THE INFLUENCE OF THE KU80 CARBOXY-TERMINUS ON ACTIVATION OF THE DNA-DEPENDENT PROTEIN KINASE AND DNA REPAIR IS DEPENDENT ON THE STRUCTURE OF DNA COFACTORS

Derek S. Woods

Submitted to the faculty of the University Graduate School
in partial fulfillment of the requirements
for the degree
Doctor of Philosophy
in the Department of Biochemistry and Molecular Biology,
Indiana University

November 2013

Accepted by the Graduate Faculty, of Indiana University, in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

John J. Turchi, Ph.D., Chairman

Maureen A. Harrington, Ph.D.

Doctoral Committee

Anna L. Malkova, Ph.D.

August 6, 2013

Yuichiro Takagi, Ph.D.

ACKNOWLEDGEMENTS

I would like to thank my family for their unfaltering support over the last several years. My parents have played such a key role in my development and well being throughout my life. I am very lucky to have their unconditional love and dedication. My brother and sister-in-law, who have expressed their support of science and my research in particular over the last 5 years. My affectionate and supportive wife, Carly Woods, whose love and commitment to me knows no end. She sacrifices so much in order to be with me and this does not go unnoticed. I can't imagine my life without her. She is my rock, my biggest cheerleader, and my best friend. There is no way I can ever repay her for all that she has done for me but I intend to try everyday for the rest of my life. I also have to thank my advisor, Dr. John Turchi, for taking a risk by accepting me into his lab with basically no experience to work on a project that was not funded. Over the past 5 years he has given me direction when it was required as well as the freedom to explore my passion. I look forward to our next endeavor at NERx Biosciences, Inc. where he has once again decided to take a chance on me. Finally I would like to thank Dr. Katherine Pawelczak whose previous work is the basis for my thesis. She has taught me numerous valuable lessons in science and in life. I truly appreciate her friendship and mentorship.

ABSTRACT

Derek S. Woods

The Influence of the Ku80 Carboxy-Terminus on Activation of the DNA-Dependent Protein Kinase and DNA Repair is Dependent on the Structure of DNA Cofactors

In mammalian cells DNA double strand breaks (DSBs) are highly variable with respect to sequence and structure all of which are recognized by the DNA-dependent protein kinase (DNA-PK), a critical component for the resolution of these breaks. Previously studies have shown that DNA-PK does not respond the same way to all DSBs but how DNA-PK senses differences in DNA substrate sequence and structure is unknown. Here we explore the enzymatic mechanism by which DNA-PK is activated by various DNA substrates. We provide evidence that recognition of DNA structural variations occur through distinct protein-protein interactions between the carboxy terminal (C-terminal) region of Ku80 and DNA-dependent protein kinase catalytic subunit (DNA-PKcs). Discrimination of terminal DNA sequences, on the other hand, occurs independently of Ku 80 C-terminal interactions and results exclusively from DNA-PKcs interactions with the DNA. We also show that sequence differences in DNA termini can drastically influence DNA repair through altered DNA-PK activation. Our results indicate that even subtle differences in DNA substrates influence DNA-PK activation and ultimately Non-homologous End Joining (NHEJ) efficiency.

John J. Turchi, Ph.D., Chairman

TABLE OF CONTENTS

ABBREVIATIONS

DSB	double-strand breaks
bp	base pairs
py	pyrimidines
pu	purines
NHEJ	non-homologous end joining
HR	homologous recombination
DDR	DNA damage response
DNA-PK	DNA dependent protein kinase
DNA-PKcs	DNA dependent protein kinase catalytic subunit
PIKKs	phosphatidylinositol-3 kinase-like protein kinases
FAT	FRAP, ATM, and TRRAP
C-terminus	Carboxy terminus
dsDNA	double-strand DNA
ssDNA	single-strand DNA
SAP	SAF-A/B, Acinus, and PIAS domain
VHS	Vps27p/Hrs/STAM
XLF	XRCC4-like factor
RPA	Replication Protein A

1. Introduction

1.1 DNA Damage and Repair

DNA is under constant assault which threatens the integrity of the genome. Damaged DNA interferes with several cellular processes including transcription, DNA replication, and chromosome separation. In addition DNA damage impedes the accurate inheritance of genomic information from one generation to the next. In order to overcome the potentially harmful effects of DNA damage, DNA repair mechanisms have evolved to eliminate or at the very least, minimize DNA damage.

Specific pathways have evolved to deal with distinct forms of DNA damage. Base excision repair is the main pathway used to repair non-bulky base damage. This pathway becomes activated in response to damaged base residues and nucleotides as well as in response to abasic sites. The removal of 8-oxo-G lesions requires the utilization of base excision repair. Mismatched bases can occur during DNA replication or from non-fidelitous repair and are corrected via the mismatch repair pathway. Mismatched bases are a particularly interesting form of DNA damage in that the cellular machinery must differentiate between the strands are recognize the patental DNA strand which presumably contains the correct DNA sequence. Bulky, helical distorting lesions are recognized and repaired via the nucleotide excision repair pathway. This pathway is responsible for the removal of several forms of DNA damage including thymidine dimers and several alkylating agents[1]. These lesions are initially detected through transcription stalling at the site of the distortion[2]. Additionally, these bulky lesions may be detected by scanning proteins which detect helical distortion of the DNA[2].

While the previous pathways have involved damage to one strand, specific types of agents cause damage to both strands of the DNA. Among these are interstrand crosslinks in which covalent bonds form between bases on separate strands which are usually irreversible[3]. These crosslinks are extremely toxic because they prevent the separation of DNA strands which is essential for cellular processes such as replication and transcription. To resolve such lesions the interstand crosslink repair pathway produces nicks in the strands on either side of the crosslinked bases essentially producing a DNA double strand break. The resulting break is then repaired through the homologous recombination pathway which repairs a subset of double strand breaks and is discussed in detail below.

1.2 Double Strand Breaks (DSBs)

DNA is under constant and unrelenting assault which threatens the integrity of the genome. In order to overcome these genomic stresses, DNA repair pathways have evolved to deal with a variety of DNA lesions. Of these, DNA double strand breaks are a particularly toxic form of DNA damage in that a single unresolved DSB can be lethal[4]. DSBs can result from both endogenous and exogenous sources. Endogenous sources of DNA include exposure to reactive oxygen species, which are created as by-products of metabolism, and DNA replication fork collapse which itself is caused by a variety of events. During replication, DNA is particularly vulnerable to damage. As the replication fork proceeds, if it encounters a lesion such as a backbone nick or a region of extensive supercoiling, the replication machinery dissociates resulting in a one-ended DSB[5].

Exogenous sources of DSBs include exposure to ionizing radiation and radio-mimetic drugs both of which are used in cancer therapies. The toxicity of DSBs makes

the induction of such lesions attractive as therapies as they can be highly effective in killing cancer cells. Exposure to ionizing radiation can cause DSBs directly through at least two mechanisms both of which involve free radical formation[6]. The first mechanism would be the formation of two single strand breaks located within 15 base pairs (bp) and be on opposite strands[6]. The second occurs through a radical transfer mechanism in which the initial radical induced on the DNA produces a single strand break. The radical is then transferred to the other strand causing the second break[7]. Of these two possibilities the latter is statistically more probable[6]. Radiation can also cause DSBs indirectly through the initial creation of reactive oxygen species in the nucleoplasm which then cascade eventually colliding with DNA causing breaks[6]. Several radio-mimetics have been developed to induce DSBs including bleomycin and etopiside[1]. Bleomycin is a group of glycopeptides which were first isolated from *Streptomyces verticillus* in 1966[8]. These drugs consist of a disaccharide-modified metal-binding domain connected through a methylvalerate-Thr linker to a bithiazole C-terminal tail[9]. Chen and Stubbe have proposed a model by which bleomycin can cleave both strands which necessitates the partial intercalation and flexibility of the bithiazole tail[10]. In addition to the bithiazole tail, the pyrimidines moiety plays a role in DNA binding[11]. The N3 and N4 groups of the pyrimidine moiety provide sequence specificity of DNA cleavage by binding to the N3- and N2-amino groups of the guanine 5' to the pyrimidine (py) cleavage site[12]. This sequence specific cleavage is commonly referred to as the "5'-G-Py-3' rule." Cleavage is initiated when bleomycin removes the 4'-hydrogen atom from the C4' of the deoxyribose moiety of the pyrimidine 3' to the guanine[13]. The resulting termini can have either a one base 5' overhang or blunt ends depending on the

base 3' to the Py. 5' overhangs are generated when the adjacent 3' base is a purine while blunt ends are generated when the adjacent 3' base is a Py[12]. Unlike etoposide, bleomycin has not been shown to directly inhibit DNA replication[14]. During DNA replication Topoisomerase 2 functions to relieve torsional strain in DNA caused by supercoiling, DNA catenation, and DNA knotting[15]. Etoposide is a topoisomerase 2 inhibitor which has proven an effective anticancer drug for a variety of cancers[1]. To create DNA damage tyrosines on topoisomerase 2 attack the phosphodiester bonds four bases apart on the opposite strands of G duplex DNA yielding a DSB[15]. Etoposide blocks the religation activity of topoisomerase 2 thus leaving a DSB with a covalently bound protein at the terminus.

To resolve DSBs two main pathways are utilized in mammalian cells. The homologous recombination (HR) pathway involves extensive 5' resection at the site of the break. The resulting 3' single stranded DNA overhang is utilized to seek homologous regions within the genome. These homolgous regions are used as a template to repair the break. Under basal conditions, the HR pathway is restricted to S and G2 phases of the cell cycle[16]. Because the HR pathway uses homologous regions of DNA as a template, it is somewhat fidelitous. In contrast the Non-homologous End Joining (NHEJ) pathway does not use homologous regions of the genome and instead resolves the lesion through direct ligation of the break following limited processing of the DNA termini, if necessary[17]. Despite being more "error prone" the NHEJ pathway repairs DSBs much quicker than HR and is initiated within seconds of the formation of a break. NHEJ is also not restrained to one stage of the cell cycle and instead can occur at all stages[17].

Initiation of NHEJ requires the formation and activation of the DNA-dependent protein kinase (DNA-PK). Once activated, DNA-PK functions to regulate NHEJ pathway progression as well as initiates the DNA damage response (DDR) signaling cascade (**Figure 1**). The DDR initiates cell cycle arrest which is thought to allow time so that the lesions can be resolved. In addition to its direct role in NHEJ, recent work has demonstrated that DNA-PK also regulates HR and likely plays a critical role in DSB repair pathway choice through autoregulation[18-20]. This will be discussed in detail below.

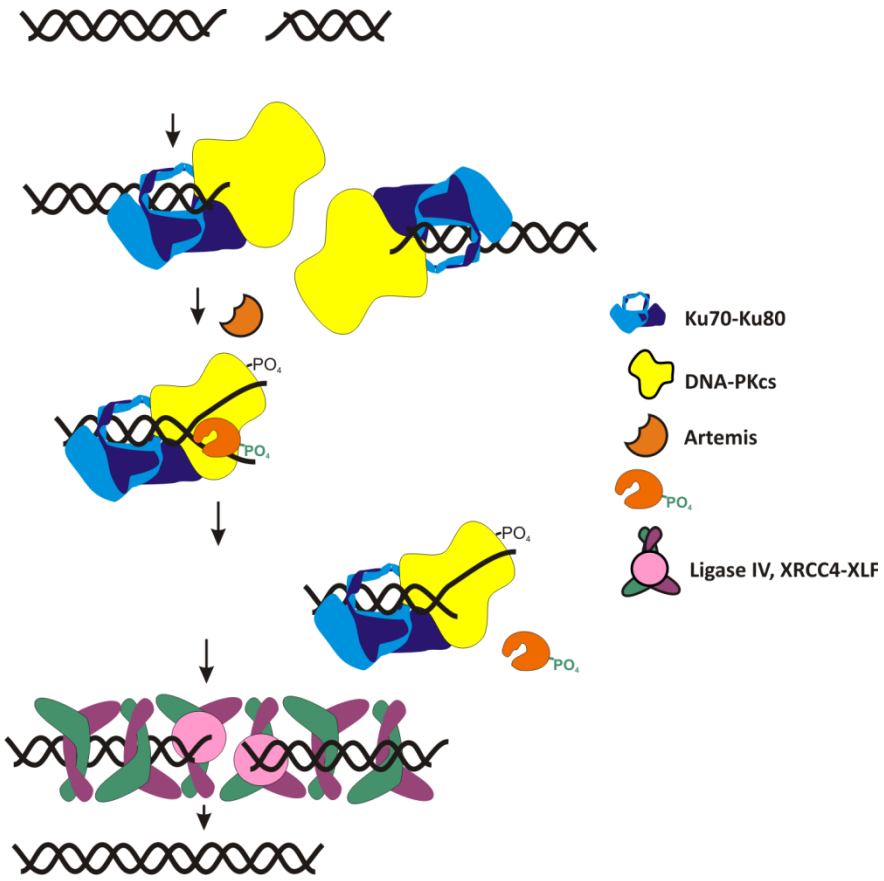


Figure 1. Model of Non-homologous End Joining DNA Repair Pathway. Following a DSB the pathway is initiated by the Ku70/80 binding to the DNA termini. DNA-PKcs is recruited to form the DNA-PK heterotrimer. DNA-PK undergoes autophosphorylation as well as phosphorylation of downstream molecules. DNA termini are processed and DNA-PK dissociates from the DNA. DNA Ligase IV/XRCC4-XLF complex facilitates ligation. Adapted from reference[17].

1.3 DNA Damaging Agents in Cancer Therapy

DNA damaging agents have been used historically to treat cancer and remain the most widely used approach in treatment[1]. Unlike the vast majority of cells in the human body, cancer cells are rapidly dividing and thus can be targeted with genotoxic agents. Ionizing radiation is well established producer of DSBs and is a mainstay in cancer therapy. The toxicity associated with DSB production makes the induction of such lesions attractive as therapies as they can be highly effective in killing cancer cells. Exposure to ionizing radiation can cause DSBs directly through at least two mechanisms both of which involve free radical formation[6]. The first mechanism would be the formation of two single strand breaks located within 15 base pairs (bp) and be on opposite strands[6]. The second occurs through a radical transfer mechanism in which the initial radical induced on the DNA produces a single strand break. The radical is then transferred to the other strand causing the second break[7]. Of these two possibilities the latter is statistically more probable[6]. Radiation can also cause DSBs indirectly through the initial creation of reactive oxygen species in the nucleoplasm which then cascade eventually colliding with DNA causing breaks[6].

Several radio-mimetic cancer drugs have been developed to induce DSBs including bleomycin and etopiside[1]. Bleomycin is a group of glycopeptides which were first isolated from *Streptomyces verticillus* in 1966[8]. These drugs consist of a disaccharide-modified metal-binding domain connected through a methylvalerate-Thr linker to a bithiazole C-terminal tail[9]. Chen and Stubbe have proposed a model by which bleomycin can cleave both strands which necessitates the partial intercalation and flexibility of the bithiazole tail[10]. In addition to the bithiazole tail, the pyrimidines moiety

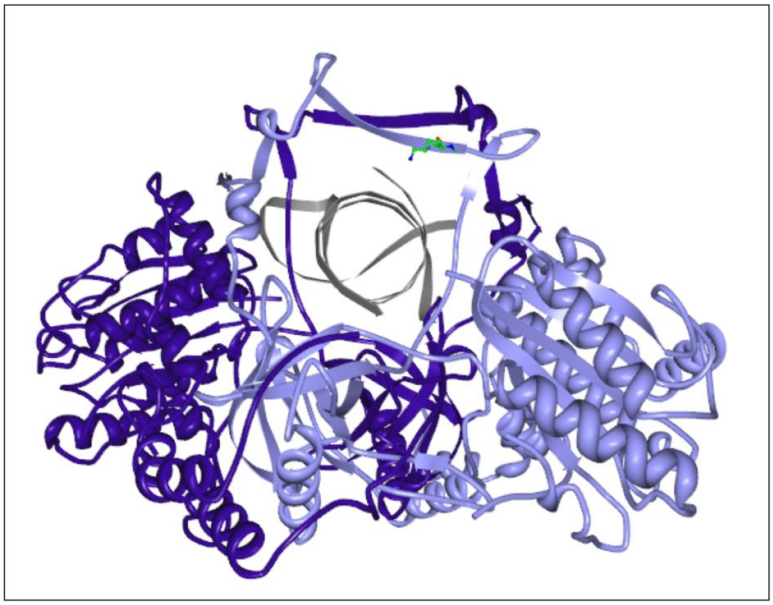
plays a role in DNA binding[11]. The N3 and N4 groups of the pyrimidine moiety provide sequence specificity of DNA cleavage by binding to the N3- and N2-amino groups of the guanine 5' to the pyrimidine (py) cleavage site[12]. This sequence specific cleavage is commonly referred to as the "5'-G-Py-3' rule." Cleavage is initiated when bleomycin removes the 4'-hydrogen atom from the C4' of the deoxyribose moiety of the pyrimidine 3' to the guanine[13]. The resulting termini can have either a one base 5' overhang or blunt ends depending on the base 3' to the Py. 5' overhangs are generated when the adjacent 3' base is a purine while blunt ends are generated when the adjacent 3' base is a Py[12]. Unlike etoposide, bleomycin has not been shown to directly inhibit DNA replication[14]. During DNA replication Topoisomerase 2 functions to relieve torsional strain in DNA caused by supercoiling, DNA catenation, and DNA knotting[15]. Etoposide is a topoisomerase 2 inhibitor which has proven an effective anticancer drug for a variety of cancers[1]. To create DNA damage tyrosines on topoisomerase 2 attack the phosphodiester bonds four bases apart on the opposite strands of G duplex DNA yielding a DSB[15]. Etoposide blocks the religation activity of topoisomerase 2 thus leaving a DSB with a covalently bound protein at the terminus.

