type of repair. In G1/M oligomers of 53bp1 flank on the either side of the DSB bound to the histones modified by acetylases and methyl transferases. The chromatin conformation is thus maintained, with limited access to transcription and HR factors\textsuperscript{[121]}, promoting NHEJ. Whenever there is a DSB, irrespective of the cell cycle Ku is the initial protein\textsuperscript{[122]} to translocate to the site and binds dsDNA ends with extremely high affinity\textsuperscript{[82]} and encircles the DNA with its preformed channel\textsuperscript{[123]}. This binding occurs in an ATP dependent manner\textsuperscript{[104]} and prevents the DNA ends from further damage. Ku functions as a signal molecule for the recruitment of further NHEJ proteins. Recently, it has been shown that Ku can process the ends of the DNA by the
virtue of its 5′-dRP/AP lyase activity, removing the abasic sites by nicking DNA 3′ of the abasic site which involves a Schiff-base covalent intermediate \[124\]. After its binding to the ends, there will be inward translocation of Ku and recruitment of DNA-PKcs, to form a Ku-DNA-PKcs synaptic complex (DNAPK) which tethers the DNA ends together and offers protection from premature nucleolytic degradation \[125, 126\]. Synapsis of the juxtaposed DNA ends stimulates auto phosphorylation of DNAPKcs by DNA-PK kinase activity, induces a conformational change in the DNA-Protein complex, providing access to nucleases and ligases as DNA-PKcs moves away \[126-128\]. As the IR produces complex DSB’s with 3′-phosphate or 3′-phosphoglycolate groups, 3′ and 5′ overhangs with damaged bases and/or ribose units in the vicinity of the break ends \[129, 130\], the DNA ends must be processed prior to ligation. Enzymes, such as Artemis, FEN-1, PNKP, WRN, MRN, hTDP1, Polymerases belonging to Pol X family aid in the processing step. Artemis removes the hairpin bend at the end at the V(D)J associated DSB, metnase trims the damaged ends \[131\], nucleases remove several nucleotides from single-stranded overhangs at the DSB ends termini, aprataxin removes adenylate groups \[132, 133\], the phosphodiesterases TDP1 and TDP2 process DNA topoisomerase adducts \[134\]. DNA polymerases belonging PolX family fill in the gap. Mre11/Rad50/ Nbs1 (MRN) complex plays a role in damage signaling and protection of the ends from degradation \[135\] Khanna et al., 2001). FEN-1 has an exonuclease function; WRN (Werner syndrome helicase) and BLM (Bloom syndrome helicase) are the unwinding enzymes \[136\].

In addition to DNAPKcs, DNA bound Ku also recruits XRCC4/Ligase IV complex and XLF independently \[137\]. The recruitment of the proteins does not require any protein-protein interactions but the assembly of the proteins does. XRCC4 requires DNAPKcs and XLF requires XRCC4 to be retained at the repair loci. Once recruited, XLF and XRCC4 form a nucleo-
filament around the DNA ends by interacting in a head-head manner. The architectural channel thus formed with the synergy of Ku-XLF-XRCC4 interaction stabilizes DSBs and provides accessibility and activation of LigIV, which ligates the DNA, ends \(^{138}\). Other proteins recruited to the DSB such as Aprataxin and PNK-like factor (APLF) enhance the stability of the filament \(^{133,139,140}\). As expected of it, NHEJ is an error prone pathway \(^{141}\), as there will be often removal of the nucleotides and fill in without the guidance of a template. Loss of nucleotides during the process makes the DSB repair inaccurate. Except for the recruitment of the Ku heterodimer, the order of recruitment of other factors is flexible depending on the complexity of the DNA damage \(^{142}\). The recruitment of the NHEJ machinery to the site of the DSB occurs via a dynamic assembly \(^{137,143}\), rather than a step-wise sequential manner. Simple Ku heterodimer, XRCC4, LigaseIV and XLF are sufficient for the repair of simple DSB, whereas complex breaks requires DNA-PKcs and ATM activation.
Fig: 1.3: Schematic overview of NHEJ: 1) Recognition of the DSB produced by an heterodimer protein Ku 2) Recruitment of the DNA PKcs by Ku and formation of DNA PK complex by the activation of its kinase activity 3) Autophosphorylation of DNA PKcs shifts the complex inwards exposing the DNA ends for the processing enzymes. 4) XRCC4/LigIV complex in combination with XLF forms a filament around the broken ends and once the ends are processed DNA LigIV ligates the ends restoring the DNA integrity.
NHEJ Proteins:

1.10.1 Ku Protein:

Ku is the first of the core NHEJ proteins to sense the DSB and initiate the repair\textsuperscript{144}. The protein occurs in abundance and binds to the DNA ends with a strong binding constant of $2 \times 10^9 \text{M}^{-1}$. Ku protein exists as a heterodimer, composed of a 73 kDa subunit (Ku70) and an 86 kDa subunit (Ku80). Ku70 and Ku80 share similarity in the central DNA binding region (Ku core), with unique carboxy and amino-terminal regions. Ku heterodimer binds to the ends of the DNA in an ATP dependent manner with high affinity\textsuperscript{109,124}. The three-dimensional structures of both subunits, have regions that contribute to a basket structure to completely encircle the dsDNA molecule\textsuperscript{123}. Binding site for Ku on DNA is the sugar back bone and thus sequence independent\textsuperscript{123}. Though there can be binding of multiple Ku proteins on to the DNA in a length dependant manner in-vitro\textsuperscript{146,147}, only a single heterodimer can bind to each end in vivo\textsuperscript{148}. At the two DNA ends of the break Ku binds to each of them forming a Ku: DNA complex. This serves as a dock for further NHEJ protein assembly\textsuperscript{149}. The Ku heterodimer, either directly or indirectly, recruits the main NHEJ factors, including DNA-PKcs\textsuperscript{145}, X-ray cross complementing protein 4 (XRCC4)\textsuperscript{150,151}, DNA Ligase IV\textsuperscript{150}, XRCC4-like factor (XLF)\textsuperscript{152}, and Aprataxin-and-PNK-like factor (APLF)\textsuperscript{139, 140, 153} to DSBs. Ku undergoes a structural conformation in order to interact with DNAPKcs\textsuperscript{154}, polymerases mu and lambda and XRCC4:DNA ligase IV\textsuperscript{151, 154}. On recruitment, binding and activation of DNA-PKcs by the Ku-DNA complex they form “DNA-PK” complex\textsuperscript{156}. There is activation of its kinase activity which phosphorylates four sites in the unique regions on Ku: serine at 6 th position of Ku70 and serine 577, serine 580
and threonine 715 sites of Ku80 \cite{157} and various other NHEJ proteins. Protein-protein interactions are promoted by the putative von Willebrand (vWA) domains on Ku 70 and Ku 80. SAP domain (SAF-A/B, Acinus and PIAS), a putative DNA-binding motif on the carboxy-termini of Ku70 proteins is important in chromosomal organization. Ku 70 interacts with XRCC4 directly\cite{144} and the heterodimer with DNA ligase IV in a DNA dependent manner via its tandem BRCA1 C-terminal (BRCT) domains found in C-terminus of DNA Ligase\cite{150}. The binding of XLF to the Ku heterodimer is also in a DNA dependant fashion mediated between the heterodimeric domain of Ku and the C-terminal region of XLF from amino acids 270–299 \cite{158}. The protein APLF which stabilizes the NHEJ complex, interacts at its conserved MID domain (182–191), directly with the vWA domain of Ku80 \cite{192}. Structurally, the Ku heterodimer performs a general role in binding to and maintaining the stability of the ends at DSBs in all cell cycle phases \cite{159}. By maintaining the two ends of the broken DNA molecule together in a synaptic complex, it produces positional stability in both in vitro and in vivo \cite{146,160,161}. Ku protein also blocks the DNA end processing enzymes including exonuclease 1 and the Mre11/Rad50/Nbs1 complex in vitro \cite{162}. Recently discovered enzymatic activity of Ku is 3’-dRP/AP lyase that removes abasic sites by nicking DNA 3’ of the abasic site, catalyzed via several lysine residues in the Ku70 vWA domain\cite{124}.