1.4 Ku70/80

DNA-PK consists of three proteins: Ku70, Ku80, and the DNA-dependent protein kinase catalytic subunit (DNA-PKcs). Ku70 and Ku80 form a tight heterodimeric complex referred to as Ku[21]. Ku was originally identified in 1981 when following the characterization of sera from patients with polymyositisscleroderma overlap syndrome when more than half of the patients' sera recognized Ku as an autoantigen[22]. *In vivo* Ku is found as an extremely stable complex. Ku80 deficient rodent cells have very low levels of Ku70 and Ku70 deficient mice have very low levels of Ku80[23,24]. In addition expression of Ku80 in Ku80-deficient cells restores both Ku80 and Ku70 protein levels implying that the heterodimeric form of Ku is the stable and functional form[25]. Together Ku70/80 bind to dsDNA termini and serve as a scaffolding complex. The Ku70/80 complex promotes several accessory proteins including DNA-PKcs to localize to the site of a DSB. Among these proteins is the DNA Ligase IV/Xrcc4/XLF complex which is responsible for ligating the ends of the break together (**Figure 1)**. Importantly, biochemical studies have shown that Ku stimulates this ligation activity of the complex[26]. Recent structural data has further defined the nature and role of Ku/Xrcc4/XLF interactions demonstrating that Xrcc4/XLF forms a filamentous complex which, when in complex with Ku, can tether DNA termini across a break[27]. The tethering of DNA termini across a DSB is commonly referred to as the synaptic complex. Ku itself has also been implicated in synaptic complex formation *in vitro*. Previous studies utilized atomic force microscopy to visualize protein/DNA complex formation, which show that Ku binds to DNA termini and aligns termini independent DNA-PKcs[28,29]. Interestingly, the presence of DNA-PKcs does not stimulate the alignment of these termini suggesting that

Ku is the main factor in the alignment activity[29]. More recent data has demonstrated that Ku has enzymatic activity on DNA. Specifically Ku is a 5'-dRP/AP lyase and can excise abasic sites near DNA termini which would otherwise interfere with the ligation step of NHEJ[30]. Specifically Ku was shown to have the greatest lyase activity on substrates where the abasic site is within a short 5' overhang[31]. On the other hand, excision activity is strongly suppressed by as little as two paired bases 5' of the abasic site[31]. While Ku stimulates the processing of abasic sites near the terminus, it has also been shown to limit endonuclease and exonuclease resection of DNA termini which may limit the amount of lost DNA during repair[32]. Obviously Ku serves several functions in NHEJ, all of which are absolutely dependent on the binding of Ku to DNA termini which occurs with relatively high affinity as indicated by Kd values of 2.4×10^{-9} to 5×10^{-10} M-1[33,34]. The crystal structure of Ku has provided considerable information regarding the mechanism by which Ku binds DNA in a sequence-nonspecific mechanism with high affinity. The structure of Ku was solved in the absence of the N-terminal and C-terminal regions of both Ku70 and 80 which were subsequently solved separately (**Figure 2**)[35]. Despite having poorly conserved primary sequences (about 15% identity), Ku 70 and 80 have high structural homology.

A.



B.

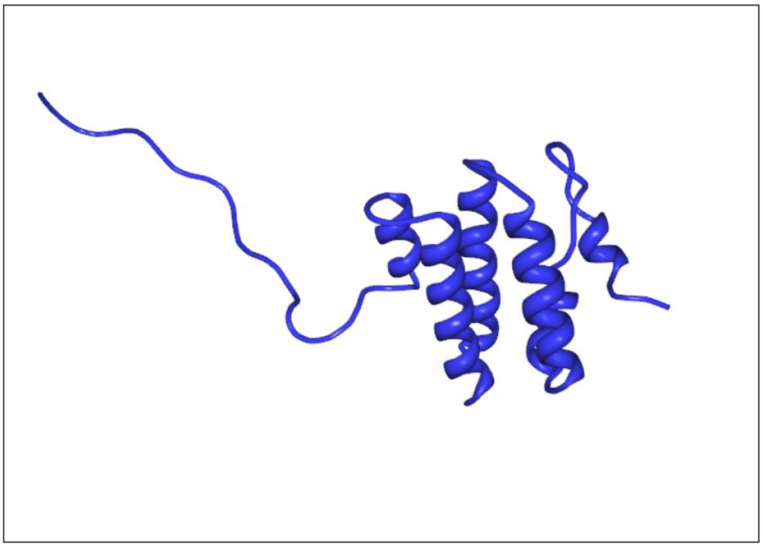


Figure 2. Ku Structures A. Ku70/80 Structure adapted from PDB file 1JEY[35]. Ku80 is shown in dark blue and Ku70 is shown in light blue. Structures were solved in the absence of the Ku80 C-terminus and bound to DNA(Grey). B. The C-terminal region of Ku80 adapted from PDB file 1Q2Z[36].

Both proteins contain three distinct regions: an amino terminal von Willebrand A domain, a central core dimerization/DNA binding domain, and divergent C-terminal regions. Crystal structures determined from X-ray crystallography revealed that together the two Ku subunit asymmetric ring with an extensive base and a narrow bridge and pillar region[35]. Ku70/80 crystals were solved in the presence and absence of dsDNA and consistent with results from our lab, DNA binding was not shown to cause major structural changes in the core structures of the protein[37]. These data also indicate that the ring structure of Ku accommodates a dsDNA molecule which validates early models of Ku/DNA binding in which Ku acts like a bead and is threaded onto DNA. The DNA binding channel is lined with positively charged residues. These amino acids bind to about two turns of DNA along the extensive base of the molecule. Further, the structure reveals that the interactions between Ku and DNA are exclusively between the protein and the sugar/phosphate backbone, which helps explain how Ku binds dsDNA in a sequence non-specific manner. The asymmetrical binding also allows the distinction between Ku70 and 80 in terms of localization of binding. Consistent with previous work involving covalent cross-linking, the high resolution structure clearly shows that Ku70 makes major contacts with DNA toward the terminus, while Ku80 occupies adjacent bases toward the continuous strand of the DNA[38].

As is often the case, obtaining crystal structures of full length Ku70 and 80 proved difficult and thus only the core domain was used for crystallography. Thus the structure of the C-terminal regions of Ku70 and 80 were solved independently by different groups. Zhang et al solved the structure of the Ku70 C-terminus encompassing amino acids 536-609[39]. Using nuclear magnetic resonance it was found that residues

536-560 are highly flexible and could be considered extensively disordered. Residues 561-609 on the other hand form a well defined structure with three alpha helices which create a putative DNA binding domain. The putative DNA binding domain does not resemble the helix-turn-helix DNA binding motifs or other common sequence specific DNA-binding domains seen in zinc fingers or leucine zippers[39]. Instead it contains a SAP domain which was named for three protein containing this motif: SAF-A/B, Acinus, and PIAS[40]. The SAP domain is a slight variation of a helix-extension-helix fold which is often involved in binding unusual DNA structures such as holiday structures, three-way junctions, heteroduplex loops, base mismatches, bulky adducts and curved DNA[41]. Exactly what role this domain plays in DNA repair remains unknown. A reasonable hypothesis is that this regions DNA binding activity is responsible for Ku's "pausing" at certain regions when translocating along DNA[21]. Consistent with this hypothesis, a recent molecular modeling study predicts that the SAP may bind to specific sequences with a relatively high binding affinity ($\Delta G \approx -20$kcal/mol)[42]. The authors speculate that the SAP domain may be involved in gene regulation on intact chromosomes however more studies are needed in order to validate this claim.

Similar to the C-terminus of Ku70, in order to obtain high quality structural data of the C-terminus of Ku80 nuclear magnetic resonance was employed[36]. Results indicate that the C-terminus of Ku80 has three distinct structural regions: a disorder region between amino acids 550-593, a helical bundle between amino acids 594-704, and a disorder region at the extreme C-terminus between amino acids 705-732. Importantly for findings in our study, all of the C-terminal constructs used in this study were well behaved and did not form self aggregates. Within the six helices there are two significant

invaginations between helix two and helix four that the authors predict to be docking sites for a putative protein or peptide ligand. Overall the C-terminal helical bundle resembles conserved superhelical repeats seen in MIF4G from the human nuclear cap binding protein 80kDa subunit and Vps27p/Hrs/STAM (VHS) domain from the *Drosophila melanogaster* protein Hrs. Of these two, the Ku80 C-terminus is more similar to the VHS domain in size with the inclusion of the disordered region at the extreme carboxy end; however the VHS domain contains eight helices. The authors speculate that the extreme C-terminus of Ku80 may adopt the "missing" helical elements of the homologous VHS domain structure upon binding to DNA-PKcs for which this region of the protein has been implicated[36,43].

Our interest in the C-terminus of Ku80 first peaked while studying DNA-dependent conformation changes in the Ku heterodimer[37]. Consistent with Ku crystallography studies, Lehman and Turchi found that the core domain of Ku does not undergo considerable conformational changes upon binding DNA. However, the C-terminal regions of both Ku70 and Ku80 undergo extensive changes following Ku binding to dsDNA.

1.5 DNA-PKcs

DNA-PKcs is the largest single polypeptide consisting of 4128 amino acids and a staggering mass of 469kDa[17]. The large size of DNA-PKcs hinders several aspects of study but most notably high resolution structural studies. Despite recent advances in crystallography to date, the best resolution achieved is at 6.6 Angstroms which is far from ideal[44]. At this resolution side chains, autophosphorylation sites, DNA binding specific atoms cannot be resolved. Nevertheless results from this study do provide important

information regarding the structure of this massive protein. DNA-PKcs is comprised of several distinct structural domains. A large ring structure containing many alpha-helical HEAT repeats (helix-turn-helix motifs) surrounds a putative dsDNA binding domain. Unlike the ring structure of Ku, the ring structure in DNAPKcs contains a clearly defined gap[44]. Adjacent to the gap are at least 66 helices arranged as HEAT repeats which comprise the majority of the ring structure which is folded into a concave shape when viewed from the side. The ring most likely consists entirely of the amino terminal region of DNA-PKcs. While such HEAT repeats have been observed in other structures these are structurally irregular in that the polypeptide chain has its amino terminus on one side of the gap, circumnavigates the ring, and then reverses direction to the other side of the gap[44]. These authors predict that these two irregular helical regions are points of conformational flexibility that could widen the gap following phosphorylation of specific residues thus facilitating DNA-PKcs release from DNA. This model is consistent with biochemical studies discussed later in this chapter. At the top of the ring is the "forehead" domain which is angled forward and is partially responsible for the concave shape of the protein. This region supports a head domain also referred to as the crown domain. Using homologous structures, it was found that the head domain contains the kinase domain which is flacked by the 500 amino acid FAT domain (named after PI3kinase family members FRAP, ATM, and TRRAP) and a highly conserved 35 amino acid domain called the FATC domain. Accordingly, the FAT and FATC domains are thought to interact with each other with the kinase domain wedged between. Interestingly, unlike previous cryo-electron microscopy data, the presence of a putative single stranded DNA binding site is absent in the crystal structure. Whether both distinct

dsDNA and ssDNA binding sites exist in DNA-PKcs is an important question that deserves further investigation.

1.6 Ku/DNA-PKcs Interactions

There is a plethora of data, both biochemical and structural, supporting protein/protein interactions that occur between DNA-PKcs and Ku, primarily involving the C-terminal region of Ku80. As mentioned above, Ku is highly conserved and can be found in all kingdoms of life[21]. Additionally, several of the accessory proteins involved in NHEJ are present across species as well including Ligase 4/Xrcc4 and XLF homologs. While the core /DNA binding domain of Ku exists across species, only vertebrates contain the extensive C-terminus of Ku80[43]. To date DNA-PKcs has also only been identified in vertebrates. This evidence alone is enough to speculate that the C-terminus of Ku80 may influence DNAPKcs. Because of its central importance to my work, evidence for these interactions will be extensively discussed below.

One of the first studies to provide evidence for a direct interaction between DNA-PKcs and the C-terminus of Ku80 was completed by Gell and Jackson in 1999[43]. Here relatively simple "pull-down" methods were employed using purified DNA-PKcs and N-terminal and C-terminal deletion constructs of Ku80. Results defined the extreme C-terminus of Ku80 as being necessary and sufficient for a DNA-PKcs interaction. It should be noted however, that these studies were completed in with Ku80 in the absence of Ku70 and thus it cannot fully represent the heterodimeric complex with DNA-PKcs. That same year another group tested the influence of the C-terminus of Ku80 on DNA-PKcs activation using a variety of *in vitro* and *in vivo* techniques. Using a Ku80 construct with amino acids 554-732 deleted, the authors demonstrate that this deletion

dramatically reduces DNA-PKcs kinase activity *in vitro* but partially rescues the radio sensitive phenotype of Ku80 deficient cell lines[45]. Further they found that this C-terminal deletion does not influence the ability of Ku to bind dsDNA. Elegant follow up work from Jackson and colleagues further defined what they termed "the PIKK interaction motif" first identified by their study reported in 1999. DNA-PKcs is a member of the phosphatidylinositol-3 kinase-like protein kinases (PIKKs) which also includes the DNA Damage Response (DDR) initiation kinases ATM and ATR[1]. Nbs1 and ATRIP have been implicated in activating ATM and ATR respectively. Sequence alignments with Ku80, Nbs1, and ATRIP show significant sequence similarity among the C-termini of these proteins[46]. Extensive biochemical analysis showed strong evidence to support that this PIKK interaction domain is required for the recruitment of each of the kinases to damaged DNA and activation of the respective kinase[46].

In contrast to these reports, a more recent study added another level of complexity to these findings. Weterings and colleagues used a variety of *in vitro* and *in vivo* technique to characterize Ku80 mutants which did not contain the extreme C-terminal "PIKK binding motif"[47]. Similar to the result reported from the Jeggo laboratory, introduction of C-terminal deletion mutants into Ku80 null cells partially rescued the radiosensitive phenotype[45]. Interestingly no distinguishable difference was observed in this assay between cells complimented with Ku80 containing a C-terminal 24 amino acid deletion or a 163 amino acid deletion. Further, *in vitro* electro mobility shift assays (EMSAs) and *in vivo* fluorescent monitoring indicate that the final 163 amino acids of Ku80 are dispensable for DNA-PKcs recruitment to dsDNA breaks. Subsequent monitoring of DNA-PK kinase activity showed that in the absence of the helical bundle

and the extreme C-terminus, DNA-PK activity was reduced by 50%. Clearly these data differ from previous studies and suggest that the PIKK interaction domain located at the extreme C-terminus is not necessarily required for the Ku/DNA-PKcs interactions driving DNA-PKcs recruitment to DNA, retention on the DNA, or DNA-dependent kinase activity. This does not diminish the role of Ku70/80, however, as it is still required for stablizing DNA-PKcs-DNA interactions and stimulation of kinase activity.

While there is an abundance of biochemical evidence for Ku80 C-terminus/DNAPKcs interactions, structural data are lacking. As previously stated, the sheer size of DNA-PKcs makes obtaining any information about structure difficult to obtain and even harder to interpret. However in-depth analyses in the few structural studies that have been completed provide insight into the nature of these potential interactions. The first piece of data we should consider is that the crystal structure of DNA-PKcs discussed above was solved in the presence of the C-terminus of Ku80 spanning residues 539-732[44]. In fact DNA-PKcs did not crystallize in the absence of the C-terminus. The authors suggest that this is indicative of a role for the C-terminus in stabilizing DNA-PKcs conformations. Further, small angle X-ray scattering data indicate upon DNA-PK/DNA-PK synaptic complex formation (**See Figure 1**) the C-terminus of Ku80 becomes extended enough to interact with the DNA-PKcs molecule on the opposite site of the break[48]. This study also indicates that an extensive interaction interface exists between the core DNA binding/dimerization domain of Ku and DNA-PKcs. This suggests that multiple protein/protein interactions exist between Ku and DNA-PKcs that do not involve the Ku80 C-terminus, however whether the Ku80 C-terminus is required to allosterically modulate these interactions remains unknown.

1.7 DNA/DNA-PK Interactions

Less well characterized, but equally important to understanding DNA-PK biochemistry, are the DNA/DNA-PK interactions. Multiple electron microscopy studies and atomic force microscopy studies have determined that DNA-PKcs interacts with the terminus of dsDNA[29,49,50]. This is consistent with the DNA binding activities of its binding partner Ku which does not interact with circular DNA but instead only interacts with dsDNA termini. Interestingly, data obtained from surface plasmon resonance analysis suggests that DNA-PKcs in the absence of Ku shows preferential DNA binding of some terminal DNA structures with non duplexed single stranded bases[51]. This same group using highly sensitive photo cross-linking methods determined that major DNA-PKcs-DNA interaction sites are confined to the most terminal 10 bases[38]. It should be noted however, that the DNA substrates used in these experiments were 32 base pairs and thus differences in binding may be observed on longer substrates that are less constricting.

How these interactions influence DNA-PK activity has been a significant research interest in our lab. We were the first to report that subtle changes in DNA sequence and structure can have profound influences on DNA-PK activation. Using relatively short DNA substrates of 30 base pairs or less, we determined that DNA-PK is preferentially activated by duplex ends containing a 3' pyrimidine-rich and 5' purine-rich strands[52]. The role of single strand overhangs adjacent to duplex DNA was also investigated and interestingly distinct roles for both 5' and 3' DNA overhangs were identified[53]. 5' single stranded overhangs seem to preferentially stimulate DNA-PK activity. The role of 3' single stranded overhangs is slightly more complex in that these stimulate synaptic

complex formation through DNA micro-homology mediated end tethering. The formation of this synaptic complex in and of itself stimulates kinase activity[53]. Together these data suggest that multiple factors concerning DNA including sequence influence stimulation of DNA-PK.

Interesting structural studies have shown that DNA/DNA-PKcs interactions influence protein complex confirmations. Multiple lines of evidence suggest that DNA-PKcs molecules interact across DNA termini during NHEJ and are discussed below. Using small angle X-ray scattering Hammel and colleagues found that when bound to 40base pair DNA containing hairpin structures, DNA-PKcs dimers interact through their respective head domains[48]. In contrast, upon binding to DNA with non-complementary terminal bases often termed "Y" DNA, DNA-PKcs dimers interact through their large ring domains.

1.8 Downstream phosphorylation targets of DNA-PKcs

Efficient activation of DNA-PK is critical for NHEJ and the DDR[1]. Thus far over 700 targets of DNA-PK phosphorylation have been identified, many of which contain the S/T-Q consensus sequences[54], however several important signaling sites have been identified outside of these consensus sequences[55]. Many of the proteins involved in NHEJ, including Ku, have been shown to be phosphorylated by DNA-PK, however mutating these sites results in no marked change in NHEJ as measured by radio-sensitivity[56]. The notable exception to this is DNA-PKcs itself which has been shown to undergo trans-autophosphorylation at over 30 sites and phosphorylation of these sites has distinct influences on NHEJ and DDR[54,57,58]. For example phosphorylation of the ABCDE cluster centering on S2056 promotes end processing of the DNA termini which

has been suggested to open the protein/DNA complex to allow proteins to access the DNA. Conversely phosphorylation of the PQR cluster has been suggested to limit end processing by sterically hindering the DNA termini from processing. Phosphorylation of these clusters also appears to have opposing roles in DSB repair pathway choice as phosphorylation of the ABCDE cluster promotes NHEJ while phosphorylation of the PQR cluster promotes HR[19]. Further complicating DNA-PKcs auto-regulation is the recently identified N and T sites which ablate kinase activity upon autophosphorylation[19]. Whether these and other distinct signaling events result from responses to specific DNA lesion structures or as a response to other cellular processes such as cell cycle, remains unknown.

Replication protein A (RPA) is a heterotrimeric, ssDNA binding protein consisting of 70, 32, and 14 kDa subunits which is also a downstream phosphorylation target of DNA-PK[55]. Due to its ability to bind ssDNA, RPA has been shown to be involved in several nuclear pathways including DNA replication and DNA repair[2]. In terms of the repair of double strand breaks, RPA has been shown to play an integral role in HR. As mentioned above recent work demonstrates that DNA-PK has been implicated in playing a major role in HR through autophosphorylation. Additionally DNA-PK targeting of RPA may regulate HR. At least 7 sites located at the N-terminus of RPA 32 undergo phosphorylation during the DNA damage response. The phosphorylation status regulation of each of said sites and their influence on the neighboring sites within the N-terminus have recently been elucidated by Oakley and colleagues[55]. Using in vivo and in vitro data they show that Serine 4 and Serine 8 are phosphorylated by DNA-PKcs. The phosphorylation of these serines significantly stimulates the phosphorylation of Ser12 by

ATM and DNA-PKcs while moderately stimulating the phosphorylation of Ser 33 by ATR. These phosphorylation events in RPA likely play a major role in regulating ATR activity which in and of itself is a major regulator of the HR pathway. Besides regulating ATR activity there are important cellular consequences of RPA hyperphosphorylation following DNA-PKcs activation. A recent study showed that RPA interacts with the tumor suppressor p53; however, this interaction is ablated upon hyperphosphorylation of RPA 32 at the N-terminus by DNA-PK[59]. In this study the authors provide a model in which the dissociation of RPA and p53 may be important for RPA to function in DNA repair pathways. Others have shown that RPA 32 hyperphosphorylation facilitates NHEJ by suppressing sister chromatid exchanges[60].

DNA-PK has also been shown to phosphorylate p53 on serine 15 and serine 37[61]. Under basal conditions p53 levels in the cell remain low and p53's transcriptional activity are inhibited through its interaction with Mdm2. The phosphorylation sites targeted by DNA-PK are located in p53's activation domain and upon phosphorylation, the interaction of p53 and Mdm2 is ablated thus releasing p53 from its Mdm2-mediated inhibition. Activation of p53 leads to the expression of several cell cycle arrest and DNA repair factors. Thus p53 is an important downstream target of DNA-PK which signals to the cell that a DSB has occurred.

1.9 Significance

Owing to the importance of DNA-PK in response to DSBs, I sought to better understand the factors which influence DNA-PK activation. The major contributors to this activity thus far identified are protein-protein interactions and protein-DNA interactions. As discussed in detail above, evidence suggests that the C-terminus of Ku80

may play a major role in DNA-PK holoenzyme formation, however the location of potential protein-protein interactions and their influence is debated. Protein structure-function studies provide significant insight into biochemical mechanisms. This investigation used this approach in order to determine which of the structurally distinct regions of the Ku80 C-terminus influence DNA-PK activation. Even less well understood are the protein-DNA interactions which influence DNA-PK activation. Our lab has published that the structure and sequence of DNA can modulate DNA-PK activity. This is of great interest because DNA-PK must respond to a large variety of DNA structures and sequences in order to efficiently repair DSBs. How this is accomplished by one protein complex remains unknown. Here I define the mechanism by which DNA-PK is modulated by various DNA substrates through a variety of protein-protein interactions involving the C-terminus of Ku80 and protein-DNA interactions involving DNA-PKcs.

2. Materials and Methods

2.1 Ku Mutant Design

The C-terminal region of Ku80 contains 3 distinct structural regions[36]. A disordered linker region exists between amino acids 550 and 594. Amino acids 595-704 fold into a helical bundle and an additional disordered region is found at the extreme C-terminus from amino acids 705-732. Ku80 mutants were designed to specifically delete these distinct structural features. Mutants will be referred to by their last amino acid. For example the 550 mutant contains only amino acids 1-550 and does not contain amino acids 551-732. The 704 mutant did not contain a disordered region at the extreme C-terminus while the 594 mutant did not contain this disordered region or the helical bundle.

2.2 Ku Mutant Construction

WT Ku80 and the 550 mutant constructs were already available in the lab and were used in this study with no notable changes[62]. Site-directed mutagenesis was used to insert stop codons into the Ku80 sequence to yield the desired C-terminal deletion constructs using standard PCR techniques. Parental plasmid containing WT Ku80 gene was pFastBac1 and was used to generate the 704 and 594 mutants (Invitrogen). Primer sequences can be found in **Table 1**. 100µl PCR reactions were used containing ~250ng of each primer, 200µM dNTPs, 250ng of template DNA (pFastBac1-Ku80), 10µl of 10x Pfu buffer, and 5 units of Pfu polymerase. Reactions were cycled sixteen times for 30 seconds at 95^{o}C, 1 minute at 55^{o}C, and 10 minutes at 68^{o}C. After the sixteenth cycle, incubation was continued at 68^{o}C for an additional 10 minutes. Following PCR, products were transformed into XL1-blue competent *E.coli* cells. 1µl of PCR product was added

to 14µl of sterile water. DNA was then added to 200µl of bacteria cells and placed on ice for 20 minutes. Tubes were heat shocked for 45 seconds at 42^{o}C. Following heat shock cells were placed back on ice for an additional 2 minutes. 1ml of LB broth was added to tube and incubated at 37^{o}C for 1 hour with shaking. Transformation reactions were then diluted and plated on LB agar plates complemented with ampicillin. Following overnight incubation at 37^{o}C, colonies were picked and 5ml LB tubes were inoculated. Mini preps were completed to isolate plasmid DNA according to manufactures instructions (Qiagen). Correct insertion of desired codons was completed using T7 forward and reverse primers provided by Indiana University DNA Sequencing Core Facility.

Once insertions of stop codons were confirmed from DNA sequencing, pFastBac1-Ku80-594 and pFastBac1-704 were transformed into DH10Bac *E. coli* cells which contain a Bacmid genome (Invitrogen Bac-to-Bac Baculovirus Expression System). Transformation into DH10Bac cells can completed as before except that S.O.C. media was used following the heat shock and transformations were plated on LB agar plates containing 50µg/ml kanamycin, 7µg/ml gentamicin, 10µg/m tetracycline, 100 g/ml X-gal and 40µg/ml IPTG. Plates were incubated for 48 hours at 37^{o}C and white colonies were picked for analysis by PCR. Bacmid DNA was purified using PureLink HIPure Plasmid Miniprep Kit according to manufactures instructions (Invitrogen).

Recombinant bacmid DNA was then transfected into SF9 insect cells using Bacfectin. 1x 10^{6} cells were plated in a 35 mm dish an allowed to adhere to plate for 1 hour at 27^{o}C. Media was removed and cells were washed with serum-free media and incubated for an additional 30 minutes. 2µg of bacmid DNA was diluted into a final

volume 96µl with sterile water. 4µl of Bacfectin was added to the diluted DNA and incubated at room temperature for 15 minutes. The mixture was added dropwise to the cells while gently swirling. Transfected cells were incubated at 27^{o}C for 5 hours before media containing 10% FBS protein was added. Following addition of media containing serum, cells were incubated at 27^{o}C for 72 hours to allow production of viruses and the tissue culture supernatant was collected and saved in sterile tubes.

Recombinant virus clones were isolated from the virus containing supernatant via plaque assays with serial dilutions ranging from 10^{-1} to 10^{-3}. 1x 10^{6}cells were plated in each well in a 6 well-35mm dish. 100µl of virus dilutions were added to each well and incubated for one hour at room temperature. Sterile 4% low melting agarose was melted and diluted with Grace's media to a final dilution of 1%. Diluted virus inoculums were removed and 1.5 ml of the 1% agarose/media mixture was added to each well. Following media solidification 1ml of Grace's media was added on top of the agarose. Cells were incubated for 5 days at 27^{o}C. Media was removed and 1ml of 0.03% neutral red dye diluted in sterile PBS was added on top of the agarose and incubated at room temperature for one hour. Dye was removed from plates and dishes were inverted and stored overnight in the dark to allow plaques to clear. Plaques appearing as clear were picked using a glass Pasteur pipettes and resuspended in 1ml of Grace's media. Viruses were allowed to diffuse out of agarose plug overnight and the resulting supernatant was termed as the primary virus. Isolated virus was then amplified to generate a passage one viral stock by infecting 5 x 10^{5} SF9 cells in 6-well dishes for five days and removing supernatant. The resulting P1 virus was further amplified until a P3 viral stock was generated. Following each amplification, cells lysates were analyzed by western blot to

ensure that cells were expressing the Ku80 contructs. The titer of the P3 viral stock was determined via plaque assay as before however virus was serial diluted further and dilutions tested ranged from 1 x 10^{-5} to 1x10^{-9}. Resulting plaques were counted and titers were determined using the following equation.

Titer ____(pfu/ml) = (average number of plaques * 10) / (dilution factor)

Determining an accurate titer of the Ku80 constructs viruses was critical for protein production. Titers for P3 stocks typically ranged from 5 x 10^{7} to 1.5 x 10^{8}pfu/ml.

2.3 Protein expression of Ku70/80

Ku80 protein constructs were coexpressed with Ku70 and purified at a desired ratio of 1:1 therefore it was necessary to infect SF9 cells at the same multiplicity of infection (M.O.I.) Using titrations we determined that the optimal M.O.I. for Ku expression in SF9 cells is 10. P3 stocks of each of the Ku80 constructs (WT, 704, 594, and 550) were used in co-infections of SF9 cells with P3 stocks of Ku70 virus. Following infection or 100ml of cells at 1 x 10^{6} cells/ml, SF9 cells were incubated for 48 hours at 27^{o}C. Infected cells were pelleted at 4,000 x g at 4^{o}C for 15 minutes.

2.4 Protein Purification

2.4.1 Purification of Ku70/80

In total four different Ku constructs were expressed and purified from SF9 cells. These constructs only differed in the C-terminal region of Ku80 and the same purification scheme was used for all Ku purifications.

Pelleted cells were resuspended in 15 ml of Extraction Buffer containing: 50mM sodium phosphate (pH 7.8), 1M potassium chloride, 10% glycerol, 0.25% Triton X-100,

7mM 2-mercaptoethanol, and 20mM imidazole. Importantly, all buffers used in purifications were supplemented with protease inhibitors (0.5mM PMSF and 1µg/ml each of leupeptin and pepstatin). Resuspended cells were then dounce homogenized and sonicated at micro-tip limit (50%) for 5 pulses on ice. The solution extract was then sedimented at 10,000 x g at 4^{o}C for 30 minutes and cell free extract collected for purification.

Ku70 constructs were designed to contain a N-terminal hexa-histidine tag to simplify purification. Accordingly, cell free extracts were loaded onto a 2ml NTA-Ni column and flow through was collected. The flow through was then loaded onto the column again to increase protein concentration. The column was then washed with 5 column volumes (10ml) of Extraction Buffer with 10mM imidazole to wash away unwanted proteins. Protein was then eluted off column with Extraction Buffer supplemented with 350mM imidazole and fractions were collected in 1ml aliquots using a fraction collector. The fractions containing the highest protein concentrations were determined via Bradford assay using standard techniques. Fractions containing significant amounts of protein were further analyzed via SDS-PAGE with coomassie staining. Due to the "bump" nature of the elution, peaks tended to cluster in 3ml of elution. Peak fractions were polled and dialyzed overnight in Buffer A containing: 25mM Tris (pH8.0), 75mM potassium chloride, 10% glycerol, 0.005% Triton X-100, and 2mM dithiothreitol (DTT).

After at least 12 hours of dialysis, protein was removed from dialysis and loaded onto a 2ml Q-Sepharose column equilibrated in Buffer A at 0.5ml per minute. Column was then washed with 5 column volumes (10ml) of Buffer A. To eluted the protein from

the column, a linear salt gradient was used from a 75mM potassium chloride and ending with 1M potassium chloride Buffer A in ten column volumes (20ml). The elution was collected in 1ml fractions and the protein was analyzed via Bradford assay. Peak fractions were analyzed using SDS-PAGE and coomassie staining. Fractions with the most pure Ku at the highest concentrations were pooled and dialyzed in a final hepes buffer containing: 20mM Hepes (pH 7.0), 75mM potassium chloride, 10% glycerol, 0.005% Triton X-100, and 2mM DTT.

2.4.2 Purification of DNA-PKcs

Our lab has used several schemes to purify heterotrimeric DNA-PK to varying degrees of success. In all of the experiments described in this investigation, it was essential that DNA-PKcs be purified free from WT Ku70/80. Thus the method I used was unique in order to obtain purified DNA-PKcs. It should also be noted that this scheme gave us the most active DNA-PKcs in terms of kinase activity of any other attempted in the Turchi Lab. To date there is no over-expression system for DNA-PKcs and thus it was purified from Hela cell pellets. Unlike other purification schemes for DNA-PK which may require using between 15 and 40 liters of cells, we found that using significantly less cells, ~4L Hela pellets containing 4×10^9 cells, was optimal. In addition we found that cell pellets purchased from the National Cell Culture Center yielded more active DNA-PKcs than what we were able to obtain from cells we grew ourselves. Cell pellets were resuspended in 26ml of hypotonic buffer containing 10mM Tris-HCl (pH8.0), 1mM Ethylenediaminetetraacetic acid (EDTA), and 5mM DTT and incubated on ice for 20 minutes. All buffers were complemented with protease inhibitors with final concentrations of 0.5mM PMSF and 1μg/ml each of leupeptin and pepstatin.

Following incubation, resuspended cells were dounce homogenized with 10-20 strokes of a tight pestle. Lysate was then incubated at 4^{o}C and 26ml of high salt buffer containing 50mM Tris-HCl (pH 8.0), 10mM magnesium chloride, 2mM DTT. 25% sucrose, and 50% glycerol was added dropwise with gentle stirring. 3ml of saturated ammonia methyl sulfate was added with gentle stiring and incubated for 30 minutes at 4^{o}C. The lysate was sedimented at 30,000 rpm for 1 hour at 4^{o}C. The supernatant was then dialyzed overnight without potassium chloride in Buffer A. Buffer was changed twice with approximately 5 hours in each condition. A total volume of 4 liters was used in this dialysis.

The following morning lysate was spun for 15 minutes at 10,000 rpm in a JA-14 rotor. The supernatant was then loaded onto a 50ml cyanogen bromide DNA- Sepharose column. Importantly the DNA attached to the sepharose beads was obtained from calf thymus and was platinated. This column was run in tandem with a 1ml DEAE column to remove DNA. Flow through was collected and run over column again at 3.5ml/min. Column was washed with 250ml of Buffer A containing 75mM potassium chloride. DNA-PK was eluted from column using 500ml of Buffer A containing 0.5M potassium chloride. 10ml fractions were collected and were individually assays for protein concentration and DNA-PK kinase activity using Bradford assays and Signatech DNA-PK Kinase assay respectively. Signatech DNA-PK Kinase assay was performed according to manufacturer's instructions. Due to the bump nature of the elution, peak DNA-PK fractions tended to be just over 1 column volume of the elution as would be expected. Peak fractions were pooled and dialyzed overnight in Buffer A containing 50mM potassium chloride.

The following day sample was removed from dialysis and loaded onto a 10ml heparin column equilibrated with 50mM potassium chloride Buffer A at a flow rate of 3ml/min. Column was then washed with 50ml of Buffer A containing 50mM potassium chloride at a flow rate of 3 ml/min. Protein was eluted off the column using a 100ml linear gradient starting at 50mM potassium chloride and ending at 0.6M potassium chloride both in Buffer A. 3ml fractions were collected and assayed for protein concentration and kinase activity. Unlike with the fractions assayed from the 50ml DNA column, the heparin fractions were assayed using our standard kinase assay method detailed below. Importantly Ku was complemented into the reactions and salmon sperm DNA (675ng) digested with EcoRI and BamHI was used to stimulate the kinase. Peak fractions tended to center around 350mM potassium chloride in the elution. During the analysis of this column, kinase activity was the major determinant of peak fractions because DNA-PK concentration is relatively low at this point of the purification. Peak fractions were dialyzed overnight in Buffer B containing 10mM potassium phosphate (pH 7.5), 50mM Tris (pH 7.5), 10% glycerol, and 1mM DTT.

The following day sample was removed from dialysis and loaded onto a 5ml hydroxyapatite column at a flow rate of 0.5ml/min. Column was washed with 25ml of Buffer B containing 10mM potassium phosphate. Protein was eluted from column using a 50ml linear gradient starting at 10mM potassium phosphate and ending at 500mM potassium phosphate in Buffer B. Fractions were collected in 0.75ml aliquots and assayed for protein concentration and kinase activity. Again the kinase activity was the major factor in determining peak fractions. This column is used to separate Ku from DNA-PKcs and thus fractions were assayed for kinase activity under conditions with and

without Ku complementation. Importantly, fractions which were most stimulated by the addition of exogenous Ku were determined to be peak fractions and were pooled. Sample was then dialyzed in a final hepes buffer containing 20mM Hepes (pH=7.0), 75mM potassium chloride, 10% glycerol, 0.005% Triton X-100, and 2mM DTT.

2.5 Electrophoretic mobility shift assays (EMSA)

In order to ensure that equal amounts of Ku were used in kinase assays, EMSAs were performed to determine the DNA binding activity of WT Ku and the mutants. EMSAs were performed in 20µl reactions containing 50mM sodium chloride, 10 mM Tris-HCl, 10mM magnesium chloride, 1mM DTT, and 300fmol of 5' radiolabeled double-strand 30mer DNA with varying amount of Ku. Protein-DNA complexes were separated on a 6% non-denaturing polyacrylamide gel and visualized using phosphorimager radiography.

2.6 SDS-PAGE and Western Blot Analysis

Protein samples were separated by SDS-PAGE using 8% Tris-glycine gel according to manufacturer's instructions (Invitrogen). Gels were either stained with Coomassie Blue or transferred to Immobilon-FL membranes (Millipore, Bedford, MA), probed and then visualized using chemiluminescent detection, and the LAS-3000 imaging system (FujiFilm) was used to document and quantify blots. Relevant information regarding antibodies used for western blot analysis are as follows: Ku70, NeoMarkers, catalog number MS-329-P1; Ku80, NeoMarkers, catalog number MS-285-P1; DNA-PKcs, Calbiochem, catalog number PC127.

2.7 DNA-PK kinase assays

Kinase assays were performed at 37º C in a final volume of 20µl containing 20 mM HEPES, pH 7.5, 8 mM $MgCl_2$, 1 mM DTT, 5 % glycerol, 125 µM ATP, [γ-^{32}P] ATP (0.5µCi), 42.4ng of DNA-PK, 500 µM p53 synthetic peptide and varying amounts of DNA. Reactions with 30bp and 60bp substrates were performed with 500fmol of DNA per reaction to a final concentration of 25nM while reactions with ~400bp and plasmid DNA were performed with 154fmol of DNA. Single stranded oligonucleotides were purchased from Integrated DNA Technology (IDT, Coralville, IA, USA). DNA sequences of oligonucleotides are shown in **Table 1**. 154fmol of ~400bp substrates and plasmid DNA were used in kinase assays. The ~400bp substrates were purchased as double stranded "gBlocks" Gene Fragments from IDT. The sense strand of these sequences is reported in **Table 2**. The single stranded overhangs are highlighted in red. The blunt ended substrate is 421bp. To generate the 5' overhangs the Blunt-ended gBlock was digested with EcoRI. The 3' overhangs were generated with KpnI digestion. G50 columns were used to separate the 397bp digested DNA from the terminal fragments. Kinase assays performed with plasmid DNA were completed with pcDNA3.1 (XhoI, BamHI, EcoRV, and KpnI) and pCAG-GFP (XbaI and EcoRI) digested with the indicated restriction enzyme (New England Biolabs). High fidelity versions of the restriction enzymes were used when applicable. The specific sequences recognized by the restriction enzymes are shown in **Table 3**. Arrows indicate location of cleavage. To ensure complete digestion reaction products were analyzed by native agarose gel electrophoresis and enzymes were heat inactivated prior to use in kinase assays. Kinase assay reactions were initiated with addition of ATP, incubated at 37ºC for 30 minutes and

terminated by addition of 20µl of 30 % acetic acid. Reactions products were spotted on P81 phosphocellulose cation exchange filter paper. The paper was washed three times, 5 minutes each, in 15 % acetic acid and allowed to dry. Samples were quantified by PhosphorImager® analysis using ImageQuant software (Molecular Dynamics).