In a treatment with combination of camptothecin and ataxia-telangiectasia mutated (ATM) protein inhibitors \cite{148} or treatment of HR-deficient cells with Poly (ADP-ribose) polymerase (PARP) inhibitors \cite{165} or in treatment of Fanconi Anemia (FA)-deficient cells \cite{164}, Ku mediated NHEJ is required to kill the cells.
Fig 1.4: Schematic representation Ku heterodimer subunits 70 and 80 with the common “Ku core” regions of Ku70 (amino acids 261–505) and Ku80 (amino acids 251–509), vWA at the Ku70 (amino acids 37–260) and Ku80 (amino acids 9–235), phosphorylation sites Ku70 (serine-6) and Ku80 (serines 577 and 580 and threonine 715) indicated by “P” and DNA-PKcs interacting region (amino acids 720–732) on Ku80, SAP domain in Ku70 (amino acids 559–609) and a Nuclear Localization Signal domain. S.P. Lees-Miller, K. Meek 2003
1.10.2 DNA-PKcs

DNA-PKcs is a serine/threonine protein kinase, a member of the phosphatidylinositol-3 kinase related protein kinases (PIKKs)\textsuperscript{166-168}. The DNA-PKcs tertiary structure, as determined by the Cryo-electron micrograph studies is composed of a globular head or crown region and a tail or base \textsuperscript{169-171}, with cavities large enough to accommodate dsDNA. It is composed of a large catalytic subunit, DNA-PKcs (encoded by the gene PRKDC), and a heterodimeric protein, Ku. Like the other PIKKs it has a carboxyl-terminal PI 3, 4-kinase domain. Flanking the kinase domain are a large amino-terminal domain and FAT (FRAP, ATM, TRRAP) and FATC domains \textsuperscript{172}. The FAT and FATC domains may interact to stabilize the catalytic domain. The large N terminal domains have a highly helical surface that could provide multiple sites for protein–protein interactions \textsuperscript{173}.

At the DSB, Ku70/80 heterodimer recruits DNA-PKcs to the DNA ends to form the DNA-PK holoenzyme and activates the kinase activity of DNA-PKcs. Activated DNA PKcs can phosphorylate each of the canonical NHEJ factors including Ku70/80 \textsuperscript{157,175}, XRCC4 \textsuperscript{176,177}, DNA Ligase IV \textsuperscript{178}, and XLF \textsuperscript{179}, in vitro, though not essential for their function. Other factors such as Artemis, polynucleotide kinase/phosphatase (PNKP), the histone H2AX and p53 are also phosphorylated in a redundant manner by both DNA PKcs and ATM\textsuperscript{180-182}. DNA-PKcs has no to limited kinase activity in the absence of Ku70/80 and DNA, thus making it truly a DNA-dependent protein kinase \textsuperscript{183,184}. Both autophosphorylation and transphosphorylation of the DNAPKcs can occur as a function of the DNA damage response\textsuperscript{185,186}. Autophosphorylation at the DNA-PKcs S2056 cluster of DNA-PKcs is required for relieving the blockade\textsuperscript{186} it creates for the formation of XRRC4-XLF filament across the DNA break\textsuperscript{185,187,188}, thus allowing the end ligation to occur. Transphosphorylation at the T2609 clusters on DNAPKcs by ATM
primarily required the recruitment of Artemis endonuclease to open the hair pin bends at the DNA ends. ATR mediated phosphorylation also promotes end processing. DNAPKcs can phosphorylate the other self-molecule at the same cluster as ATM. Studies show that the protein with mutation at three of the five potential phosphorylation sites with in the T2609 cluster (3A allele) does not block hairpin opening, but renders it hypersensitive to both ATM and DNA-PK kinase inhibitors[189]. The DNAPKcs null mice are viable but immunodeficient, and human cell lines such as M059J with frame shift mutations of DNAPKcs[190,191] lack kinase activity[192], rendering them radiosensitive[193].
Fig 1.5 Schematic representation of DNA PKcs showing the FAT (FRAP, ATM, TRRAP), FATC domains, kinase domain and the phosphorylation sites. Davis et al. 2014.
1.10.3 XRCC4/DNA Ligase IV:

After processing of the DSB ends by nucleases and polymerases, the terminal and the central joining step of NHEJ is carried out by XRCC4-DNA Ligase IV complex. XRCC4 – X-Ray Cross Complementing protein -4 interacts and forms a tight complex with DNA ligase IV [118,194,195] and this complex, often referred to as X4L4 [224]. This complex is formed in the cytoplasm and localized to the nucleus in response to DSB [243]. During this localization process XLF is also recruited [243]. XRCC4 protein is a tetramer resembling a dumbbell with two globular ends separated by a long thin stalk. The tetramer is composed of two dimers, and each dimer is made up of two similar subunits. The first subunit (L) contains amino acid residues 1 – 203 and has a longer stalk than the second subunit (S) which contains residues 1 – 178. These two subunits share similarity in the globular N-terminal domains. XRCC4 can interacts with DNA [196], DNA-PKcs [197], Ku [157] and DNA polymerase μ [198]. Upon recruitment of XRCC4 and XLF to the site of DSB, these proteins interact co-operatively to form nucleoprotein super helical filaments via alternating XRCC4 and XLF head domains. The interface of the filament interacts and bridges the ends of the broken DNA [199-201]. Hyper-phosphorylation by DNA-PKcs at the C-terminal alpha helical domains of XRCC4 facilitates this interaction. The so formed filament interacts with the DNA at the positively charged surfaces formed by conserved lysine/arginine clusters, XRCC4: Arg-71, Lys-72, Lys-99, Lys-102, and Arg-107; XLF: Lys-31, Arg-57, Lys-59, Lys-63, Arg-64, Arg-81, Arg-107, and Arg-109) [202,203]. XRCC4/LigaseIV complex is formed by the interaction of BRCT2 domain of Lig IV at the C-terminal stalk of XRCC4 via multiple residues, enhanced by the Cernunnos(XLF) [199]. DNA Ligase IV is specific to c-NHEJ and is an ATP-dependent DNA ligase (ATP-dependent transfer of phosphate bonds that results in strand ligation) [204]. Ligase IV at its N-terminus, carries a DNA binding
domain (DBD) with an adenylation domain that has the active site of the enzyme and an oligo-binding domain (OBD)\textsuperscript{204}. At the C-terminus, it has two breast and ovarian cancer susceptibility protein C-terminus (BRCT) domains separated by a linker containing the XRCC4-interacting region (XIR). DNA Ligase IV is capable of ligating incompatible DNA ends and DNA across gaps\textsuperscript{205}. LigIV is recruited to DSBs through its interaction with the stalk of XRCC4\textsuperscript{206,207}. Ligase IV protein is not stable in XRCC4-deficient cells, either in mammals or yeast\textsuperscript{208,209}. Though DNA ligase IV has activity on its own, its interaction with the XRCC4 stimulates Ligase IV activity by the stabilization and adenylation\textsuperscript{194}. Ligase IV mediated ligation of the mismatched and non-cohesive ends is facilitated by XLF. By promoting readenylation, XLF also primes DNA Ligase for the next ligation event\textsuperscript{210-213}. APLF protein was recently proved to stimulate ligation by XRCC4-DNA LigaseIV only in the presence of Ku70/80\textsuperscript{139}. Studies show that localization of XRCC4 and XLF to the DNA and the formation of filaments depends Ligase IV\textsuperscript{214,215}. And formation of Ligase IV/XRCC4 complex is needed for the efficient interaction of XLF with XRCC4\textsuperscript{215}. In addition, Ligase IV/XRCC4 complex contributes to DNA-PKcs autophosphorylation, DNA end synapsis\textsuperscript{216} and recruiting and modulating the activity of nucleases and polymerases thus aiding in processing of the ends\textsuperscript{204, 217-219}. As with the other NHEJ proteins, mutations of the XRCC4/Ligase IV result in immunodeficiency and increased IR sensitivity due to the dysregulated NHEJ\textsuperscript{220}. 
Fig 1.6: Schematic representation of XRCC4 showing core functional and homodimerization domains of XRCC4 as well as the ligase IV binding region. Adapted from Kara et al. 2005