Name	Annealed with	Overhangs	Sequence (5'-3')
Primer 590 Forward	pFastBac1-Ku80	--	GCATGCATGCCATATGGGAAGTGTGA
Primer 590 Reverse	pFastBac1-Ku80	--	TCACACTTCCCATATGGCATGCATGC
Primer 704 Forward	pFastBac1-Ku80	--	AAGTTTTAGTAGGGATCCFACAAACC
Primer 704 Reverse	pFastBac1-Ku80	--	GGTTTGTCGGATCCCTACTAAAACTT
30mer Top	30mer Bottom	None	CTGAAGGCAGTGTCACCTCTGTTGGAAGTG
30mer Bottom	30mer Top	None	CACTTCCAACAGAGGTGACACTGCCTTCAG
30mer 3' Top	30mer 3' Bottom	4 base 3'	CTGAAGGCAGTGTCACCTCTGTTGGAAGTGTGCA
30mer 3' Bottom	30mer 3' Top	4 base 3'	CACTTCCAACAGAGGTGACACTGCCTTCAGTGCA
30mer 5' Top	30mer 5' Bottom	4 Base 5'	TGCACTGAAGGCAGTGTCACCTCTGTTGGAAGTG
30mer 5' Bottom	30mer 5' Top	4 Base 5'	TGCACACTTCCAACAGAGGTGACACTGCCTTCAG
60mer Top	60mer Bottom	None	GTTAAGTATCTGCATCTTACTTGACGCATGCAATCGTCACGTGCTAGACTACTGGTCAAG
60mer Bottom	60mer Top	None	CTTGACCAGTAGTCTAGCACGTGACGATTGCATGCGTCAAGTAAGATGCAGATACTTAAC
60mer 3' Top	60mer 3' Bottom	4 Base 3'	TGCAGTTAAGTATCTGCATCTTACTTGACGCATGCAATCGTCACGTGCTAGACTACTGGTCAAG
60mer 3' Bottom	60mer 3' Top	4 Base 3'	ACCAGTAGTCTAGCACGTGACGATTGCATGCGTCAAGTAAGATGCAGATACTTAACTGCATGCA
60mer 5' Top	60mer 5' Bottom	4 Base 5'	TGCAGTTAAGTATCTGCATCTTACTTGACGCATGCAATCGTCACGTGCTAGACTACTGGTCAAG
60mer 5' Bottom	60mer 5' Top	4 Base 5'	TGCACTTGACCAGTAGTCTAGCACGTGACGATTGCATGCGTCAAGTAAGATGCAGATACTTAAC

Table 1. DNA Primers Used for PCR and Substrates Used in Kinase Assays.

Name	Length (bp)	Overhangs	Sequence (5'-3')
Blunt-ended gBlock	421	None	TGA CAT TGA ATT CTC TCC ACC ATA TGG AAT TTG ATT ATG TAA TAT GCG AAG AAT GTG GGA AAG AAT TTA TGG ATT CTT ATC TTA TGA ACC ACT TTG ATT TGC CAA CTT GTG ATA ACT GCA GAG ATG CTG ATG ATA AAC ACA AGC TTA TAA CCA AAA CAG AGG CAA AAC AAG AAT ATC TTC TGA AAG ACT GTG ATT TAG AAA AAA GAG AGC CAC CTC TTA AAT TTA TTG TGA AGA AGA ATC CAC ATC ATT CAC AAT GGG GTG ATA TGA AAC TCT ACT TAA AGT TAC AGA TTG TGA AGA GGT CTC TTG AAG TTT GGG GTA GTC AAG AAG CAT TAG AAG AAG CAA AGG AAG TCC GAC AGG AAA ACC GAG AAA AAA TGA AAC AGA AGA AAT TTT AGT AGG GAT CCG TGG ATT G AA TTC CCT CCG A
KpnI gBlock	397	4 base 3'	CTT TCC ACC ATA TGG AAT TTG ATT ATG TAA TAT GCG AAG AAT GTG GGA AAG AAT TTA TGG ATT CTT ATC TTA TGA ACC ACT TTG ATT TGC CAA CTT GTG ATA ACT GCA GAG ATG CTG ATGATA AAC ACA AGC TTA TAA CCA AAA CAG AGG CAA AAC AAG AAT ATC TTC TGA AAG ACT GTG ATT TAG AAA AAA GAG AGC CAC CTC TTA AAT TTA TTG TGA AGA AGA ATC CAC ATC ATT CAC AAT GGG GTG ATA TGA AAC TCT ACT TAA AGT TAC AGA TTG TGA AGA GGT CTC TTG AAG TTT GGG GTA GTC AAG AAG CAT TAG AAG AAG CAA AGG AAG TCC GAC AGG AAA ACC GAG AAA AAA TGA AAC AGA AGA AAT TTT AGT AGG GAT CCG TGG ACC GGT AC
EcoRI gBlock	397	4 base 5'	AATT CTC TCC ACC ATA TGG AAT TTG ATT ATG TAA TAT GCG AAG AAT GTG GGA AAG AAT TTA TGG ATT CTT ATC TTA TGA ACC ACT TTG ATT TGC CAA CTT GTG ATA ACT GCA GAG ATG CTG ATG ATA AAC ACA AGC TTA TAA CCA AAA CAG AGG CAA AAC AAG AAT ATC TTC TGA AAG ACT GTG ATT TAG AAA AAA GAG AGC CAC CTC TTA AAT TTA TTG TGA AGA AGA ATC CAC ATC ATT CAC AAT GGG GTG ATA TGA AAC TCT ACT TAA AGT TAC AGA TTG TGA AGA GGT CTC TTG AAG TTT GGG GTA GTC AAG AAG CAT TAG AAG AAG CAA AGG AAG TCC GAC AGG AAA ACC GAG AAA AAA TGA AAC AGA AGA AAT TTT AGT AGG GAT CCG

			TGG ATT G

Table 2. ~400 base pair DNA Substrate Sequences. Top strands of substrates are shown. Reactions were completed with their respective complements. Non-duplexed bases are shown in red.

Enzyme	Sequence and Site of Cuts
EcoRV	5′...GAT▼ATC...3′ 3′...CTA▲TAG...5′
KpnI	5′...GGTAC▼C...3′ 3′...C▲CATGG...5′
EcoRI	5′...G▼AATTC...3′ 3′...CTTAA▲G...5′
XhoI	5′...C▼TCGAG...3′ 3′...GAGCT▲C...5′
BamHI	5′...G▼GATCC...3′ 3′...CCTAG▲G...5′
XbaI	5′...T▼CTAGA...3′ 3′...AGATC▲T...5′

Table 3. Sequences Recognized by Restriction Enzymes Used. Arrows indicate sites of cleavage by indicated restriction enzyme.

We and others have previously reported that the C-terminus of Ku80 is dispensible for Ku/DNA binding. In order to control for slight variations in binding activity which may occur between protein purifications, equal amounts of active Ku was used in each reaction, measured as the amount necessary to bind 50fmol of DNA. Titration experiments were performed and demonstrate that these concentrations of Ku and DNA were in excess and DNA-PKcs is the limiting factor.

2.8 DNA-PK DNA Binding/Recruitment Assay

2pmol of 5' biotinylated 30mer top annealed to 30mer bottom DNA was bound to streptavidin coated wells of a 96 well plate overnight at 4^{o}C in 100µl of Hepes Buffer. Wells were then blocked with 200µl of 2% BSA in TBS/Tween for 1 hour at room temperature. 1pmol of WT and Ku mutants was added to DNA coated wells with 424ng of DNA-PKcs in a total volume of 100µl of Hepes Buffer. Protein complexes were allowed to bind DNA at room temperature for 1 hour. The supernatant was removed and again wells were blocked with 200µl 2% BSA in TBS/Tween for 1 hour at room temperature to avoid non-specific antibody binding. 100µl 2% BSA in TBS/Tween and a 1:750 diluted primary antibody specific for DNA-PKcs (Calbiochem Catalog number PC127) was added to each well and incubated at room temperature for 1 hour. Then wells were washed with 200µl of 2% BSA in TBS/Tween three times for 5 minutes each. 100µl of 1:2500 diluted goat anti-rabbit HRP secondary antibody in 2% BSA in TBS/Tween was added to each well and incubated for 1 hour at room temperature. Wells were then washed again three times with 200µl of 2% BSA in TBS/Tween for five minutes in each wash. 100µl of 1-Step TMB ELISA reagent was added to each well and plate was immediately read at 37^{o}C at wavelength of 370nm every 5 minutes for 60

minutes total. Average change in optical density at 370nm per minute is reported at 15 minutes. Each condition is reported as an average of triplicates.

2.9 Host Cell Reactivation Assay

Linearized pCAG-GFP was used in a host cell reactivation assay as pervious reported with only minor differences[63]. The same linearized plasmid was used for kinase activation and was linearized by indicated restriction enzyme digestion as recommended by manufacturer (New England Biolabs). Covalently closed circular pCAG-Red was used as a transfection control in all experiments. All transfections were completed using 2ug of indicated plasmid DNA and at least 100 cells were counted in each condition. Images were captured and evaluated 48 hours after transfection as previously reported[63]. Results reported are the average of 6 independent transfections.

2.10 Statistical Analysis

All statistical analysis was completed using Sigma Plot 11.0 software using a Student's t- test. The $P<0.05$ was used as a significance threshold for all tests unless otherwise indicated. All statistical analyzes of kinase assays used an n=3.

2.11 Protein Structure Prediction

Disordered protein regions and sites of protein/protein interactions were predicted using PONDR software (http://www.pondr.com)[64] with VSL1 and VLXT algorithms using Ku80 accession number AAA36154.1 GI:307094.

3. Different Structural Regions of the Ku80 C-terminus Influence DNA-PKcs Activity Depending on the Structure of the DNA Substrate

3.1 Introduction

Due to its importance in DNA repair and the DDR we sought to better understand the basic mechanism of DNA-PK activation. The two primary components contributing to DNA-PK activation are protein/protein and protein/DNA interactions. While obviously not mutually exclusive, previous studies have investigated them as such. For several years the protein/protein interactions have been under investigation and C-terminal region of Ku80 has been implicated in DNA-PK stimulation[21,45]. Strong evidence for stimulatory interactions between the extreme C-terminus of Ku80 (amino acids 720-732) and DNA-PKcs has been reported in vitro and in vivo[43,46]. Further, this interaction was reported to be indispensable for DNA-PK activity. More recent investigations, however, have shown that the C-terminus of Ku80 plays a stimulatory role in DNA-PK activation but these interactions are not essential for activation[47]. In fact, in the absence of the Ku80 amino acids 599-732, DNA-PKcs still retains 56% of the WT levels of kinase activity[47]. While Ku80 C-terminus/DNA-PKcs interactions may stimulate DNA-PK activity, we have reported that they are not sufficient for activation thus other Ku70/80/DNA-PKcs interactions independent of the C-terminus of Ku80 must contribute to activation[62]. In accordance with this observation, structural analysis reveal an extensive interface between the core of the Ku70/80 dimer and DNA-PKcs[48]. It is likely that these interactions as well as those involving the C-terminus of Ku80 are necessary for maximum activity of DNA-PK. The protein/DNA interactions are equally important in DNA-PK activation as DNA is an essential substrate required for activation.

Cross-linking experiments suggest that DNA-PKcs binds primarly to terminal 10 bp of DNA while the Ku dimer interacts with the adjoining 14 bases[65]. Our group was the first to report that relatively small changes in the structure of the DNA substrate result in drastic differences in complex formation and DNA-PK activation [52,53]. Interestingly, structural data has revealed differences in DNA-PKcs/DNA-PKcs interactions which occur across the synapse of the DSB when these proteins are bound to different DNA termini[48].

It should be noted that recently an investigation has revealed that DNA-PKcs can be activated in the absence of both DNA and Ku when it is bound by an N-terminal antibody[66]. However, this does not change the model of DNA-PK activation being dependent on both protein/protein interactions and protein/DNA interactions but instead highlights the importance of the N-terminus of DNA-PKcs in activation. This interpretation shared by the author suggests that the antibody provides N-terminal constraint on the catalytic subunit which likely occurs via interactions with Ku and DNA *in vivo* and *in vitro*[66].

In order to better understand how protein/protein and protein/DNA interactions influence DNA-PK activation, we created and purified Ku80 C-terminal truncation mutants based on structural data and predicted disordered protein regions. We found that different structural regions of the Ku80 C-terminus are involved in stimulating DNA-PK activity and are dependent on the structure of the DNA substrate to which the complex is bound. Importantly, the Ku80 C-terminus was not found to influence DNA-PKcs recruitment/binding to DNA suggesting that it influences catalysis and not complex formation.

3.2 Results

Despite the fact that structural data exists for Ku70/80 core, the Ku80 C- terminal region, and DNA-PKcs, clear structure/function relationships involved in the DNA-PK holoenzyme formation are still undefined. Accordingly, we designed a series of Ku80 C-terminal truncation constructs based on known and predicted structural features. PONDR, a protein software prediction tool that predicts disordered protein regions, identified two highly disordered regions exist in the C-terminal region of Ku80 between amino acids 550-594 and 704-732 (**Figure 3**). Further these regions are both predicted to be sites of protein/protein interactions (**Figure 3** see black bars). Interestingly the disordered region at the extreme C-terminus has already been implicated in protein-protein interactions by Jackson and colleagues[43,46].

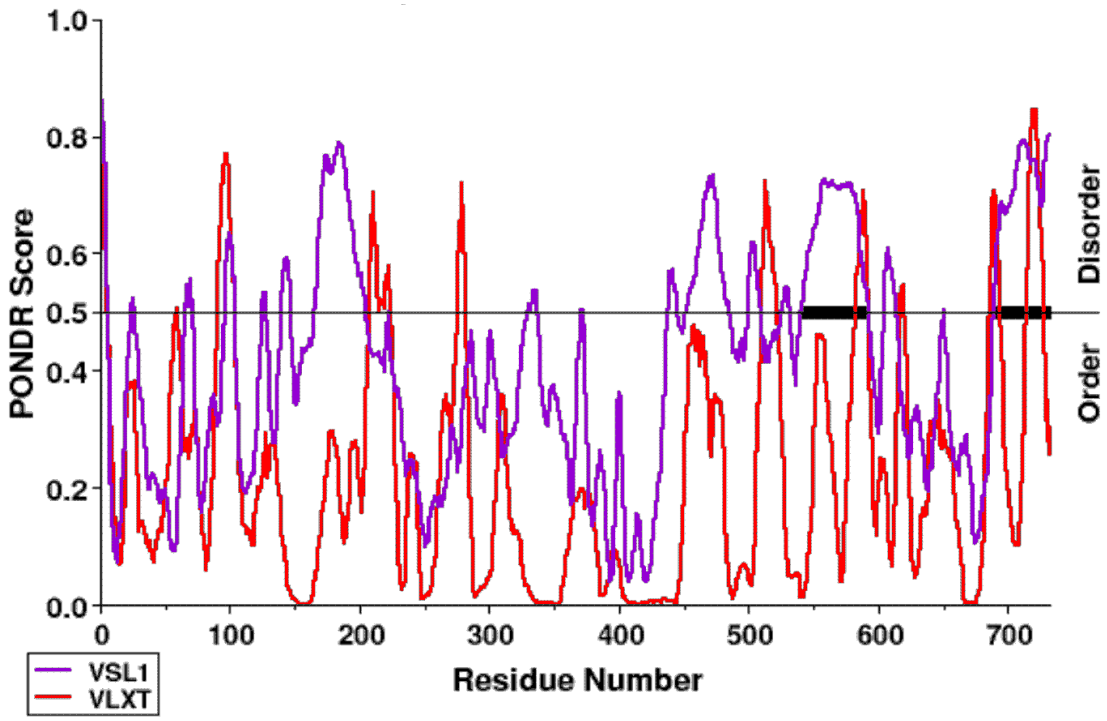


Figure 3. PONDR Prediction of Ku80 Structure. Amino acid number is shown on the X-axis against structure score on the Y-axis. Threshold defining disordered and ordered prediction is indicated by horizontal line on graph. Dark black bars along the threshold line indicate areas predicted to be involved in protein/protein interactions.

To test the influence of these regions on DNA-PK activation, Ku80 truncation constructs were created after each of the following: the alpha helical bundle (amino acid 704), the disordered linker domain (amino acid 594), and the core DNA binding/Ku70 dimerization domain (amino acid 550). The mutants are named according to the domains they contain. The C-L-H mutant contains the core, linker, and helix domains. The C-L mutant contains the core and linker region. Finally the Core mutant contains the minimal Ku70 dimerization/DNA binding domain. A linear representation of the constructs is shown in **Figure 4A**. These constructs along with wild type Ku80 were co-expressed with wild type Ku70 and purified from SF9 insect cells (**Figure 4B**). A structural model of the Ku constructs is shown in **Figure 4C**. DNA-PKcs was purified from HeLa cells as described in the materials and methods. I was able to purify DNA-PKcs to apparent homogeneity as shown in lane 9 of **Figure 5A**. Importantly, western blots were preformed with highly sensitive antibodies to ensure that no detectible endogenous wild type Ku is present in this purification (**Figure 5B**). It should be noted that the hydroxyapatite column was the crucial step in separation of endogenous Ku from DNA-PKcs as indicated in the western blot. This step is the main difference between this purification method and those previously published from our lab.

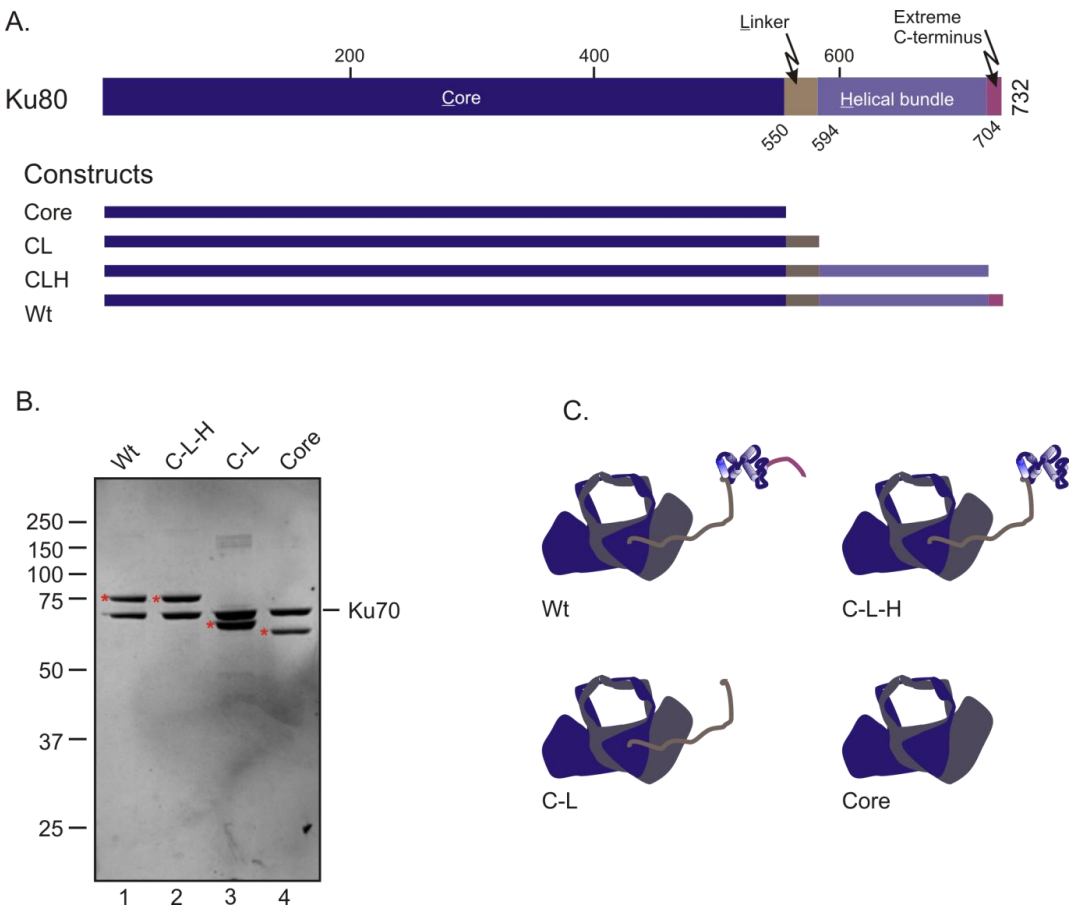


Figure 4. Ku80 Protein Structure Prediction, Mutant Construction, and Protein Purification. A. Linear representation of Ku80 structural domains and mutant construction B. Coomassie stained gel of purified Ku and DNA-PKcs proteins used in this study. All Ku80 mutants were purified with WT Ku70 which remains consistent in lane 1-4. Ku80 mutants are indicated along the top of the gel. Ku80 bands are denoted with red asterisks. C. Model of WT and Ku mutant constructs used in this study. Ku80 depicted in dark blue and Ku70 is depicted in grey. The structural regions of the Ku80 C-terminus are depicted in their corresponding color from A.

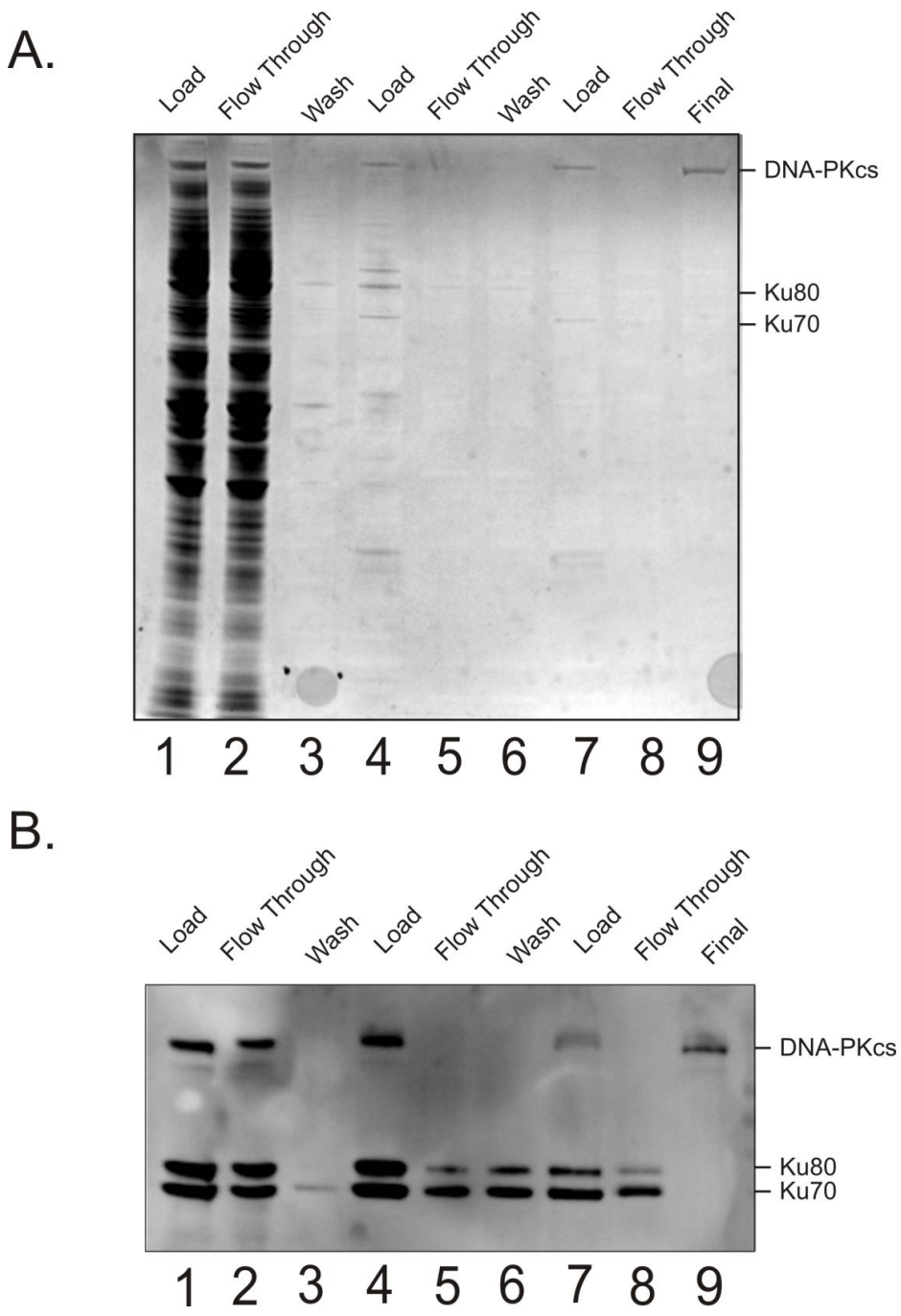


Figure 5. DNA-PKcs Purification. A. Coomassie stained gel of samples from purification of DNA-PKcs. Lanes 1-3 are samples taken from during the 50ml DS DNA column step of the purification. Lanes 4-6 are samples taken from the 10ml heparin column step of the purification. Lanes 7-8 are samples taken from the 5ml hydroxyapatite column step of the purification. Lane 9 is the final pool of the purification. B. Western blot of the same samples from 5A. Anti-bodies specific for Ku70, Ku80, and DNA-PKcs were used to visualize corresponding proteins. Note how Ku70 and Ku80 are detectible in the lane 8 but not in lane 9. This is indicative of highly purified DNA-PKcs.

We have previously reported that DNA-PK activation is influenced by the structure of the DNA substrate to which the holoenzyme is bound [52,53]. Thus we tested if distinct Ku80 structural regions contribute to differential activation of DNA-PKcs, as a function of DNA substrate. Full WT levels of activation using a 30 base pair fully duplex DNA required the disordered linker and the helical bundle, but occurs independent of the extreme C-terminus as no activation was observed with Ku80 C-L and Core mutants while complete rescue, defined as activation levels observed with WT Ku, was observed with the C-L-H mutant (**Figure 6A**). While the C-L-H mutant, containing both the helical bundle and the disordered linker, completely restores activity on the 30bp substrates with 4 base 5' and 3' overhangs, the C-L mutant containing the disordered linker region (amino acids 551-594) can also stimulate DNA-PKcs activity above background on these substrates (**Figure 6B and C**). In contrast, this C-L mutant did not stimulate DNA-PKcs activity above background on the 30 base pair blunt substrate (Figure 6A) suggesting that the stimulatory effect of the single stranded overhangs (both 3' and 5') partially compensate for the lack of the helical bundle. The Core mutant did not stimulate kinase activity significantly above the no Ku control on 30bp substrates regardless of the structure of the termini **(Figure 6A-C)**. Results from the 30bp substrates indicate that C-L mutant is able to stimulate DNA-PK activity when 5' or 3' overhangs are present but not on blunt-ends (**Figure 6**). This is the first demonstration of the DNA substrate influencing structure/function relationships involved in DNA-PK activation by the Ku80 C-terminus.

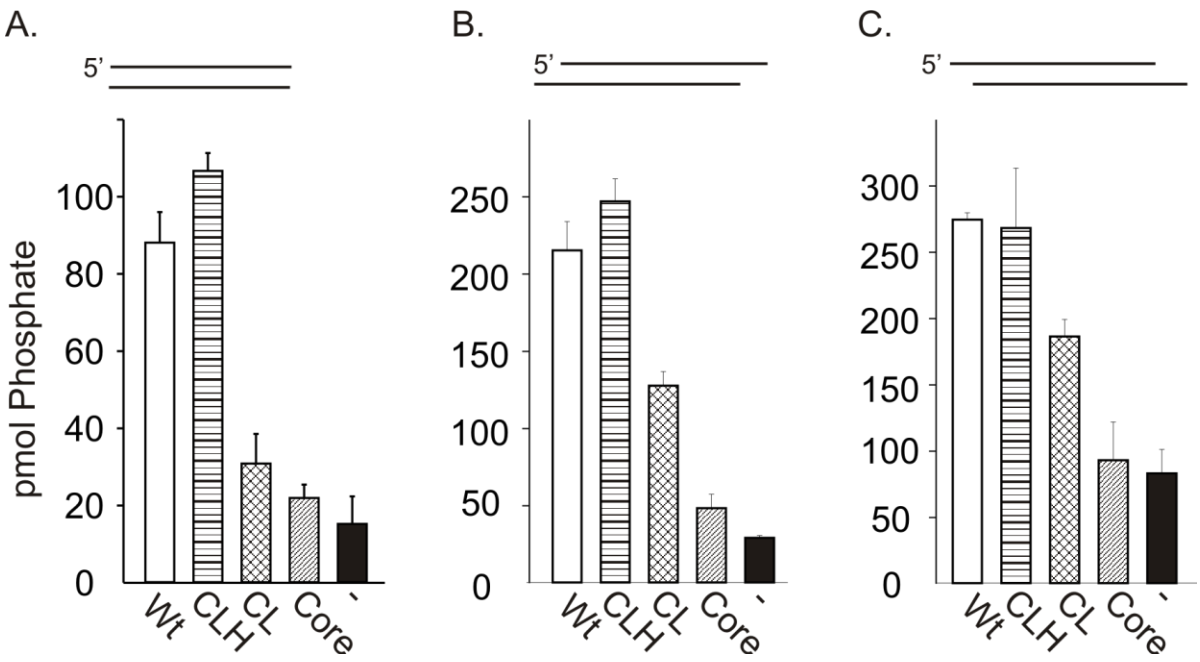


Figure 6. Stimulation of DNA-PKcs Activity By Different Structural Regions of the Ku80 C-terminus on 30bp DNA Substrates. The same 30bp DNA sequence with blunt end termini (A), 4 base 3' overhangs (B), or 4 base 5' overhangs (C) was used to stimulate DNA-PK as indicated above the bar graphs. Results are reported as pmol of Phosphate transferred. Ku80 constructs was represented as follows: WT bars are white, C-L-H bars contain horizontal lines, C-L bars contain crosshatching, core bars contain diagonal lines, and no added Ku (-) bars are solid. Averages of 3 samples are shown. Error bars indicate 1 standard deviation from the mean.

Interestingly the influence of the C-terminus on activation was distinct when reactions performed using 60bp substrates. Intermediate stimulation was observed from the Core mutant using the blunt ended and 4 base 3' overhang substrates while wild type levels of kinase activity were observed using the 60bp substrate with 4 base 5' overhangs (**Figure 7 A-C**). This suggests that the C-terminus plays a less stimulatory role in the presence of 60bp substrates and is completely dispensable when the complex is bound to 60bp substrates with single stranded 5' overhangs. The fact that the Core mutant was unable to stimulate the kinase above background on any of 30bp substrates and yet can cause at least intermediate stimulation on all of the 60bp substrates is intriguing. Since the Core mutant is capable of wild type levels of stimulation and the addition of the other disordered domains and helical bundle do not further stimulate the kinase, it is likely that these regions do not participate in stimulatory interactions on this DNA substrate. Such stimulatory interaction likely do exist however, between the Core mutant and DNA-PKcs as the no Ku control has drastically less kinase activity compared to the Core mutant. The level of activation by the C-L mutant containing the disordered linker region was statistically indistinguishable from the WT construct ($p = 0.139$) in the presence of 60bp blunt-ended substrate while only intermediate activation was observed with this mutant in reactions performed with the 60bp substrate with 4 base 3' overhangs (**Figure 7 A-C**). The average activation by the C-L-H mutant is greater than the average activation by the WT on the 60bp blunt-ended substrate however, these difference were not shown to be statistically significant. Together these data indicate that the helical bundle and disordered linker region are necessary for wild type levels of DNA-PK stimulation when the complex is bound to certain DNA structures while being dispensable on others.

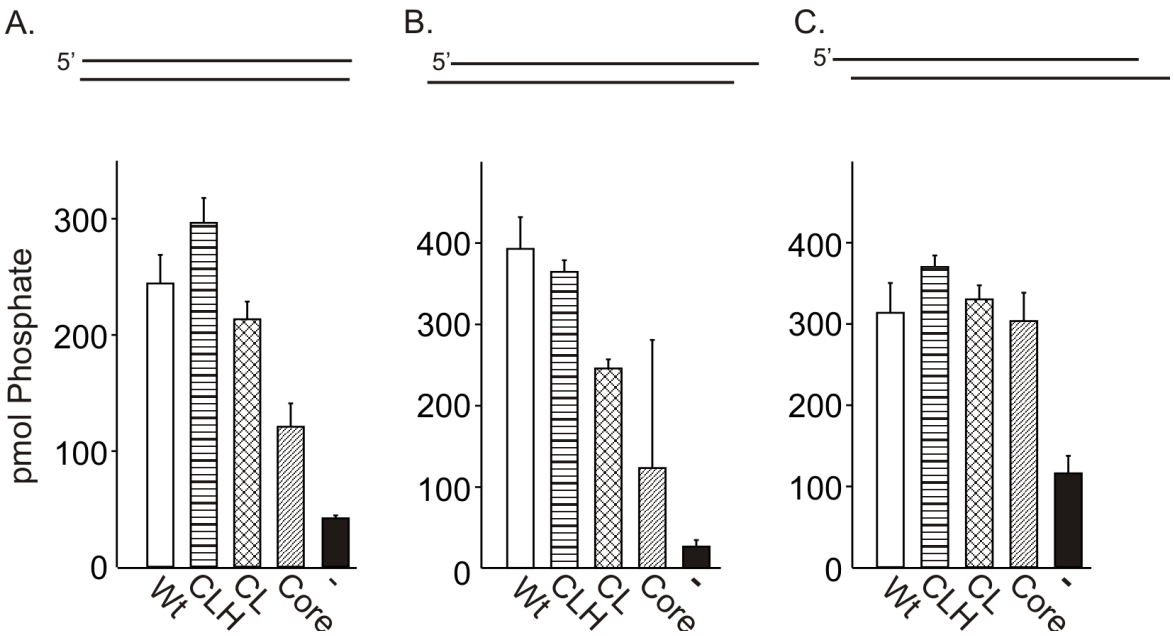


Figure 7. Stimulation of DNA-PKcs Activity By Different Structural Regions of the Ku80 C-terminus on 60bp DNA Substrates. The same 60bp DNA sequence with blunt end termini (A), 4 base 3' overhangs (B), or 4 base 5' overhangs (C) was used to stimulate DNA-PK as indicated above the bar graphs. Results are reported as pmol of Phosphate transferred. Ku80 constructs was represented as follows: WT bars are white, C-L-H bars contain horizontal lines, C-L bars contain crosshatching, core bars contain diagonal lines, and no added Ku (-) bars are solid. Averages of 3 samples are shown. Error bars indicate 1 standard deviation from the mean.

One possible mechanism to explain our observations of differential DNA-PK activation is that these result from differential DNA-PKcs-DNA binding. The role of the Ku80 C-terminal region in recruiting DNA-PKcs to the site of a DSB has been debated. While some evidence shows that the extreme C-terminus of Ku80 is necessary and sufficient for DNA-PKcs to bind to DNA, others show that the C-terminus has only a modest impact on DNA-PKcs-DNA binding [46,47]. We observed substantial differences concerning the influence of the C-terminal structural regions of Ku80 in DNA-PKcs activation between 30 base pair substrates and 60 base pair substrates. Accordingly we asked if these differences were due to alterations in binding or catalysis. The DNA-dependence of the kinase suggests that kinase stimulation observed on the 60 base pair substrates must result from efficient DNA-PKcs-DNA binding. Therefore, an ELISA based DNA-PKcs-DNA binding assay was employed using the 30 base pair blunt-end substrate used in **Figure 6A** which did not support activation in the presence of the Core or C-L constructs. This DNA was bound to the well and incubated with purified Ku and DNA-PKcs. Unbound protein was removed and DNA bound complexes were detected using a DNA-PKcs specific antibody. Results illustrate that the presence of Ku increases DNA-DNA-PKcs binding up to 4 fold (**Figure 8)**. Each of the mutants and WT Ku were able to effectively recruit DNA-PKcs to DNA, and no statistically significant differences were observed between the mutants and WT. The C-L and Core mutants did not stimulate DNA-PKcs kinase activity above background in the presence of this substrate (**Figure 6**); however, these mutants were able to recruit DNA-PKcs effectively (**Figure 8**). These data suggest the differential DNA-PKcs stimulation result from differences in catalysis and not differences in DNA-PKcs-DNA binding efficiency. Further the 30 base

pair duplex is the most constrictive substrate in terms of lateral movement of the protein complex on DNA which in turn would also be expected to lead to the highest level of molecular crowding events. Such events would cause the most restriction to DNA-PK-DNA complex formation and retainment of the complex on DNA. Despite this, Core mutant was able to recruit and retain DNA-PKcs on the 30 base duplex DNA at levels that were statistically indistinguishable from WT. This provides strong evidence that the entire C-terminus of Ku80 is dispensable for Ku-dependent DNA-PKcs-DNA binding.