Fig 1.7: Schematic representation of LigasIV showing the DBD, catalytic domain, NLS, BRCT I, XIR and BRCT II Adapted from D.B. Francis et al. 2014
1.10.4 XLF:

XRCC4-like factor (XLF), also known as Non-homologous end-joining factor 1 (NHEJ1) and Cernunnos, is a novel protein encoded by the NHEJ1 gene\(^\text{[221]}\). It is named as XLF due its similarity of the tertiary structure to XRCC-4 rather than the sequence homology\(^\text{[222]}\). XLF is a constitutive dimer with a N-terminal globular head, an alpha-helical stalk and a compact folded C-terminal region\(^\text{[202]}\). XLF has no enzymatic function. It interacts with DNA, XRCC4, Ku and Ligase IV. XLF interacts with DNA in a length dependant and protein concentration dependent manner\(^\text{[119,202,211]}\) at Lys-31, Arg-57, Lys-59, Lys-63, Arg-64, Arg-81, Arg-107, and Arg-109 along with the XRCC4 residues Arg-71, Lys-72, Lys-99, Lys-102, and Arg-107 in its filamentous forms\(^\text{[202,203]}\). XLF and XRCC4 interact in a head-head fashion at residues Leu-115 residues of XLF and Met-59, Met-61, Lys-65, Lys-99, Leu-108, and Phe-10 of XRCC4 forming a “Leu-lock”\(^\text{[202]}\). These two proteins form long super-helical filaments that interact in parallel to form a grooved U-shaped channel\(^\text{[200]}\). XRCC4/XLF head domains form the base and the stalks of the proteins form the sides\(^\text{[223]}\). DNA-PK complex phosphorylates XLF at C-terminal sites in vitro aiding the ability of the XRCC4:XLF filaments to bridge DNA molecules in C-NHEJ or V(D)J recombination\(^\text{[224]}\). Studies show that XLF requires a DNA of 83 bp or longer to bind and cannot form a stable complex with 60 bp DNA, in gel shift assays\(^\text{[211]}\). But in the presence of Ku, XLF can form a stable complex with shorter DNA\(^\text{[152]}\). The 10 terminal residues in XLF(289–299) interact with the central DNA binding domain of the Ku heterodimer\(^\text{[152]}\). XLF once recruited and bound to XRCC4 enhances the activity of DNA ligase IV by 20–200-fold (in vitro)\(^\text{[119,211, 212]}\), especially at mis-matched DNA ends\(^\text{[264]}\). XLF may enhance ligation by promoting re-adenylation of LigIV\(^\text{[225]}\). In response to the damage in the DNA, XLF responds quickly and starts to accumulate locally within a few seconds\(^\text{[137]}\) in a Ku-dependent and
XRCC4-independent reaction. IR sensitivity was detected in XLF deficient human fibroblasts \cite{210,213} and mouse ES cells \cite{226} along with DSB repair defects and severely impaired V(D)J recombination. The first clinical case of SCID due to XLF deficiency was reported in 2003\cite{227}.
Fig 1.8: Structure of the human Cernunnos/XLF protein. (A) Schematic representation of the full-length Cernunnos/XLF protein (299 amino acids). Clinically and functionally important residues are shown. NLS means a nuclear localization signal. A gray region represents the N-terminal globular domain. Adapted from Yoni et al. 2009
1.11 Base Excision repair pathway

Oxidative DNA damage can result from the attack by reactive oxygen and nitrogen species (ROS/RNS) which can be produced by either exogenous sources of oxidation such as ionizing radiations like X, γ and cosmic rays, radon decay, oxidizing chemicals and UV-A solar light or endogenous O$_2$ metabolism (electron transport chain (ETC)), immune responses and inflammation. This includes base modifications such as oxidized bases, abasic (AP) sites, and single-strand breaks (SSBs). BER is the primary nuclear and mitochondrial DNA repair pathway for small base modifications such as alkylation, deamination and oxidation. It predominantly processes small base lesions derived from oxidation, alkylating agents and genotoxic chemicals. The BER pathway functions in five steps utilizing four main enzymes, DNA glycosylase (DG), AP endonuclease (APE1), DNA polymerase and DNA ligases. Other essential proteins that are required for the repair are X-ray repair cross-complementing 1 (XRCC1); proliferating cell nuclear antigen (PCNA); replication protein A (RPA) (Mitra et al., 2001) and poly ADP ribose polymerase-1 (PARP1). The poly ADP ribose polymerase-1 (PARP1) and X-ray repair cross-complementing protein 1 (XRCC1) are the two major scaffolding proteins in BER. The repair pathway proceeds through five steps. First step is the recognition and removal of the damaged base, which is performed as a function of glycosylases such as uracil/thymine excision enzymes (UNG, SMUG1, TDG, and MBD4), 8-oxo-G repair enzymes (OGG1 and MYH), oxidized pyrimidine repair enzymes (NTH1 and NEIL1–3) and methyl-purine glycosylase (MPG). The glycosylation leaves an abasic site which can stall transcription and must be processed for the repair to continue. This apurinic/apyrimidinic site is cleaved by the DG itself, if it is bifunctional or by APE1, if DG is monofunctional. After the removal of the abasic sugar, the strand
break formed is approximated by the scaffolding proteins poly ADP ribose polymerase-1 (PARP1) and X-ray repair cross-complementing protein 1 (XRCC1) \[234-238\]. There is further recruitment of the other BER proteins, required for the completion of the repair \[234-238\].

The fourth step is clearing of the 3' block formed by 3' phosphoglycolates, by APE1 and/or PNKP or Pol β and addition of 5' phosphate groups by 5’ polynucleotide kinase (PNK) in an XRCC-1 and DNA-PKCS-dependent manner\[253\]. As the ends are cleaned and made compatible for ligation, DNA polymerase-β (pol-β) fills in the gaps\[254\]. The insertion can be either a single base (short-patch repair) or a stretch of bases inserted at and beyond the initial site of damage (long-patch repair) displacing some of the nearby bases\[255\]. These displaced stretch of DNA bases are cleaved by a specialized flap-endonuclease, FEN1 \[256\]. Finally, the nick created is sealed by DNA Ligase, DNA ligase I (LIGI, long-patch BER) or DNA ligase III (LIGIII, short-patch BER) \[257\].
Fig. 1.9: Oxidative DNA base damage repair through BER-pathway. Initiation of BER is by removal of the modified base by either a monofunctional (M) or bifunctional (B) DNA glycosylase (DG), which leaves an AP site. If monofunctional enzyme cleaves the glycosidic bond APE 1 incises the DNA backbone 5′ to the AP site. The single strand breaks (SSBs) left by the both DG and MG contain either a 3′ or 5′ obstructive termini. These end breaks are recognized by PARP and processed by pol-β, APE1 or PNKP depending on the specific nature of the terminus. When 3′-OH and 5′-P termini result from end processing, the BER and SSBR diverge into two sub-pathways, short-patch (SN) and long-patch (LP) BER/SSBR. In SN-BER/SSBR, the single nucleotide gap is filled in by pol β aided by the XRCC1 scaffold in the presence of PNKP and ligated by Lig III. In LP-BER/SSBR, pol δ/ε with the aid of PCNA 2 to 13 nucleotide gap is filled in and the resulting 5′ flap is removed by FEN-1. As with the SN-BER final ligation step is accomplished by Lig I.
1.12 Human EndoIII like protein (NTH) and bacterial Endonuclease III

EndoIII like protein (NTH) is the major human DNA glycosylase that detects and removes the oxidized bases by its N-glycosylase activity \(^{(258)}\) as the first step in base excision repair pathway (BER). Thymine glycol is one of the substrates \(^{(259)}\). However, some studies documented that if the oxidative bases are located at the 3’ terminal position of a SSB, that activity of NTH1 is greatly reduced \(^{(260,261)}\). Many genes encode enzymes that are involved in the repair of damaged bases that are evolutionally conserved in many species ranging from bacteria to human \(^{(262)}\).