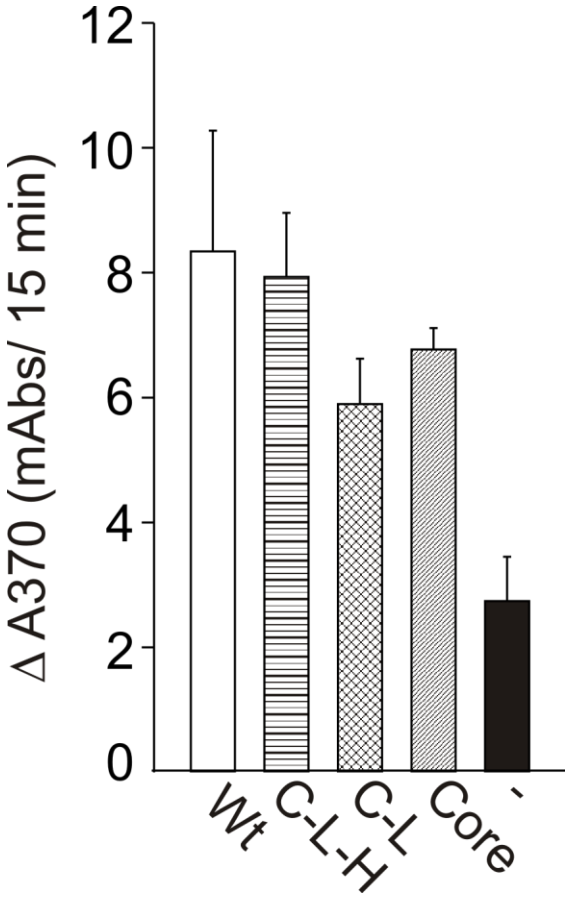


Figure 8. The Ku80 C-terminus is Dispensable for DNA-PKcs-DNA Binding. Results of ELISA are reported as the rate of change in absorbance at OD_{370nm} in 15 minutes. Ku80 constructs was represented as follows: WT bars are white, C-L-H bars contain horizontal lines, C-L bars contain crosshatching, core bars contain diagonal lines, and no added Ku (-) bars are solid. Averages of 3 samples are shown. Error bars indicate 1 standard deviation from the mean. There are no statistically significant differences among the Ku mutants or WT. Each of the conditions are statistically significant from the No Ku control.

After demonstrating that the differences in activation between the 30bp substrates and the 60bp substrates are due to altered catalysis and not differences in DNA-PKcs recruitment to DNA, we further interrogated DNA influence on DNA-PK activation. The observed effect of length led us to test the influence of significantly longer DNA substrates on activation relationships involving the C-terminus of Ku80. Overall, the affect or the Ku80 C-terminus on DNA-PK activation was similar in reactions performed with blunt ended and 3' overhang ~400bp substrates as their respective 60bp counterparts (**Figure 9A and B**). The C-L-H mutant was able to fully activate the kinase to wild type levels while the C-L mutant was able to partially stimulate the kinase. Additionally the Core mutant, which completely lacks the C-terminal region, can stimulate the kinase above background and these differences are statistically significant. In contrast, the ~400bp substrate containing 4 base 5' overhangs alters the influence of the C-terminus on DNA-PK activation in such a way that is distinct from both 30bp and 60bp substrates containing 4 base 5' overhangs (compare **Figure 9C to 6C and 7C**). Here the C-L-H mutant is sufficient to activate DNA-PKcs to wild type levels, but in the absence of the helical bundle (see C-L mutant) no activation is observed above background (**Figure 9C**). Overall these data illustrate the importance of C-terminus of Ku80 on DNA-PKcs activation as distinct structurally defined protein regions differentially influence kinase activation depending of the structure of the DNA substrate to which they are bound.

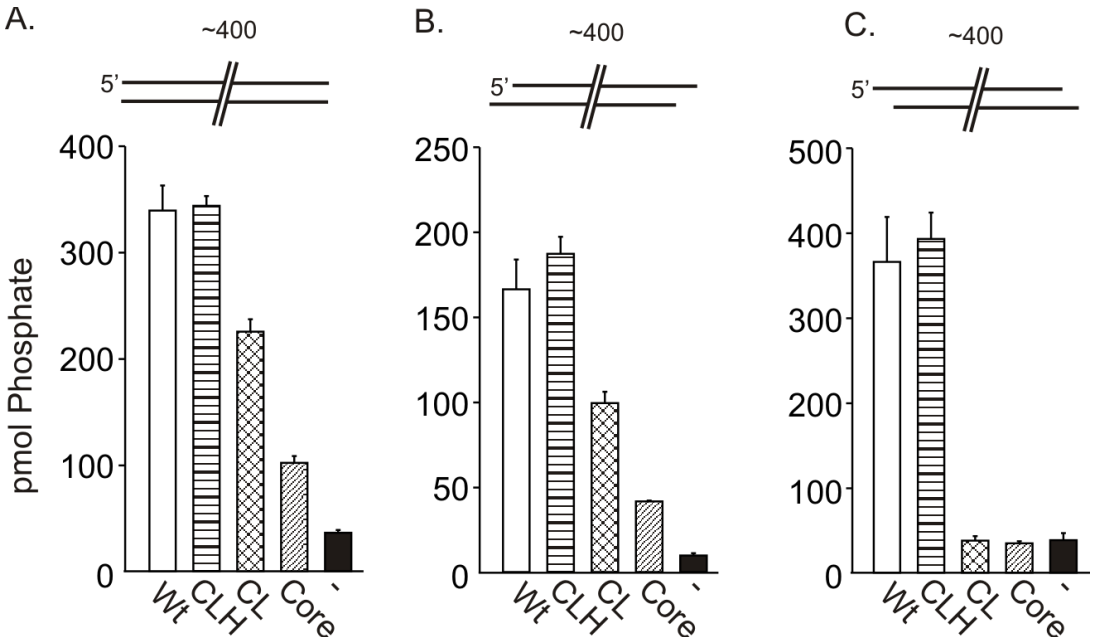


Figure 9. Stimulation of DNA-PKcs Activity By Different Structural Regions of the Ku80 C-terminus on ~400bp DNA Substrates. The ~400bp DNA sequence with blunt end termini (A), 4 base 3' overhangs (B), or 4 base 5' overhangs (C) was used to stimulate DNA-PK as indicated above the bar graphs. Results are reported as pmol of Phosphate transferred. Ku80 constructs was represented as follows: WT bars are white, C-L-H bars contain horizontal lines, C-L bars contain crosshatching, core bars contain diagonal lines, and no added Ku (-) bars are solid. Averages of 3 samples are shown. Error bars indicate 1 standard deviation from the mean.

Substantial differences in activation were observed between substrates of various lengths (**Figure 6, 7, and 9**). In order to evaluate the longest DNA substrates available while maintaining a homogenous preparation of substrates, plasmid DNA was linearized with various restriction enzymes yielding blunt ends, 4 base 3' overhangs and 4 base 5' overhangs and used in kinase assays as before (**Table 3**). Additionally, using these substrates to test DNA-PKcs activation was advantageous because it is possible assess their DNA repair *in vivo.* Surprisingly the influence of the plasmid DNA was more similar to that of the 30bp substrates than the 60bp substrates. Namely the helical bundle (amino acids 595-704) is necessary for wild type levels of kinase stimulation when substrates contained 4 base 3' or 5' overhangs while the presence of the disordered linker (amino acids 551-594) yields intermediate stimulation (**Figure 10 A-C**). Two important differences were observed between the activation on linearized plasmid DNA and the ~400bp substrates. First the C-L mutant can stimulate the DNA-PKcs on 5' overhang plasmid DNA but not on the 400bp 5' overhang (Compare **Figure 9C to Figure 10C**). Second, the extreme C-terminus (amino acids 705-732) stimulates kinase activity on blunt-ended linearized plasmid DNA (**Figure 10A**) but no difference is observed on any other substrate tested including the blunt ended 421bp substrate (**Figures 6, 7, and 9**). This is unexpected as the extreme C-terminus (amino acids 705-732) contains the previously defined "PIKK interaction motif" which has been reported to be necessary and sufficient for Ku80/DNA-PKcs interactions[43,46]. In contrast, our enzymatic activity data suggest that other interactions exist between Ku80 and DNA-PKcs as evidenced by Ku dependent activation in the absence of the extreme C-terminus on all substrates.

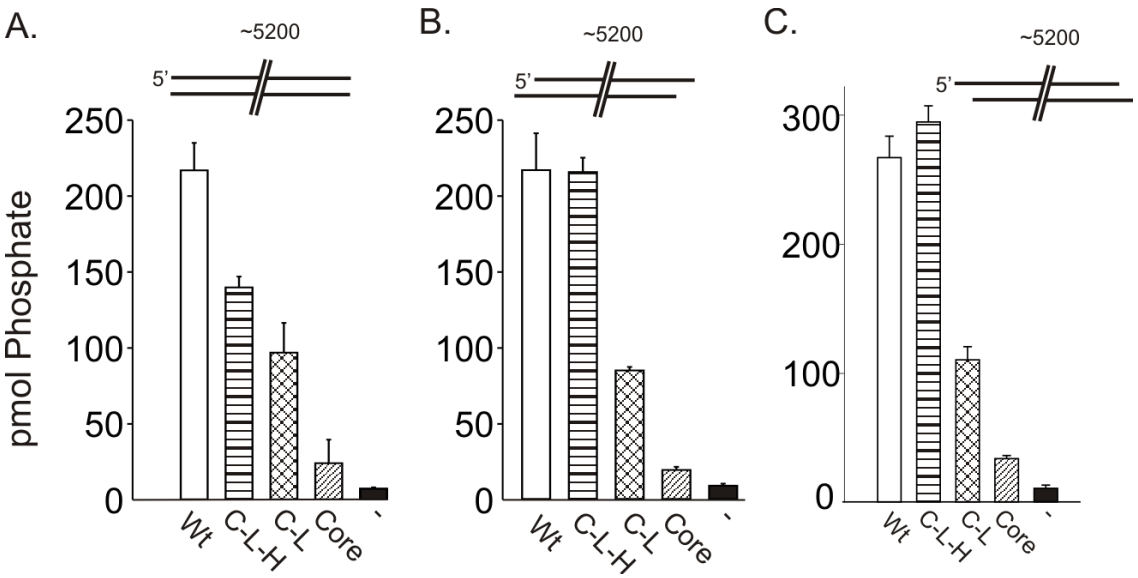


Figure 10. Stimulation of DNA-PKcs Activity By Different Structural Regions of the Ku80 C-terminus on Linearized Plasmid DNA Substrates. The pcDNA3.1 plasmid with blunt end termini (A), 4 base 3' overhangs (B), or 4 base 5' overhangs (C) was used to stimulate DNA-PK as indicated above the bar graphs. Results are reported as pmol of Phosphate transferred. Ku80 constructs was represented as follows: WT bars are white, C-L-H bars contain horizontal lines, C-L bars contain crosshatching, core bars contain diagonal lines, and no added Ku (-) bars are solid. Averages of 3 samples are shown. Error bars indicate 1 standard deviation from the mean.

Overall, the data suggest that Ku80/DNA-PKcs interactions necessary for DNA-PK activation cannot be defined to any one structural region of the Ku80 C-terminus regardless of DNA substrate. Instead different structural regions are required for kinase stimulation depending on the structure of the DNA substrate to which the complex is bound. A summary of the stimulatory effects of each of the structural Ku80 mutants normalized to WT is shown in **Table 1**.

		WT	C-L-H	C-L	Core
30bp	Blunt	++++	++++	-	-
	5'	++++	++++	++	-
	3'	++++	++++	++	-
60bp	Blunt	++++	++++	++++	++
	5'	++++	++++	++++	++++
	3'	++++	++++	++	++
~400bp	Blunt	++++	++++	++	+
	5'	++++	++++	-	-
	3'	++++	++++	++	+
Plasmid	Blunt	++++	+++	++	+
	5'	++++	++++	++	+
	3'	++++	++++	++	+

Table 4. Summary of kinase assay results from Figures 6, 7, 9, and 10. Each plus sign indicates 25% of kinase activity for each DNA substrate normalized to WT

3.3 Discussion

These data demonstrate that the structurally defined portions of the Ku80 C-terminus play differing roles in stimulation of DNA-PK when Ku is bound to different DNA substrates. On 30bp DNA substrates the helical bundle was required for wild type levels of activation and when these substrates had overhangs, the linker region was also capable of moderate stimulation of kinase activity. These results are likely due to distinct contributions from the overhangs and not due to length of the overhangs.

Perhaps the most interesting results from this study were those obtained using the 60bp substrates which were clearly atypical. The activation relationships between the Ku mutants and DNA-PKcs remained largely unchanged between the 30bp and 60 bp substrates containing 3' overhangs (**Table 4**). The activation relationships concerning the substrates containing blunt ends and 5' overhangs, on the other hand, changed drastically between the 30bp and 60bp. On the 30bp blunt substrates, no significant activation was observed with the Core or C-L mutants. On the 60bp blunt substrates however, the Core mutant stimulates DNA-PKcs to approximately 50% of WT levels while the C-L is capable of stimulating the kinase to full WT levels. Of the many substrates tested, the 60bp blunt end substrate was the only one in which the C-L mutant was the minimal construct capable of supporting full kinase activity. 60bp substrates with 5' overhangs supported full levels of activation in the presence of just the Core domain suggesting that the entire C-terminus of Ku80 is dispensable for DNA-PK stimulation when Ku is bound to these specific cofactors. This does not necessarily mean that there are no interactions between the various structural regions of the C-terminus and DNA-PKcs when these regions are present. Such interactions could very well occur however the data presented

here show that these potential interactions are not influencing kinase activity. Further analysis is required to distinguish whether or not these non-stimulator interactions exist under these conditions. In this study only the p53 substrate was used to assess kinase activity. It is possible that interactions exist that do not effect phosphorylation of this target but do alter activity toward another target.

Generally, the observed differences between the 30bp and 60bp substrates must be attributed to substrate length. The fact that any difference exists between these lengths is surprising because the 30bp substrates are longer than the minimal length required for DNA-PK formation as discovered via covalent cross-linking methods by Yoo and Dynan[38]. We have also shown that DNA-PK activation is supported on even shorter lengths of DNA including 20bp substrates albeit at a lower rate. The decreased necessity of the C-terminus of Ku80 in DNA-PK activation on 60bp substrates may result because Ku has more room for lateral movement along the DNA. This increased movement may allow for stimulatory interactions between the Core domain of the Ku and DNA-PKcs which are restricted on shorter substrates. An inconsistency with this model is that the C-terminus while dispensable on 60bp substrates with 5' overhangs is necessary for activation on the 400bp and plasmid substrates. Instead is seems that the 60 base pair length is particularly stimulatory. It is known that other complexes including MRN bind to DSBs near the terminus[67]. There is potential that these other protein complexes could crowd the terminus where Ku and DNA-PKcs bind. Therefore the 30bp substrate with restrictive movement may be more representative of the *in vivo* substrates than the 60bp substrates.

The results from the 400bp substrates were similar to the 30bp substrates in that the C-L-H mutant was the minimal mutant capable of WT levels of activation. Thus the extreme C-terminus is dispensable for kinase stimulation in the presence of these substrates. 5' overhangs on these substrates yielded unique results in that no detectible stimulation occurs with the Core or C-L mutants. These mutants were capable of at least moderate stimulation in the presence of ~400bp substrates with 3' overhangs and blunt ends. Typically the substrates with 5' overhangs are more stimulatory than substrates with blunt ends or 3' overhangs which makes the absence of activation with the Core and C-L mutants all the more interesting.

Kinase assays using plasmid DNA were used to further interrogate the effects of substrate length on DNA-PK stimulation. Due to Ku's terminal DNA binding activity and non-cooperativity of binding multiple Ku molecules on one DNA molecule, it seems logical that at a certain point increasing DNA length will not alter Ku's effect on DNA-PKcs. The differences observed between the ~400bp and plasmid substrates suggest that this threshold length is beyond 400bp. This is much longer than anyone in the field could have predicted because Ku binds to approximately 20bp of DNA. Evidence suggesting that the complex can differentiate between 400bp substrates and plasmid DNA is quite remarkable. The only other possibility to explain these observed differences is DNA methylation. The 400bp substrates are produced via PCR amplification while the plasmid DNA substrates are purified from *E. coli*. Thus the plasmid DNA is methylated while the 400bp substrates are not. Considering what we know about Ku/DNA interactions it is more likely that methylation is contributing to the differences in DNA-PK activation than DNA length. This a model would show methylated DNA with blunt

ends would necessitate the extreme C-terminus for maximum activation while the stimulatory activity of the DNA methylation allows for activation with the Core and C-L mutants on 5' substrates.

The fact that the extreme C-terminus is uniquely required for maximum kinase activity on the linearized blunt-end plasmid is interesting. During DSB formation, there are a number of ways to get overhang bases on either 3' or 5' strands, however there are fewer ways to get blunt ended DNA termini. One such way is during DNA replication when the polymerase encounters a nick in the DNA backbone. These events are further distinctive in that they yield one ended DSBs. Thus the NHEJ pathway would be of little use to the cell in this instance as there is not another double stranded terminus for the blunted DNA to be joined. Replication occurs during S-phase which is also when the majority of homologous recombination occurs[20]. Thus perhaps this unique role of the extreme C-terminus of Ku80 in activating DNA-PKcs on blunt ended DNA allows for the discrimination of these potentially one-ended DSBs that cannot be repaired via NHEJ. This does not mean that DNA-PK dependent signaling is not important for the repair of these breaks however. Evidence suggests that in fact DNA-PK activity, in particular autophosphorylation, plays an important role in promoting homologous recombination and/or NHEJ[20]. This is in agreement with our observations that DNA-PK signaling still occurs on blunt-ended DNA substrates (**Figure 10A**).

The data presented in this chapter demonstrate how the activation of DNA-PKcs by the C-terminus of Ku 80 is dependent of DNA substrate structure. Clearly substrate length, presence of overhangs, and orientation of overhangs dictate which region of the

Ku80 C-terminus is required for DNA-PKcs stimulation. Evidence is also presented that suggests these differences are due to altered catalysis and not differential recruitment.

4. Preferential DNA-PK Activation by Terminal Pyrimidines Leads to Increased NHEJ

4.1 Introduction

We have previously reported that DNA-PK activation shows preference for certain terminal DNA sequences. During these investigations it was determined that in the presence of DNA substrates with 30 base pairs or less, DNA-PK is preferentially activated by duplex ends containing a 3' pyrimidine-rich and 5' purine-rich strands[52]. The role of single-strand overhangs adjacent to duplex DNA was also investigated and distinct roles for both 5' and 3' DNA overhangs were identified. 5' single stranded overhangs preferentially stimulate DNA-PK activity while 3' overhangs are involved in bridging the synaptic complex and aligns DNA termini across the gap via microhomology. How these sequences manipulate DNA-PK activation on longer substrates has yet to be determined. In the previous chapter I identified the profound role substrate length plays in DNA-PK activation. Thus we investigated how the terminal bases influence DNA-PK activation on longer substrates. Because genomic DNA mega bases in length, it is likely that these studies are more representative of DSBs than short oligo nucleotides.

4.2 Results

Substrates with 5' overhangs are of particular interest as the influence of the C-terminal domains on DNA-PKcs stimulation varied significantly according to the length of the DNA substrate. In order to determine if overhang sequence altered how of the Ku80 C-terminal structural regions influence kinase stimulation, the same plasmid was digested with different restriction endonucleases that generate distinct sequences and

overhangs at the terminus. Despite the differences in terminal sequences, the influence of C-terminus was similar when both experiments were normalized to WT (**Figure 11A**). Under both conditions the helical bundle was required for maximum stimulation while the core DNA binding/Ku70 dimerization domain and disordered linker regions both contributed to stimulation above the no Ku control. While the relationships between the abilities of the mutants to activate the kinase remains the same (**Figure 11A**), the overall level of kinase activation from XhoI digested DNA was significantly greater than the activation observed by the BamHI digested DNA (**Figure 11B**). Despite the substantial differences in kinase stimulation, the sequence differences in the ends generated from digestion with these restriction enzymes are subtle. While digestion from both enzymes results in a 4 base 5' overhang, digestion by XhoI generates a 5'TCGA overhang and digestion by BamHI generates a 5'GATC overhang. Using additional restriction enzymes and a GFP reporter plasmid, we next tested whether the terminal pyrimidines generated by XhoI digestion were responsible for increased activation. XbaI digestion yields a 5'CTAG single stranded overhang while EcoRI digestion yields a 5'AATT single stranded overhang. Consistent with our prediction the linearized plasmid with terminal pyrimidines, (XbaI digested) resulted in greater DNA-PK activation than that of terminal pyrimidines (**Figure 12A)**. Interestingly, both BamHI and EcoRI digested DNA contain single stranded overhangs with pyrimidines adjacent to the terminal purines however these ends do not stimulate DNA-PK as high as those with terminal pyrimidines. All four of the 5' overhang-generating restriction enzymes used in this study result in single strand regions containing 2 pyrimidines and 2 purines and yet these cause drastically different levels of DNA-PK stimulation. This illustrates that subtle variations

in DNA sequence can have profound effects on DNA-PK activation. Further these subtle variations in DNA sequence are not "sensed" by Ku in the same way DNA structural variations are. Instead it seems that this ability to distinguish terminal pyrimidines from purines is intrinsic to DNA-PKcs.

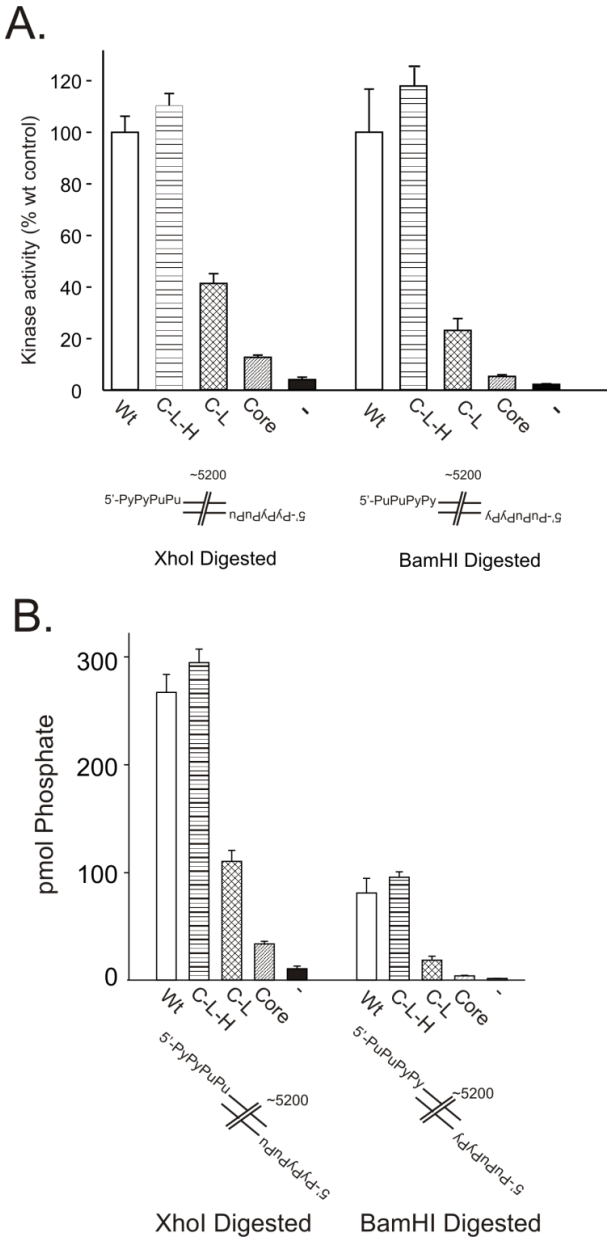


Figure 11. Preferential DNA-PK Action by Terminal Pyrimidines on 5' Overhangs. DNA-PK kinase stimulation with plasmid DNA linearized with XhoI and BamHI generating 4 base 5' single stranded overhangs. DNA substrates are depicted pictorially. DNA termini generated by digestion are depicted below graph indicating locations of pyrimidines (Py) and purines (Pu). Ku80 constructs was represented as follows: WT bars are white, C-L-H bars contain horizontal lines, C-L bars contain crosshatching, core bars contain diagonal lines, and no added Ku (-) bars are solid. A. Kinase activity is normalized to average of WT in each reactions. Data are reported as a percentage of WT. B. Kinase activity is reported as pmol of phosphate transferred. Error bars indicate 1 standard deviation from the mean.

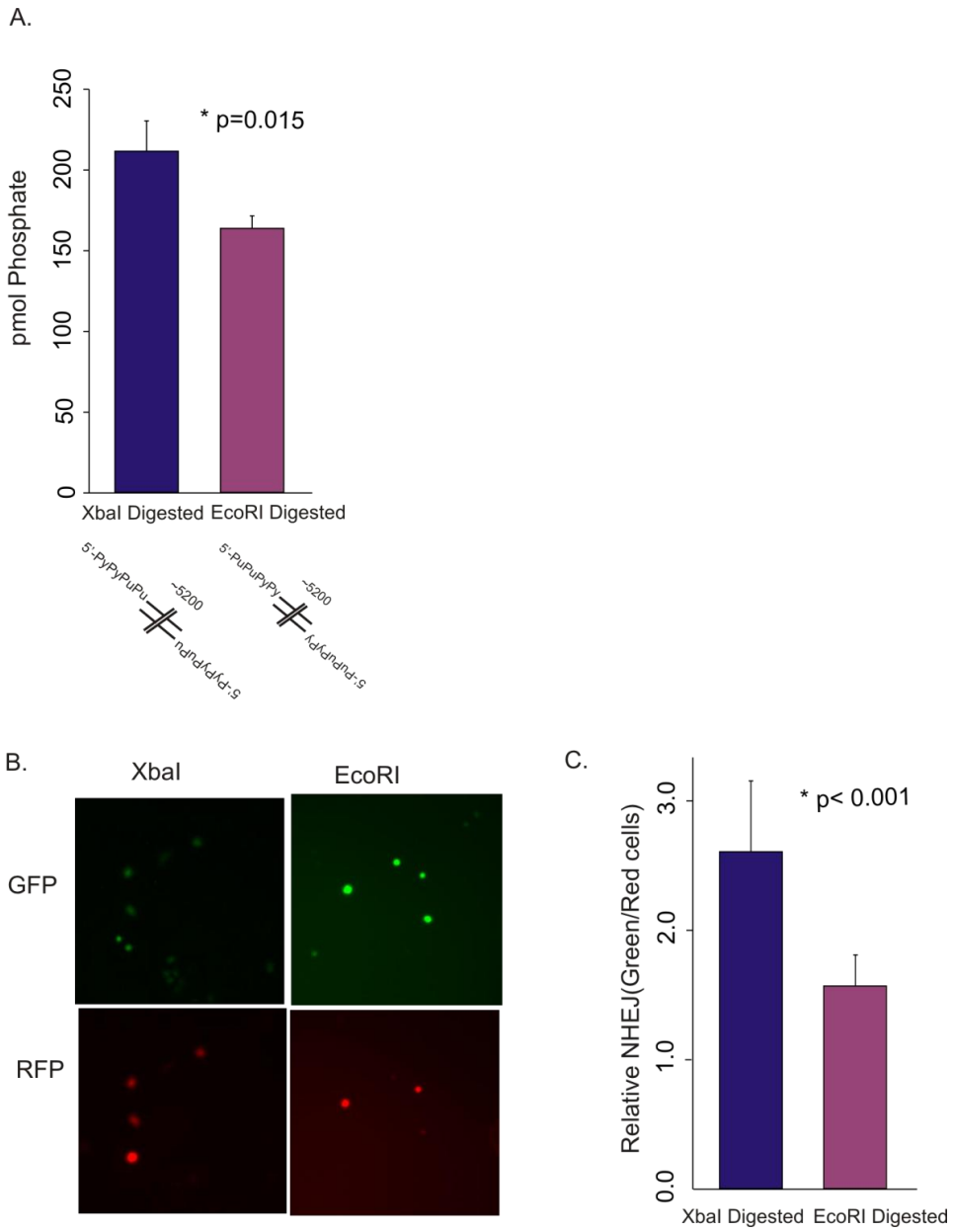


Figure 12. Preferential DNA-PK Activation by Terminal Pyrimidines Leads to Increased NHEJ. A. DNA-PK kinase activity with WT Ku and linearized reporter plasmid digested with restriction enzymes XbaI and EcoRI. DNA termini generated by digestion are depicted below graph indicating locations of pyrimidines (Py) and purines (Pu). Activity is reported as pmol of phosphate transferred. **B.** Representative images analyzed for host cell reactivation assay. **C.** Quantified results from host cell reactivation assay. Results are reported as ratio of green cells to red fluorescent cells. Results from XbaI digested DNA are shown in blue and results from EcoRI digested DNA are shown in purple. Error bars indicate 1 standard deviation from the mean.

Because DNA-PK activation is a major regulator in NHEJ[1], we asked if the sequence-specific activation by terminal pyrimidines would impact DNA repair. To investigate this question, a host cell reactivation assay was employed utilizing a green fluorescent protein(GFP) reporter gene that is expressed following NHEJ. This plasmid was linearized with either XbaI or EcoRI, both of which generate 4 base 5' overhangs between the promoter and GFP expression reporter gene, however XbaI digestion yields an NHEJ substrate with terminal pyrimidines while EcoRI digestion yields an NHEJ substrate with terminal purines. Differences in transfection efficiency were controlled via co-transfection with circular pCAG-dsRED which produces a red fluorescent protein (RFP) independent of NHEJ [63]. The results presented in **Figure 12B** are representative images of individual fields and a composite. Quantified results are presented in **Figure 12C** as a ratio of green cells: red cells and show that NHEJ substrates generated with XbaI digestion are repaired more efficiently than those generated via EcoRI digestion ($p < 0.001$). We have previously shown using this plasmid that expression of the GFP does not occur in the presence of a DNA-PK inhibitor, NU7441, nor does it occur in NHEJ deficient cells[63]. Thus expression of GFP is indicative of NHEJ and it is clear that NHEJ substrates with terminal pyrimidines are repaired more efficiently than those of with terminal purines. These results are consistent with a model where the sequences at the termini of DSBs influence DNA-PK activation and ultimately repair.

4.3 Discussion

To the best of our knowledge, we provide the first report that the sequence surrounding a DSB dictates how efficiently it can be repaired which can be attributed to DNA-PK activation (**Figure 12**). Others have shown that DSB with 3' overhangs are

repaired more efficiently than those with 5' overhangs however sequence bias was not investigated extensively or attributed to DNA-PK activation[68]. Indeed, DNA-PK displays a dramatic sequence bias for terminal pyrimidines even over sequences with terminal purines containing adjacent pyrimidines (**Figure 11B and Figure 12A**). We did not observe any differences in DNA-PKcs stimulation by the Ku mutants due to terminal sequence variations. Clearly when it comes to DNA-PK activation, not all bases are created equal and their precise location at the DSB has a considerable impact on repair.

Why terminal pyrimidines are more stimulatory than terminal purines is still unknown. Interestingly this preference for pyrimidines is also shared with other enzymes involved in NHEJ. XRCC4/Ligase IV, the predominant ligase complex in NHEJ (**Figure 1**), typically ligates duplex DNA but can ligate ssDNA if the termini are poly-dT. In addition, POL X polymerase family member Pol μ has been shown to add nucleotides in a template-independent manner during NHEJ[69]. Pol μ prefers to add pyrimidines, particularly dT when it uses a template-independent manner[70]. The terminal pyrimidines preference along with the increased activation we report with DNA-PK could help explain in why substrates with terminal pyrimidines are repaired more efficiently than those with terminal purines. Further the fact that in the absence of a template Pol μ inserts pyrimidines may be a mechanism by which DNA-PK is stimulated. It would be interesting to see the endonucleases and exonucleases trim purines more efficiently thereby leaving pyrimidines rich termini. In addition it will be important in future studies to determine if terminal sequences influence DSB repair accuracy but how these relationships may exist cannot be predicted. In example one could envision a mechanism where DSBs with terminal pyrimidines increase DNA-PK activation, allowing for quick

repair. This efficient repair would thereby limit termini from processing leading to higher fidelity at the site of the break. Contrastingly, increased DNA-PK activation due to terminal pyrimidines may increase DNA-PK-dependent activation of endonucleases such as Artemis leading to increased cleavage at the site of the break making repair less fidelitous.

The fact that Ku is not involved in sequence discrimination is consistent with structural data which shows that Ku interacts exclusively with the phosphodiester backbone of DNA. Therefore Ku's inability to differentiate sequence variation is not surprising. Ku does however recruit both Xrcc4/Ligase 4 and Pol μ which clearly do show sequence discrimination activities. Thus perhaps the discrimination of sequences is due to other proteins in NHEJ while discrimination of DNA structures is determined by Ku. Indeed, structural studies of DNA-PKcs, reveal two potential DNA binding domains[48,71]. The larger of the two located in the palm domain is almost assuredly involved in binding dsDNA while the smaller domain located in the head domain has been implicated in binding ssDNA[71]. Perhaps the single ring structure of pyrimidines is less sterically hindered thus interacts more strongly with this putative ssDNA binding domain as compared to the double ring structure of purines. This mechanism is currently under investigation. Further, there are numerous other DNA structures and conformations that exist throughout the genome, all of which are susceptible to DSBs. It will be important to evaluate how these regions influence DNA-PK activation in light of the differential activation and repair observed in this study.

5. Discussion

Non-homolgous end joining is a critical pathway for the repair of DSB and is responsible for the majority of said repair in mammalian cells[72,73]. A crucial step in regulating this pathway, activation of the DNA-dependent protein kinase, occurs immediately following the recognition of free double stranded DNA termini. The free termini are initially bound by the heterodimeric Ku protein consisting of Ku70 and Ku80. Following relatively modest but consistent conformational changes upon Ku binding to DNA, DNA-PKcs is recruited to DNA terminus[37]. This recruitment was further investigated in this study. To date the only regions of Ku which were shown to undergo conformational changes following DNA binding are the N-terminal and C-terminal regions of both Ku70 and Ku80 [37,48]. Interestingly there is no direct evidence that any of these four regions interact with DNA. Of these regions the most likely to interact with DNA is the C-terminus of Ku70 through the conserved SAP domain [42]. Previous studies have concluded that the extreme C-terminus of Ku80 is responsible for recruiting DNA-PKcs to the DNA break[46]. In our study we report contrasting data which indicate that the entire C-terminus is dispensable for DNA-PKcs recruitment (**Figure 8**). In fact the DNA-PKcs recruitment by the Core mutant and WT Ku were statistically indistinguishable. An extensive interaction interface between the core DNA binding/dimerization domain of Ku and DNA-PKcs has been identified via SAXS studies[48]. Despite the fact that we and others have not identified major conformational changes in core of Ku following DNA binding, these Ku/DNA-PKcs interactions have not been observed in the absence of DNA[37,48]. Perhaps DNA-PKcs must undergo a DNA dependent conformational change in order to interact with the Ku core which remains

largely unchanged. Another explanation is that DNA-PKcs shows weak affinity for both the core domain of Ku and DNA and can only bind to DNA termini due to the summation of these weak interactions.

Once recruited to the site of the DSB DNA-PK is activated and must discriminate a wide variety of DNA structures in order to coordinate appropriate processing[68]. The results presented here define the mechanism by which DNA-PKcs activation is stimulated through interactions with the C-terminus of Ku80. Interestingly the structurally defined regions of the C-terminus of Ku80 responsible for DNA-PKcs stimulation are different depending on the structure of the DNA cofactor to which the kinase complex is bound. We interpret these results to mean that the Ku80 C-terminus is responsible for discrimination of DNA substrates through specific protein-protein interactions with DNA-PKcs. In support of this interpretation, each of the Ku mutant constructs we designed was capable of supporting WT levels of Ku dependent DNA-PKcs stimulation on specific DNA substrates (**Table 4**). Perhaps the most interesting results were obtained from the analysis of DNA-PK activation using 60 bp substrates. The Core mutant which lacks the entire C-terminus was capable of supporting WT levels of DNA-PKcs stimulation on the 60 bp substrate with 5' overhangs suggesting that in this case the C-terminus is dispensable for activation. In the presence of 60 bp substrates with blunt ends and 3' overhangs, on the other hand, both require the linker and the helical bundle to retain WT levels of activation (**Table 4**). Why the activation relationships of the 60 bp substrates are so different than the rest of the substrates test remains unknown. These may be due to sequence variation however, the mixed nature of the substrates provide no obvious sequence patterns that would result in these activation relationships.