Endonuclease III (Nth) is a bifunctional enzyme from E. coli that acts as both an N-glycosylase and an AP-lyase. While the N-glycosylase activity releases damaged pyrimidines from double-stranded DNA, generating an abasic site (an apyrimidinic or AP site), AP-lyase activity of the enzyme cleaves 3’ to the AP site, leaving a 5’ phosphate and a 3’ ring-opened sugar. Urea, thymine glycol and uracil glycol are the damaged bases recognized and removed by Endonuclease III \(^{(263,264)}\).

1.13 Klenow Fragment:

DNA polymerase I (Pol I) is encoded by PolA gene (928bp) has 5’-3’ exonuclease activity (1–323), 5’-3’ polymerase activity (324–517) and 3’ -5’ exonuclease activity (521–928). Klenow fragment is the large protein fragment (324-928) and is produced from DNA Pol I by proteolysis with protease subtilisin, losing its 5’-3’ exonuclease activity and retaining the polymerase and 3’-5’ exonuclease activity. It functions in synthesis of double-stranded DNA from single-stranded templates, fills in recessed 3’ ends of DNA fragments to make 5’ overhang blunt by digests away protruding 3’ overhangs. Mutation in the region with 5’-3’ exonuclease activity retains only 5’-3’ polymerase activity, producing a Klenow Exo- fragment. Klenow Exo- can only polymerize
the strand in 5′-3′ direction on a single strand of DNA or at the end with a 5′ overhang. It has also the ability to add an extra nucleotide to the 3′ terminus of a blunt end because of the absence of exonuclease activity.

Studies show that presence of thymine glycol in the clustered DNA lesions influences the repair of these lesions. If Tg is opposed to the AP lesion in a bistranded cluster damage there is a substantial delay in the repair of the AP lesion by short BER\textsuperscript{[265]} Presence of Tg at the DSB with the overhangs, inhibit the gap filling by polymerases, due to its planar structure\textsuperscript{[266]}. The current study focuses on the tolerance of the presence of the Tg at the DSB blunt ends by NHEJ repair in cell extracts and core recombinant NHEJ proteins.
II. MATERIALS AND METHODS

2.1 Substrate

The oligonucleotides 5′ATCGGATCGCTTGTCTg (Tg1), 5′ATCGGATCGCTTGTCTgC (Tg2), 5′ATCGGATCGCTTGTCTgC-Tg-CT-3′ (Tg3) and 5′-ATCGGATCGCTTGT-GT-CT-3′ (Tg5) were ordered from Midland Certified Reagents (Midland, TX). They were re-suspended in TE (10 mM Tris–HCl, pH 8, and 0.1 mM EDTA). 1mCi of [γ-32P] ATP was used to label 50 pmoles of the oligonucleotide, using Polynucleotide kinase (PNK) and incubated for 1 hr at 37 °C. Each oligomer was then annealed to an equal quantity of a complementary 5′-phosphate containing, 3′GTATACGCTAGCGCAACGAP5′ or 3′GTATACGCTAGCGCAACGAP-5′, via heating to 80°C followed by slow cooling to 10°C. The duplexes generated by this annealing process included, on one end a 3-base 3′ overhang that was complementary to the 3′ overhang created by BstAPI digestion of pUC19.

2.2 Ligation of oligomeric duplexes to the KasI/BstAPI cleaved pUC19:

17μg of the plasmid pUC19 DNA is digested with 50 units of KasI (NEB) in 0.3 ml of NEB CutSmart buffer for 3 hrs at 37°C and then digested with 50 units of BstAPI in the same buffer
for 3 hrs at 60°C. The 2.6-kb KasI-BstAPI fragment was purified on a 0.8% agarose gel, electroeluted and concentrated with an ultracel 100k centrifugal filter.

10 pmole of the annealed oligomers were combined with (~1.7 pmole) of KasI/BstAPI fragment and treated with 12,000 units (4 μl) T7 DNA ligase in to ligate them at the BstAPI end in 130μl of ligation buffer for 2 hrs at 25°C. Phenol/chloroform and ethanol precipitated extraction of the ligation products is done. These products are then digested with SmaI (500 units) for 2 hours at 25°C in 50 μl CutSmart buffer. This removes 177 bp from the KasI-end and leaves a blunt end. The resulting 2.6-kb substrate has one Tg-modified end. This substrate is electro eluted and concentrated as above. T7 ligase was used for ligase instead of T4 ligase due to its preferential ligation of only cohesive and not the blunt ends (Doherty at al. 1996; Subramanya et al.1996). Some self-ligation was detected in case of the unmodified control oligomers, and therefore for these oligomers the T7 Ligase was reduced to only 2400 units. In the case of Tg5 plasmid there some ligation seen between head-head blunt ends, and cutting the plasmid with Sma I would have eliminated any head-tail ligation. The ligation products were purified by agarose gel electrophoresis and Electro elution (Bennett et al. 1996).

2.3 Electro elution of the modified substrate:

The nucleotide is electro eluted using a 11 inch segment of pretreated dialysis tubing (6,000-8,000 molecular weight) filled with the buffer composed of 20mM Tris, pH8 and 1mM EDTA (the elution buffer). The gel slice containing the desired DNA band was placed in the dialysis tubing and the other end of it was knotted after removal of air bubbles. The dialysis tube was placed in a gel box filled with the same buffer and the DNA was eluted overnight at 50 V. Then the dialysis bag was cut and the buffer containing DNA substrate was collected into 15 ml
centrifuge tubes. 0.45μm filters were used to remove any debris and the DNA was concentrated by micro-concentration to about 500 μl using centricon-100 (Amicon). The concentrate was collected into 1.5 ml Eppendorf tubes and precipitated with 1/9 volumes of NaOAc and 2.5 volumes of 100% EtOH, washed with 70% EtOH and dissolved in 50 μl TE.

2.4 Bus cell extract preparation:

Bus cells were grown in to confluence on thirty 15-cm dishes each with 20 ml alpha-MEM medium with 10% FBS. On the day of the experiment, the dishes were washed with 20 ml PBS at room temperature. 3ml of 0.25% trypsin-EDTA was added to each dish, making sure it covers all the cells. They were incubated for 5 min until the cells are detached and then 3 ml medium and serum were added. These samples were collected and pelleted in 50 ml tubes at 1500 rpm for 5 min in ultracentrifuge using SW55 rotor at room temperature. The cells were resuspended in 50 ml medium and serum and pelleted again. The pellet was again resuspended in 25 ml medium and serum and pelleted again. The pellet was resuspended in 30 ml ice cold PBS and pelleted at 4°C and packed cell volume (PCV) is recorded. The last process was repeated again. The pelleted cells were resuspended in 5XPCV cold hypotonic buffer containing 10Mm Tris-HCl pH8, 1mM EDTA, 5mM DTT and proteinase inhibitors pepstatin 1μg/ml, chymostatin 1μg/ml, aprotinin 2μg/ml, leupeptin 1μg/ml, PMSF 1mM. The cells were quickly pelleted. This pellet was again resuspended in 2XPCV hypotonic buffer and inhibitors and let set on ice 20 min. 1.5 ml of this solution was put in a small dounce homogenizer each time and stroked 20 times. After it was let set for 20 min, it was transferred to cold 5 ml centrifuge and ¼ volume of the hypertonic solution ( 83.5Mm Tris Ph 7.5, 1.65M KCl, 3.3 mM EDTA, 1mM DTT) and inverted several times to mix well. This was centrifuged at 43K for 3 hrs in SW55 at 2°C. From the top the middle layer which has cytosol was collected using 22G needle, avoiding cloudy layer on the top. The
cytoplasm thus collected was dialyzed for 3 hrs in storage buffer (20Mm Tris pH 8, 0.1 M KOAC, 1Mm DTT, 0.5Mm EDTA, 20% glycerol) with inhibitors at 4°C. The resulting solution was freezed in small aliquots(10μl-80μl) at -80°C.

2.5 End joining assays using cell extracts:

Bus whole cell extracts which are deficient in XLF were used for the reactions in a buffer containing 50 mM triethanolammonium acetate (pH8), 1 mM ATP, 1 mM dithiothreitol, 50 μg/ml BSA , 1.3 mM Mg(OAc)2 and dNTPs at 100 μM each.

Total reaction volume was 16-μl, with 10μl of extract, resulting in a final protein concentration of 8 mg/ml. 200nM of human recombinant protein XLF was substituted for some reactions. 100 ng of the substrate was added and the reaction mixed by pipetting and incubated at 37 °C for 6 hrs. Samples were deproteinized by adding an End Joining Lysis buffer containing 1mg/ml of Proteinase K for 3 hrs at 56°C. The DNA was then extracted with phenol/chloroform and precipitated with glycoblué and ethanol for one hour. After the precipitation, all samples were dissolved in 43 μl of TE and digested with 20 units of NdeI and PstI and 5 μl of 10X CutSmart buffer (New England Biolabs) at 37°C for 3 hrs. Gel electrophoresis of the samples was done in 20% polyacrylamide sequencing gels. Then the gels were exposed to phosphor imaging screen for1 to 2 days at -20°C. Data were analyzed using a Typhoon 9100 imager (GE Healthcare Bio-Sciences, Pittsburgh) and Image Quant 3.1 or 5.1software (GE Healthcare Bio-Sciences, Pittsburgh). The percentage of end joining was determined by measuring the percentage radioactivity of the joined fragments as a function of total radioactivity in the same lane.
2.6 Polyacrylamide gel electrophoresis:

For electrophoretic separation of the DNA strands, Polyacrylamide gels (20x30x0.08cm) containing 40% acrylamide: bisacrylamide in a ratio of 20:1, 8μl of 10XTBE and urea in a final concentration of 8.3 M were used. The mixture was cooled to room temperature and 0.06g ammonium persulfate and 45μl of TEMED (N’, N’, N’, N’-tetramethylethylene diamine) was added. The gel was allowed to polymerize for 1 hour. 15 μl of samples are then loaded into the wells of the gel and electrophoresed at constant power of 42 W until the marker reached the bottom of the gel.

2.7 End joining reaction with ddTTP and Endonuclease:

Bus cell extracts were substituted with XLF and ddTTP was used instead of dTTP during the incubation for some end joining reactions. After proteolysis and phenol/ethanol precipitation, each of the reactions with dTTP and ddTTP was divided into equal halves. One half was treated with Endonuclease III in Endonuclease III buffer for 2 hrs at 37°C, while the other half served as a control. After inactivation of endonuclease at 65°C for 20 min, the reactions were treated with NdeI and PstI and analyzed on the 20% polyacrylamide sequencing gels the same way as above.

2.8 Klenow Enzymes treatment:

The joined plasmids in the Bus cell extracts substituted with XLF were treated with Klenow and Klenow exonuclease with the addition of 1 mM ATP, dNTPs at 100 μM and 5μl CutSmart buffer for 30 min at 37°C, after the proteolysis and phenol/ethanol precipitation. After inactivation at 75°C for 20 min and phenol/ethanol precipitation, the reaction is digested with NdeI and PstI, and then analyzed on 20% polyacrylamide sequencing gels.
2.9 End joining by the recombinant proteins:

Ligation of the above prepared substrates was attempted using recombinant NHEJ proteins in a buffer containing 50mM Tris pH8, 200m, 100μg/ml BSA, 0.2Mm EDTA, 0.1% Triton, 4mM DTT and 10% Poly ethylene glycol), substituted with 2mM ATP and 100mM MgCl₂. Proteins Ku, XRCC4 and XLF were dissolved in 1X of the above buffer. The total reaction volume was 10μl, with the protein concentrations of 10nM, 40nM and 100nM of Ku, XRCC4 and XLF respectively and 10ng of the labelled substrates. A third of the samples had only DNA and no proteins, a third of them had DNA with Ku and XRCC4 while the rest had all the three proteins and DNA. The reactions were incubated for 4 hours at 37°C. Samples were deproteinized by adding an End Joining Lysis buffer containing 0.3 M NaCl, 10Mm EDTA, 20 mM Tris pH 7.6, 1% SDS and 1mg/ml of Proteinase K for 3 hrs at 56°C. DNA was then extracted with phenol/chloroform and precipitated with ethanol for one hour. After the precipitation, all samples were dissolved in 43 μl of TE and digested with 20 units of NdeI and PstI after addition of 5 μl of 10X CutSmart buffer (New England Biolabs) at 37°C for 3 hrs. Gel electrophoresis of the samples was done in 20% polyacrylamide sequencing gels. Then the gels were exposed to phosphor imaging screen for1 to 2 days at -20°C. Data were analyzed using a Typhoon 9100 imager (GE Healthcare Bio-Sciences, Pittsburgh) and Image Quant 3.1 or 5.1software (GE Healthcare Bio-Sciences, Pittsburgh).

2.10 Statistics:

Error bars represent standard error of mean (SEM) for at least three independent experiments. Unpaired two-tailed t-tests were performed and the data were reported as significant for P values <0.05.
**Fig 2.1: Preparation of the substrate:** The oligonucleotide with Tg was radiolabeled with $^{32}$P at 5′ end is annealed with a complementary strand. This was ligated to the BstAPI cut end of the pUC19 using T7 DNA Ligase. The substrate thus prepared had a head with BstAPI and NdeI restriction sites and a tail with SmaI and PstI restriction sites. This substrate was used in both cell extract ligation assay and assay with recombinant proteins.
Fig 2.2: Assay Protocol: Radioactively labeled oligomeric duplexes containing normal thymine or Tg at different positions were ligated with BstAPI digested pUC19. Following attempted ligation with T4 DNA ligase or with cell extract containing DNA Ligase IV, or purified proteins, plasmids were digested with either NdeI and PstI. Resulting fragment sizes on autoradiograph revealed the extent of ligation of the various substrates under the given conditions. Inter/intramolecular ligation between the oligomeric duplex end and the end cut with SmaI, generated a 44-nucleotide. Intermolecular ligation of two blunt ends of oligomeric duplex, generated a 36-nucleotide fragment upon digestion with NdeI in both sides. If there was no ligation there will be a 19-nucleotide fragment due to digestion by NdeI.
Fig 2.3: **Base excision repair of the substrates:** A) When base excision repair proteins recognize and remove the Tg from the third position in Tg3, instead of the 19-nucleotide fragment 16-nucleotide fragment was formed and there is no end joining. B) When the Tg from the Tg5 substrate is removed there was formation of 14-nucleotide fragment and there was no end joining.
III RESULTS

3.1 T4 DNA ligase assay to test interference with the end joining by thymine glycol

Thymine glycol at the DSB of DNA interferes with the efficiency of end joining, independent of the NHEJ process. T4 DNA ligase is a bacteriophage ligase that is capable of ligating cohesive or blunt termini of the DNA. Therefore, before testing for NHEJ activity, T4 ligase assay was done to determine whether the substrates could be ligated. Linearized substrate plasmids containing $^{32}$P-labeled oligonucleotide duplexes with normal thymine or Tg at different positions from the termini (Tg1, Tg2, Tg3, Tg5) were evaluated for ligation following incubation with or without T4 DNA ligase. The ligation products were digested with NdeI and PstI in order to detect the fragments that were diagnostic of the efficiency of ligation (Figure 4-1). The unjoined substrate is indicated by a 19-base labeled fragment released by NdeI. For Tg-containing substrate, this 19-mer forms a doublet due to different stereoisomers of Tg. The joined substrate would appear as a 44-base band in the NdeI and PstI digestion of the DNA, if the ligation is head-tail in intra or inter molecular ligation. In case of intermolecular ligation head-head, digestion with NdeI (with or without PstI) would yield a 36-bp fragment.

Results of the assay show that T4 Ligase could efficiently ligate the unmodified plasmid head-tail (61%) and head-head (7%). When the substrates with Tg at different positions were used
there was an increasing efficiency of joining with the increasing distance, Tg1 - 1.6%, Tg2-1.9%, Tg3-36%, Tg5-44%. The joining was only in Head-Tail (44) in Tg1 and Tg2 and some head-head (36Tg) joining was observed in Tg3 and Tg5 (1.7% and 5.2% respectively). This shows that T4 Ligase was able to ligate the plasmids with Tg at blunt ends with minimal efficiency in Tg1 and Tg2, but as the distance of the Tg from the end increased there was also an increased ligation as observed in Tg3 and Tg5.
**Fig 3.1: T4 Ligase Assay:** 1) Gel displaying ligation of the plasmids Tg0, Tg1, Tg2, Tg3 and Tg5 with and without T4 Ligase followed by digestion with NdeI and PstI 2) Graph representing the efficiency of ligation of the plasmids in percentage 44-mer indicative of head-tail end joining, 36Tg-mer indicative of head-head ligation with Tg in position. 36-mer indicative of head-head end joining.
3.2 Ability of NHEJ to ligate the substrate with Tg5 in Bus cell extracts in the presence and absence of XLF:

Previous experiments with Bus cell extracts showed significantly greater end joining with Tg3 than Tg1 and Tg2. Therefore, ligation of the substrate with Tg at the fifth position Tg5 was attempted using XLF deficient Bus cell extract in a time dependent fashion to observe the effect of further distancing Tg from the end. The experiment was done with and without XLF addition. The end joining was evaluated after digestion with NdeI and PstI enzymes. The control with inactive extract shows 36Tg band due small amount of the head-head end joining during preparation of the substrate. In the presence of XLF there was an increase in the end joining with time both head-tail (44) and head-head (36). There was head-head joining with Tg still present in the joined product (36Tg) which can be seen as a shifted band above the 36-base fragment, which has the Tg removed and substituted with thymine. The 36-base strand has palindromic bases and hence snaps back on itself to form a hair pin structure. The hair pin bend thus formed accelerates the mobility of the 36-base strands. The shift of 36Tg was because of the interference of the Tg with the snap back of the strand on itself. The formation of 36-base strand was delayed compared to that of the 44-base strand. With time there was a minimal decrease in the joined head- tail (44-base) fragment. In addition to the 19-base fragment which represents non-joining, there was a 14-base fragment, presumably as a result of BER. The appearance of this 14-base fragment was earlier and greater in magnitude than the end joining (44) fragment. In the absence of XLF, there was no detectable end joining head-tail (44-base) but there was 36Tg and early appearance of 36-base band. This was likely due to conversion of the 36Tg which was formed during preparation of substrate, to 36-base strand by BER and it was partially suppressed in the presence of XLF as shown by the delayed appearance. There was an extra band below the
36-base band due to joining of the repaired strands with a few nucleotides removed before end joining. There was an earlier appearance of the 14-bases strand in the absence than in the presence of XLF may be due some inhibition, which needs further evaluation. This data proves that presence of XLF is essential for the end joining reaction.
Fig 3.2: The ability of NHEJ to rejoin DSBs with Tg at the fifth position (Tg5): Time dependent end joining assays in whole-cell extracts of XLF-deficient Bus fibroblasts with and without XLF substitution, with plasmid with Tg at the fifth position from the break. Extracts were supplemented with 200nM XLF.
3.3 Effect of Endonuclease and ddTTP on the end joining:

Previously, studies showed that the plasmids containing proximal Tg were joined with Tg still present in the joined fragments. To test this in the case of Tg5 substrate, Bus Cell Extract assay was done in the presence of either ddTTP or dTTP. They are divided into equal halves, one half was treated with Endonuclease III (other half was used as a control), after the end ligation step but before digesting with the restriction enzymes. ddTTP functions as a DNA chain terminator and hence inhibits the end joining. Endonuclease III recognizes and removes Tg from the DNA. Substitution with ddTTP was found to reduce the intensity of the joined fragments but did not inhibit the joining in total. Treatment with Endonuclease III reduced the end joining, as shown by the reduced intensity of radioactivity of the head-tail 44-base strand. Treatment with both Endonuclease and ddTTP further reduced the intensity of the joined fragment. This experiment proves that Tg was not removed from the strand before ligation in at least some end joining events. In some cases Tg was removed and replaced with thymine, but the experiment does not distinguish whether such replacement occurred before or after end joining.
Fig 3.3: The effect of the ddTTP and Endonuclease III on the joined fragment (44): A) Autoradiograph showing the effect of substitution of ddTTP for dTTP during incubation and Endonuclease III treatment after the ligation step and before the digestion with NdeI and PstI. B) Graph displaying the percentage of 44-base strand on treatment with ddTTP and Endonuclease III.
3.4 Comparison of End Joining and Base Excision repair between Tg3 and Tg5 in cell extracts:

To compare the end joining of the plasmids Tg0, Tg3 and Tg5, all three were incubated in either heat inactivated cell extracts or Bus cell extracts supplemented with XLF. After treating with NdeI and PstI, repair of these substrates showed that Tg0 has 40% total end joining, Tg3 has 14% and Tg5 has 10%. Thus there was more end joining in Tg3. In addition to the 19-base fragment, there was 16-base fragment in the Tg3 repair lane and a 14-bp fragment in the Tg5 lane. This was assumed to be due to the base excision repair of the substrate. The intensity of the 16-bp fragment (48.5%) is greater than 14-bp fragment(7.6%). Thus, there was much more base excision repair with Tg5 compared to Tg3 evident also from the unjoined and unprocessed substrate (19-base fragment) was higher in intensity for Tg3 than Tg5 as shown in the Fig 3.5. The increased distance of Tg from the end had apparently provided access for the BER repair proteins for the excision process and thereby decreased the end joining. Unpaired Two tailed tests performed revealed that Tg0 and Tg5 had a statistically significant difference (p<0.0003) for the end joining, and Tg3 and Tg5 had p<0.02, thus statistically significant.
Fig 3.4: Comparison of the repair between Tg3 and Tg5: A) Autoradiograph showing the repair (joining and BER) fragments of Tg0, Tg3 and Tg5 in Bus cell extracts supplemented with XLF. B) Graph representing the percentage of both joining and BER fragment for Tg3 and Tg5.
3.5 Assessment for fill in and ligation after BER:

After the BER, the DNA strand repaired was expected to be filled in and ligated. However, even after 6 hours of incubation there was no or minimal change in the 14-base fragment. To assess the ability of this strand to be filled in, Klenow enzymes which can add nucleotide to the 3’ end if there is a compatible 3’ OH group and a suitable template strand were used. After the end joining reaction in inactivated extracts and 0.25 or 6 hrs in cell extracts substituted with XLF, the reaction was treated with either Klenow or Klenow exonuclease’ (Klenow fragment which has lost its 3’-5’ exonuclease activity). Klenow fragment was able to add the nucleotides to 14 base fragment producing bands corresponding to fragments between 14 and 19 bases in length. For the 0.25 hrs sample the entire 14-base band disappears to higher levels but at 6 hrs some of it is still present. Klenow exo’ added an extra nucleotide to 19-base fragment and nucleotides to 14-base fragments at 0.25 and 6 hrs. There was a complete disappearance of the 14-base band at 0.25hrs whereas at 6hrs not all of it disappears. This persistence of the 14-base fragment and the productions of base fragments between 14 – 20 bands at 6 hrs could be probably due to the degradation of the complementary stand to various lengths. There was also a shift in the head-tail joined fragment to a smaller oligonucleotide level below 44-base band on treatment with Klenow fragment and shift of the 36Tg and 36-base band to a higher level. There is no change in the 44-base band with Klenow Exo’, but a higher shift in 36Tg and 36-base strands was observed. This data reveals that there was a 3’-OH group at the 14-base fragment end which was not filled in the extracts. The shift observed in 36Tg and 36-base bands might have been due to filling of the 5’ overhangs produced by the NdeI digestion on either side of the DNA, by the 5’-3’ polymerase activity of Klenow and Klenow Exo’. Shift of the 44-base fragment with the Klenow enzyme is difficult to be explained since this band represents an NdeI/PstI fragment.
generated after Klenow removal and there should have been no free 3’ terminus available for modification during treatment with Klenow.
Fig 3.5: Testing with the Klenow and Klenow exonuclease: Autoradiographs displaying the effect of A) Klenow fragment and B) Klenow Exonuclease on the BER products generated in the end joining of the Tg5 plasmids in Bus cell extracts with XLF.
3.6 Effect of Tg position on total, head-head and head-tail end joining:

To assess the total extent of ligation, the ligation of Tg0, Tg3 and Tg5 were analyzed on a 0.8% agarose gel after the end joining step after proteolysis. This was compared to that of the same reactions treated with SmaI after proteolysis. When all the possible end joining of the substrates are considered, both Tg3 and Tg5 have similar extent of ligation. When the restriction enzyme SmaI was used, there is diminished total end joining in case of Tg5. The similar ligation observed for Tg3 and Tg5 was due to joining of the plasmids at the tail ends which are not modified and also in Tg5 the tail ends were available for end joining as they was less ligation to head. SmaI by cleaving this end joining reduced the intensity leaving the joined fragments only in head-tail and head-head. This result confirms the reduced joining of the Tg-containing end of Tg5, as compared to Tg3, due to interference with BER.
Fig 3.6: Effect of Tg position on total, head-head and head-tail end joining: Autoradiograph showing the total end joining of the unmodified (Tg0), Tg3 and Tg5 in the top lane and the result of treatment of Sma I in the below lane.
3.7 Time dependent end joining and base excision repair of Tg3 and Tg5 plasmid:

To analyze the kinetics of repair, cell extract ligation assay was performed with Tg3 and Tg5 at different time endpoints of 0.25, 0.5, 1, 2, 4, 6, 12 hours of incubation. There was an increased head-head and head-tail end joining with time. From the graph Fig 3.7 B, it is evident that BER was more active reaching its peak earlier than the end joining. There was no significant increase in the end joining even after the BER had reached the plateau. This inability of the 14-base fragment to be filled in and ligated has been already addressed above in the result 3.5 using Klenow enzymes.
**Fig 3.7: Time dependent repair kinetics of Tg3 and Tg5:** A) Autoradiograph displaying the intensities of the repair fragments as a function of time. B) Graphs displaying the percentage of the joined and BER fragments Tg3 in the top and Tg5 below as a function of time.
3.8 Recombinant protein assay:

After cell extract assays, end joining assays were performed with unmodified, Tg1, Tg2, Tg3, and Tg5 using recombinant NHEJ proteins- Ku, X4L4 and/or XLF to observe the effect of these core NHEJ proteins on the substrates in the absence of other repair factors. Data showed that the unmodified fragment was ligated both in head-tail (44) and head-head (36) with 9% ligation in the presence of XLF and 2% in the absence of XLF. Tg1 plasmid showed 0.4% ligation head-tail (44) only in the presence of XLF. Tg2 had a slightly higher, 0.6% ligation of the same. Tg3 on the other hand has ligation, 0.1% in the absence of the XLF and 2.3% in the presence of XLF. In the presence of XLF, there was also a head-head ligation 1% (36Tg), with Tg in place. Tg5 also displayed repair in both head - head (0.2%) and head-tail (2%) in the absence of XLF. XLF enhanced the ligation of Tg5 to 4% in head-tail fashion and 2% in head-head fashion. Thus ligation improved with the increased distance of Tg from the end. But the ligation as a whole was 3-5 fold less than that observed in extracts, which would suggest the requirement for other repair factors for the most efficient ligation.
Fig 3.8: Repair by the recombinant NHEJ proteins Ku, X4L4 and/or XLF of Tg0, Tg1, Tg2, Tg3 and Tg5: A) Autoradiograph showing the extent of joining of the plasmids by the proteins B) Graph displaying the percentage of joining by the proteins in the presence and absence of XLF.
IV DISCUSSION

Radiotherapy is a “double edged sword” inducing genetic mutations in the surrounding normal tissue while causing a loss of clonogenic survival of tumor cells \(^{(267)}\). IR has the propensity to induce more than a single lesion in the DNA in one or two helical turns, known as “clustered damage” including the modified bases, SSBs and DSBs. These clustered lesions show reduced reparability when compared with that of individual lesions \(^{(268)}\). And when these lesions are in combination with DSB they were the poor substrates for repair \(^{(43, 72)}\).

Thymine glycol being the most common modified/oxidized base formed during radiotherapy, \(^{(76)}\) is repaired mainly by Base Excision Repair Pathway \(^{(76)}\). NHEJ repair pathway proteins, comprising DNA-PK, XLF, the X4L4 complex, and polymerases belonging to PolX family, have the ability to ligate the termini of DNA breaks that do not require extensive processing, with high efficiency. However, many DSB are associated with complex lesions that can be barriers to the end joining \(^{(266)}\). The current study involves the evaluation of effect of the presence of thymine glycol on the NHEJ repair pathway in cell extracts and by recombinant proteins.
T4 ligase, the bacteriophage ligase has been shown to efficiently ligate the blunt, compatible, non-compatible \(^{(269)}\) and branched DNA ends \(^{(271)}\). Ligation assay with T4 Ligase was done to assess its ability to ligate modified DNA ends as compared to X4L4 which presumably evolved specifically to ligate the modified substrates. Assays revealed that T4 Ligase can ligate the unmodified ends efficiently. The modified substrates joining efficiency increase with the increasing distance of the modified base from the end. This joining was head-tail \(^{(44)}\) both in inter and intra molecular fashion, where only one end is modified, with increasing ligation as the Tg is moved to increasing distance from the end. In contrast, presence of the Tg at both the completely blocked head-head joining \((36\text{Tg bp})\) in Tg1 and Tg2, but allowed some joining in Tg3 and Tg5 with the thymine glycol in place \((36\text{Tg})\). As there were no other processes to remove Tg from the strand, the strands are ligated with Tg in place. The 36Tg band is higher than the 36-base band as there was a palindromic sequence in the strand which formed a hair pin bend in 36-base oligonucleotide. Tg interfered with the formation of this hair pin structure in 36Tg, retarding the mobility. Thus, presence of Tg near the end of the DSB distorted the ends such that T4 Ligase could not ligate the end but when moved just a few bases away from the end there was ligation.

Previous studies in our laboratory with the Bus cell extracts using substrates Tg1, Tg2 and Tg3 revealed that the ends with Tg at the first and the second position of the DSB are less joined head-tail and not joined head-head. But Tg3 showed repair almost equal to unmodified ends both head-tail and some joining head-head. In the current study substrate with Tg at the fifth position \((\text{Tg5})\) is used to observe the trend of the repair with the increased distance from the end in the cell extracts. Bus cell extracts which are XLF deficient were used to detect the ligation of the substrates. Ligation is assessed in the extracts in the presence and absence of XLF in a time
dependent fashion. Two main observations were made in this process. There was an increase in both head-tail and head-head with XLF substitution. But there was base excision repair in both the presence and absence of the XLF. But the BER appears to be inhibited initially in the presence of XLF as evident from the Fig 3.2, there was less and delayed BER in the reaction with XLF. There was a 36Tg oligonucleotide strand formed during the preparation of the substrate due to the some ligation of the blunt ends in head-head fashion. This 36Tg stand could be observed in the control treated with inactive extract. In the absence of XLF this 36Tg strand was converted to 36-base strand early by the BER and not in its presence. This might be because the presence of XLF had inhibitory effect on BER. This effect of XLF on BER should be evaluated further. There was also some ligation of the BER repaired strand and it appears below the 36 base band because of loss of some nucleotides. Base excision repair appeared to be activated almost immediately in the first quarter of the hour, repairing the end forming the 14-base oligonucleotide with the removal of Tg at the fifth position in the 19-base unjoined strand. Though there was a halt in the BER as shown by the intensity of the 14-base strand, there was no significant increase in the end joining of the plasmid.

The plasmids Tg1, Tg2 and Tg3 were shown to be ligated with Tg still in position. The same was tested with Tg at 5th position. Addition of ddTTP which blocks the extension of the strand by polymerases and ligation of the strands, instead of dTTP while incubation had reduced but did not eliminate the intensity of the radioactivity suggesting that Tg had not been totally removed from all strands and at least in some instances they were ligated with the Tg in position, proving that Tg in the fifth position is not a hindrance for the end joining. Treatment with Endonuclease III after the ligation step removed Tg from the ligated strand and reduced the radioactivity of the
joined fragment, which is also consistent with the above result. There was further decrease in the intensity of the joined fragment with ddTTP and Endonuclease III due to the presence of some normal thymidine in the extract.

Modified bases placed at different positions from the break displayed different repair characteristic depending on the type of the lesion. As discussed in the study of the effect of the glycosylases, NTHI, OGGI and NEILI on lesions 5-hydroxyuracil (5-OHU) and 8-oxoguanine (8-oxoG), the repair is variable depending on the type and position of the lesion [272]. Previous studies in the lab revealed that there was no detectable BER prior to end joining for Tg1 or Tg2, some was observed in Tg3. To evaluate the effect of increasing the distance of Tg from the DSB on the repair of the modified substrates in Bus cell extracts, Tg5 repair was studied in comparison with the Tg3 substrates. This experiment revealed that there was a reduced end joining for Tg5. In addition, there was much more extensive activation of the Base excision repair pathway for Tg5 than that observed for Tg3. This observation denotes that the positioning of the Tg at the ends in DSB can result in poor binding and low activity for the BER proteins and increased distance from the end provides a better substrate for the activity of BER. When the time-dependent kinetics for both Tg3 and Tg5 were studied, BER pathway was activated almost immediately compared to the end joining repair. While the BER reached the plateau at 1 hour time point, the end joining did not show a significant increase. This observation led to question of the end processing of the substrate by BER proteins as APE1 should leave a compatible 3′-OH group and either polλ/polµ (NHEJ) or pol β (BER) should have filled in the gap for end joining. There was less increase in the end joining than expected of it with time even in the presence of a robust BER. In order to test for the compatibility of the repaired 3′ end for the addition of nucleotides Klenow enzymes (Klenow fragment and Klenow Exonuclease-) were used. Klenow
enzyme was able to add nucleotides to the 14-bases strand resulted from BER, producing multiple bands between 14 -19- bases. At 0.25 hrs there was complete disappearance of the 14-bases strand to the higher levels while some of it was still in place at 6hrs. Though there should be a conversion of the whole 14-base strand to 19-bases at 0.25 hrs due to the polymerase activity of the Klenow fragment, its exonuclease activity dominated producing bands of various length. At 6hrs, the bands from 14 to 19-base strand indicate that the complementary strand has been digested over time. This addition of the nucleotides suggests that the 3’ end of the 14-base strand has OH group and can be filled in. When Klenow exonuclease- enzyme was used, in addition to filling the nucleotides from 14-base -19-base, it added an extra nucleotide to the 19-base strand. This was because of lack of the exonuclease activity which allowed the polymerase activity to add nucleotides to the 14- base strand and convert it completely in to 19-base or 20-base strand at 0.25 hrs, as the complementary strand is completely intact. As the complementary strand was digested to variable lengths in cell extracts by 6 hrs of incubation, multiple bands were formed due to the polymerase function. An extra nucleotide was added to the 19-base strand as there was no exonuclease activity. This result indicated that the BER product had a compatible 3’ OH for the polymerases in the cell extracts to fill in. The reason why extracts could not do it, needs to be evaluated further. The shift in the 36Tg and 36-bases fragment with both the enzymes might have been due to the filling of the ends with 5 ‘overhangs created by the NdeI digestion in 3’-5’ direction by the polymerase activity of the Klenow and Klenow Exo- enzyme.

When the total end joining efficiency of the Tg5 was compared with that of Tg3 in the extract using agarose gels, there was an equal amount of total end joining in Tg3 and Tg5. This was due to the tail- tail ligations, which do not have Tg and hence ligated efficiently. And also as the end-joining head-tail was less with Tg5, there was more availability of tails for the joining.
SmaI had reduced the intensity of the joined fragments in Tg5 by cleaving these tail-tail ligations, leaving the head-head and head-tail ligations intact. Thus joining of the Tg-containing end was less for Tg5 than Tg3, due to interference by BER in Tg5.

End joining assays in the previous studies with the recombinant proteins have shown that compatible ends can be repaired efficiently even in the absence of the XLF protein and there was moderate and strong stimulation of the ligation of the blunt and incompatible ends respectively by XLF\(^{269}\). Based on this, repair of the blunt ends with the Tg was attempted using core NHEJ recombinant proteins Ku, X4/L4 and XLF. Tg1 and Tg2 required the presence of XLF for ligation, while Tg3 and Tg5 displayed some ligation even in the absence of the XLF. The extent of ligation of Tg3 and Tg5 was boosted 5-6 fold when XLF is added. Presence of the Tg at the ends might have resulted in their distortion making it impossible for the Ku and X4L4 to join these ends. XLF along with XRCC4 forms a filament across the DNA ends perhaps stabilizing the alignment of distorted ends. In addition, XLF might have helped to overcome the limitation by stimulating the LigaseIV activity due to its re-adenylation function \(^{210-213}\), allowing the end joining. Previous studies in the lab showed that XLF was an absolute requirement in extracts for the end joining reaction by NHEJ irrespective of the position of the Tg. In contrast, there is some ligation in the absence of XLF in Tg3 and Tg5 in the study of recombinant proteins. This could be due to competition between some other repair proteins and X4L4 in extracts for the DSB or failure of the more distorted ends to stimulate the ligase activity of X4L4. Under these conditions, XLF was required to promote the ligation. These ligated products had Tg, as there was no processing of proximal Tg residues in the ends due to lack of the BER proteins required for the same. The extent of joining with the purified proteins was 3-5 fold less than the joining in
the extracts suggesting that other structural and enzymatic proteins are required to efficiently stimulate the end joining in extracts.

In future, the effect of other proteins of NHEJ such as DNAPKcs, PAXX is to be tested to analyze the effect of these proteins on the end joining. BER proteins, NTH1 and NEIL1 with/without APE1 can be added to the NHEJ proteins to study the effect on the modified plasmids. These modified plasmids can be transfected in to cells to study the repair by other repair proteins and pathways which may not have survived the extraction process. Repair of these substrates can be evaluated after the DNA isolation and RT-PCR with the primers on either side of the DNA modification and also analyzed by high-throughput DNA sequencing.
V. CONCLUSION

The position of the modified base (Tg) alters the end joining of the blunt ends. Tg also affected the activity of the base excision repair. In the cell extracts the Base excision repair pathway was more active and competes with the End joining reaction especially if the base was far from the end shown by the more joining for Tg3 and more Base excision repair for the Tg5. XLF was mandatory for the repair of the plasmids in extracts. Recombinant proteins more efficiently repaired the plasmids when Tg was located further away from the DSB, suggesting possible less distortion of the ends. The absolute requirement for XLF in the repair of plasmids with Tg near the end reiterates its role in stimulating X4L4 at the incompatible ends. XLF is required by X4L4 enzyme to overcome the competition with other proteins in the extract at the DSB. Decreased repair by purified proteins denotes the requirement of other proteins and repair mechanisms for the effective repair of these complex lesions.
LIST OF REFERENCES
**LIST OF REFERENCES**


75


129. Sutherland BM, Bennett PV, Sutherland JC, Laval J. Clustered DNA damages induced by X-rays in human cells, Radiat. Res 157(2002) 611–616.


221. "Entrez Gene: NHEJ1 nonhomologous end-joining factor 1"