Further all of the sequence bias which we investigated extensively demonstrate that terminal sequences only changed overall activation but not protein/activation relationships. Thus it is likely that the distinct results obtained from using the 60 bp substrates are due to the duplex length of these substrates. The Jackson Lab has recently demonstrated for the first time that two Ku dimers bind to each side of DNA termini during NHEJ. This possibility is interesting because the 60 bp substrates could uniquely accommodate 2 Ku dimers but may limit lateral movement of the proteins complexes along the DNA as the Ku molecules may collide causing either to slip off the DNA terminus. Ma and Lieber have previously reported that Ku/DNA binding kinetics indicate that Ku binds DNA by a non-cooperative mechanism on substrates that are 18 and 22 bp in length, however it exhibits cooperative binding on substrates 45 bp and longer[74]. This provides an important possible explanation for our results involving the 60 bp substrates. Thus the 30 bp substrates may accommodate only 1 Ku dimer while the 60 bp and longer substrates accommodated 2 or more. While the length of the 30 bp substrates prevents multiple Ku molecules from binding to each terminus, the 400 bp and plasmid DNA substrates are certainly capable of binding multiple Ku molecules. These substrates however differ from the 60 bp substrates in that the steric crowding events on the longer substrates are greatly reduced. It should be noted that the experiment was designed to avoid multiple Ku molecules binding to one terminus which we predicted would complicate interpretation. As stated in Chapter 2 a 10-fold excess of DNA to Ku was used during the 30 bp and 60 bp experiments. The DNA to Ku species was also in vast excess to DNA-PKcs as evidenced by the lack of increased kinase activity in the presence of higher concentrations of Ku (data not shown). Through EMSA analysis we

were never able to observe 2 Ku molecules on a single 60 bp substrate under these conditions. That being said it is still possible that this species could exist at an undetected level or as an unstable intermediate and it is this species that contributes to DNA-PKcs activation in the presence of the 60 bp substrates. It is impossible to give a thorough assessment of the unpublished dimer of Ku dimers on each DNA terminus which Jackson has recently reported; however, these data along with DNA length dependent cooperative binding model presented by Ma and Lieber may be the key in explaining the atypical activation relationships observed with the 60 bp substrates[74]. The differences in activation relationships among the DNA substrates of various lengths suggest that Ku plays a role in distinguishing DNA length. Despite the fact that Ku has been extensively studied for decades, these data indicating Ku's DNA length discrimination are novel. Based on structural studies which clearly indicate that Ku binds to DNA termini, it is not obvious that Ku could distinguish various DNA lengths[35,48,75]. Additionally, 3 atomic force microscopy studies have been used to directly visualize Ku binding to DNA[28,29,76]. Data from both studies show that Ku primarily interacts with DNA termini. Further, Cary and colleagues are the only ones to show any Ku bound to internal DNA regions, but this only occurs when Ku is in vast excess to DNA and is independent of DNA-PKcs[29]. Thus it is not clear how a terminal DNA binding protein such as Ku is able to discriminate the difference in length of a 400 bp substrate versus a 5200 bp substrate. Since the interior of the DNA binding, Ku core contacts approximately 20 bases or 2 turns of DNA, it is unlikely that distinction of DNA length occurs in this domain or at the very least through this domain alone[35]. One way in which this may be accomplished is through protein/DNA interactions where the DNA loops

back upon itself to interact with Ku bound to the terminus at a region besides the primary DNA binding domain. While the 400 bp substrate is longer than the persistence length or the minimum DNA length necessary for DNA to circularize[53,77], it seems obvious that the longer the DNA is the more molecular conformations it can assume and the higher the possibility of it to loop. In addition Ku contains a second putative DNA binding domain located at the C-terminus of Ku70, the SAP domain, which may allow for this molecular differentiation. Another explanation is that length discrepancy between the 30 bp, 60 bp, and 400 bp substrates occurs through differences in lateral motion of Ku or the ability to slide along the DNA. Such sliding could allow for multiple Ku molecules to bind to a single piece of DNA. Using this model the differences between the 400 bp and plasmid DNA would not be explained through length as both provide ample room for multiple Ku molecules to bind and slide along DNA. Thus, the distinction between the 400 bp substrates and the plasmid DNA may not be due to length at all and instead involves methylation status. The plasmid DNA used for the **Figure 10** was purified from *E. coli* and as such is methylated. This may contribute to the differences in Ku dependent DNA-PK stimulation between 400 bp and linearized plasmid substrates. The most notably of these differences is that linearized plasmid substrates with blunt ends uniquely require the extreme C-terminus for maximum kinase activation while other substrates allow for maximum activation using the C-L-H mutant (**Figures 9 and 10**). Determining which of these is correct should be part of the immediate follow-up work for this research. To do this relatively straight forward approaches could be taken to eliminate methylation as a factor contributing to kinase activation. First, PCR could be used to amplify the plasmid DNA followed by digestion of template DNA with DpnI. This would yield an

unmethylated plasmid DNA that could be isolated from template fragments and digested further to generate desired DNA termini. Additionally the 400 bp substrates could be cloned into a plasmid and purified from *E. coli.* These 400 bp substrates could then be digested from the plasmid vector and isolated yielding the original substrate that is now methylated. If methylation does indeed influence activation of DNA-PKcs by Ku then this could have interesting implications as DNA methylation plays an important role in chromatin structure and gene regulation in eukaryotic cells. Because eukaryotic DNA is much longer than most of the substrates investigated here, it is likely that the linearized plasmid DNA best represents DSBs that occur *in vivo* and thus our model is primarily based on the activation relationships generated with these substrates. When Ku is bound to DSBs with 3' or 5' single stranded overhangs the linker region and helical bundle stimulate DNA-PKcs activity (**Figure 13A**). In general substrates with 5' overhangs are more stimulatory than those with 3' overhangs. Based on these results and the previous work in our lab concerning DNA substrate influence on DNA-PK activation, the 5' overhang is depicted as interacting directly with the head module of DNA-PKcs (**Figure 13A**)[53]. The 3' overhangs, on the other hand, has been shown to be involved in synaptic complex formation through microhomology dependent DNA alignment[53]. When the complex is bound to DNA substrates with blunt ends the linker region, helical bundle, and extreme C-terminus contribute to kinase stimulation, however whether these interactions occur in *cis* (**Figure 13B**) or in *trans* (**Figure 13C**) remains unknown. The existence of interactions between the Ku 80 C-terminus and DNA-PKcs occurring in *trans* was first suggested following SAXS studies which indicate that the flexibility and

dimensions are sufficient to allow such interaction, but no direct evidence has suggested such an interaction exists[48].

The outstanding question remains, "how does the C-terminus of Ku80 influence DNA-PK synaptic complex formation?" In this investigation we report enzymatic data that implicates each of the three structural regions of the C-terminus in DNA-PKcs stimulation. While we do not report any evidence for interaction between the C-terminus and DNA-PKcs occurring in *trans* there is equally no evidence that these interactions must occur in *cis*. Additionally it is possible that some interactions may occur in *cis* while others occur in *trans.* For example, **Figure 13A** depicts in *cis* interactions occurring between the linker region and the helical bundle while the extreme C-terminus is extended in the nucleoplasm and free to interact with a DNA-PKcs molecule across the synapse. As stated in chapter 1, this region has been reported to be "necessary and sufficient" for DNA-PKcs interaction[43].

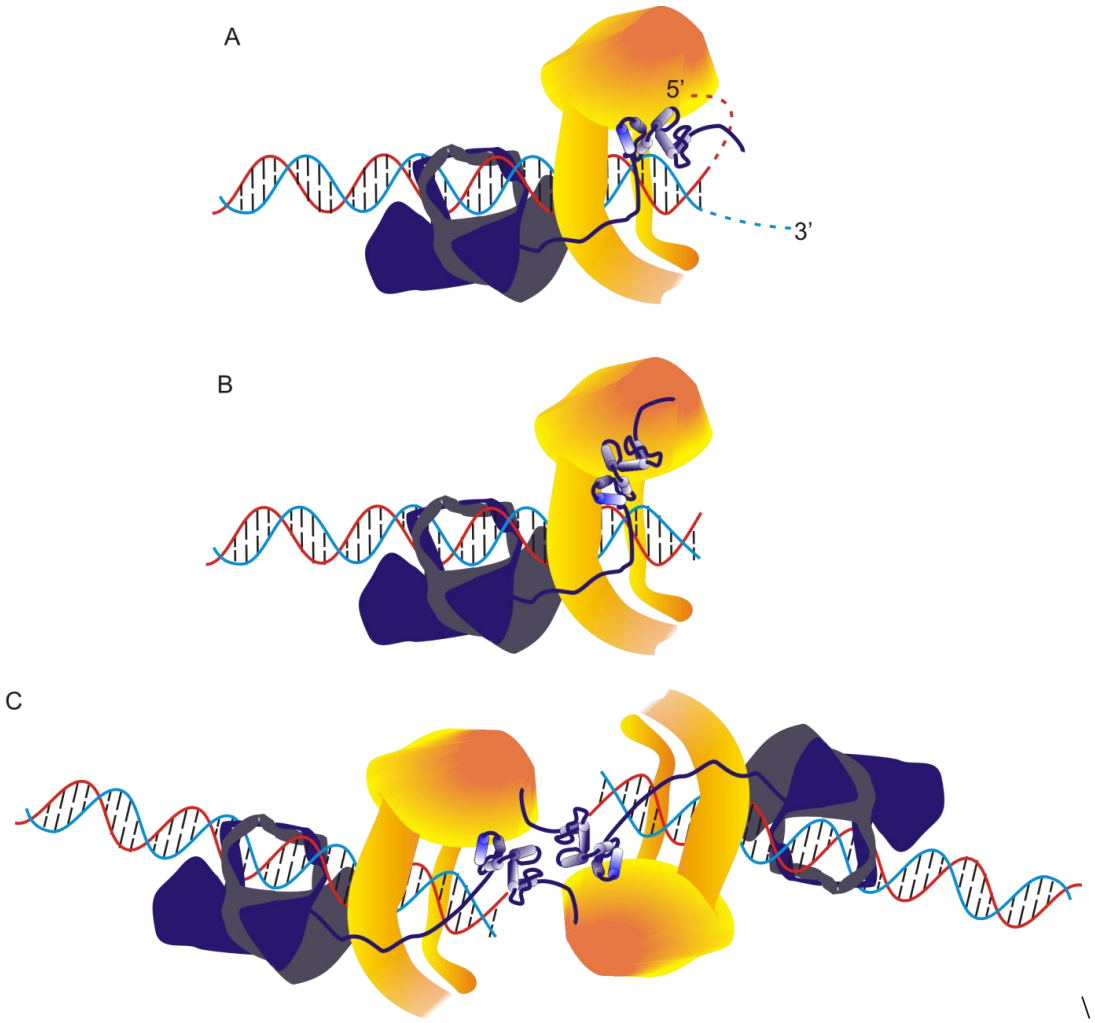


Figure 13. Model of DNA-PK Activation by the C-Terminus of Ku80. A. DNA-PK stimulation occurs through interactions with the linker region and helical bundle on DNA substrates with 5' or 3' overhangs. Ku70 is shown in grey. Ku80 is shown in dark blue. DNA-PKcs is shown in yellow and orange. **B-C** DNA-PK stimulation on DNA substrates with blunt ends occurs through interactions with the linker region, the helical bundle, and the extreme C-terminus. These interactions may occur in-*cis* or in-*trans*.

While the data in this investigation clearly show that this region is not necessary for DNA-PKcs stimulation on all substrates and only effects kinase activity in one of the many substrates tested, this certainly does not eliminate the possibility of this region having non-stimulatory interactions with DNA-PKcs. To date this region of Ku has only been implicated to interact with DNA-PKcs and the WRN protein[78] despite the fact that Ku has been shown to interact with several other proteins[21,79]. Thus there is at least a potential that the extreme C-terminus could be involved in synaptic complex formation either directly by reaching across the synapse to interact with DNA-PKcs in *trans* or indirectly by interacting with DNA-PKcs in *cis* allowing for a conformation which favors synaptic complex formation. While we predict synaptic complex formation to be influenced by the C-terminus of Ku80 it is certainly multi-factorial. Our lab has reported that synaptic complex formation is influenced by DNA substrate structure particularly by those containing 3' overhangs[53]. Others have shown through electron microscopy studies that synaptic complexes are influenced by the phosphorylation status of DNA-PKcs[49]. It is important to note that in both conditions dimers of DNA-PK were observed but were in lower abundance in the phosphorylated state. Since we and others have shown that DNA-PK kinase activation dramatically influenced by the C-terminus of Ku80[45,62] and the phosphorylation status of DNA-PK influence synaptic complex[49] it seems plausible that C-terminus of Ku80 could regulate synaptic complex formation to some degree. Interestingly, another studied using X-ray scattering revealed that in the absence of Ku, DNA-PKcs assembles into dimers but these become disassociated in presence of Ku when DNA is absent[48]. This suggests that DNA-PKcs has at least some intrinsic affinity to form synaptic complexes. While we have shown the DNA substrate influences which

structural region of the Ku80 C-terminus stimulates kinase activity, we observed that even non-stimulatory effects of the C-terminus were not inhibitory. An interpretation of these data along with the reported effect of DNA-PKcs phosphorylation on dimer disassociation would conclude that the C-terminus of Ku80 would likely destabilize synaptic complexes via kinase stimulation. This may be an over simplistic expectation however as DNA-PKcs autophosphorylation and its effects are much more complex than illustrated in these studies. As discussed in chapter 1, autophosphorylation sites can have reciprocal influences[54]. For example mutational analysis has shown that autophosphorylation of the ABCDE cluster promotes DNA-PK dissociation from DNA while phosphorylation of PQR cluster reduces dissociation[80]. Further many autophosphorylation sites including J, K, L, and N or combination of these sites have been shown to not influence dissociation regardless of phosphorylation status[54]. DNA-PK autophosphorylation is a critical target of DNA-PK kinase activity, thus determining if and how DNA substrate structure influences phosphorylation of specific sites will be important understanding DSB repair. Again our investigation clearly shows that stimulation of DNA-PKcs is intimately tied to the protein structure of Ku80 C-terminal regions and the structure of the DNA substrate to which the holoenzyme is bound. Thus if DNA structure influences phosphorylation of particular sites we should be able to identify which structural regions of the Ku80 are responsible for directing kinase activity to these specific targets. There is some evidence to suggest that the C-terminus of Ku80 may direct phosphorylation of some sites but not others. Using a Ku80 mutant containing amino acids 1-598 which is similar to our C-L mutant construct, Weterings and colleagues determined that phosphorylation of threonine 2647 was reduced compared

to WT Ku controls while phosphorylation of serine 2056 and threonine 2609 was unchanged[47]. This is interesting because similar to our results presented in Chapter 3 with the C-L mutant, the 598 mutant stimulated phosphorylation of a p53 based peptide substrate to approximately 50% of WT levels. Despite overall reduction of DNA-PK activity, at least a portion of downstream targets were phosphorylated to the same extent in the absence helical bundle and extreme C-terminus. One major flaw with this study is that different DNA cofactors i.e. sheared calf thymus DNA and 250bp were used to stimulate phosphorylation of the p53 peptide and the autophosphorylation experiment respectively[47,77]. Regardless, these data provide interesting basis for further investigation. If autophosphorylation is altered by DNA substrate structure, then pathway choice may also be influenced by the structure of break and should be interrogated. All known autophosphorylation sites of DNA-PKcs should be investigated with regard to DNA substrate structure; however, there is a subset of sites which may be particularly pertinent to pathway choice. For example, phosphorylation of the ABCDE cluster tends to promote homologous recombination (HR) while phosphorylation of the PQR cluster promotes repair via NHEJ[19]. These activities are coordinated with their influence on DNA-PK dissociation with DSBs discussed above. HR required extensive resection from the 5' end of the break yielding a long 3' overhang[67]. This single stranded DNA is then used in to seek homologous regions of undamaged DNA[81]. This extensive resection requires nucleases to access the DNA termini, which is significantly limited upon phosphorylation of the PQR cluster[80]. Additionally, phosphorylation of the JK cluster has been shown to inhibit NHEJ while promoting HR but do not influence overall kinase activity[19]. Autophosphorylation sites including threonine 3950 and the N site

cluster limit kinase activity are also of interest in terms of DNA substrate bias[19]. Our results from the 400 bp substrates with 5' overhangs indicate no measurable kinase activity with the Core and C-L mutant despite that these mutants stimulate activity on the 400 bp substrates with 3' overhangs and blunt ends (**Table 4**). One explanation of these results is that these mutants direct autophosphorylation of threonine 3050 and the N cluster thereby ablating the potential kinase activity.

Data from several groups have established that limiting NHEJ promotes HR while efficient NHEJ limits HR[16,82]. Our results from the host-cell reactivation assay clearly demonstrate that the sequence of the DNA terminus influences NHEJ efficiency (**Figure 12**). While our assay does not measure HR, it is only logical that the increased NHEJ efficiency that we observe will reduce HR effectiveness. The Meek laboratory has clearly demonstrated a relationship between NHEJ efficiency and DNA-PKcs autophosphorylation[19,20]. Therefore it is likely that autophosphorylation of specific sites of DNA-PK will be influenced by the structure and sequence of DSB. The extent to which the DNA substrate contributes to pathway choice repair via HR or NHEJ, however is unknown.

Testing the activation relationships between DNA substrate structure and Ku80 protein regions *in vivo* remains an important goal. Several factors make accomplishing this complex. The first is that Ku80 is an essential protein and as such few cell lines exist that are devoid of Ku80. This limits our experiments to the Chinese Hamster ovarian cell line Xrs6 which is Ku80 deficient. Based on work of others we would expect that our Ku80 deletion constructs will complete rodent Ku70 sufficiently for DNA-PKcs stimulation[45,46]. A major drawback with this cell line, however, is that they have

rampant, Ku independent, alternative-NHEJ activity capable of repairing DSB. Thus measuring what could be small changes in NHEJ efficiency based on Ku80 protein mutations and DNA substrate structures may be difficult to discriminate if the alternative-NHEJ pathway is compensating for these reductions. To circumvent these problems we would need to adapt an assay that allows for us to distinguish repair via classical and alternative NHEJ. It would also be helpful to determine the relative amount of HR which occurs when NHEJ efficiency is altered. In addition, these studies may be strengthened with the use of alternative-NHEJ inhibitors such as the MRN inhibitor Mirin[83]. This would eliminate a variable which complicates interpretation of repair efficiencies. This approach is not without its own weaknesses. Mirin is not a very potent inhibitor and requires the use of 1mM in biochemical assays[84]. When experiments require such high concentrations of inhibitors the likelihood of influencing other pathways in cells is raised. The use of Mirin therefore should only be used to complement these experiments and the study should not be designed around their use.

In conclusion, in this study we have defined a novel mechanism by which the DNA-PK holenzyme complex is formed. We determined that specific protein/protein interactions occur between three structural regions found at the C-terminus of Ku80 and DNA-PKcs. These interactions are dictated by the DNA cofactor to which the complex is bound. Through an unknown mechanism Ku is able to differentiate multiple structural features of DNA including the length, the presence of overhangs, and the orientation of overhangs. To date the functions defined for Ku have centered of its DNA terminus binding activity[72,79]. Here data are presented that despite the fact that Ku binds to DNA termini, it can distinguish the length difference between 400 bp and 5.4 kb DNA. While

the ability of Ku to "sense" overhangs which are located at the terminus is in accord with its end binding activity, how this activity relates to DNA length discrimination is not obvious. Regardless Ku translates this information to DNA-PKcs through specific, stimulatory interactions which vary according to the structure of the DNA cofactor. We provide evidence that these interactions explicitly control kinase activation and do not influence DNA-PKcs-DNA binding activity. Further, the mechanism by which DNA-PK is influenced by DNA sequence was investigated. It was determined that on linearized plasmid substrates with 4 base 5' overhangs, those with terminal pyrimidines cause much greater kinase stimulation than those with terminal purines. This DNA sequence specific increase in kinase activation corresponds to increases in NHEJ in vivo. To our knowledge this is the first report that suggests that the sequence surrounding a DSB influences its repair. Unlike differences in overhang structure and length, the differential action does not occur through Ku but instead is intrinsic to DNA-PKcs. It has been established that following activation DNA-PK is highly regulated through autophosphorylation[72]. Additionally structural data indicate that DNA-PKcs/DNA-PKcs dimeric interactions are substantially influenced by the structure of DNA cofactors to which DNA-PKcs is bound[48].

This work provides a significant advancement to the general field of DSB repair. Here we provide examples where the structure at sequence surrounding the DSB influence repair. The majority of this work was demonstrated in vitro with DNA substrates which remain static. Considering the fact that DNA-PK coordinates processing of DSBs, it is important to realize that during NHEJ the structure of the DSB may fluctuate due to processing. The interactions which contribute to activation of DNA-

PK likely change as the structure of the DNA cofactor changes. In addition the level of DNA-PK activity likely changes as DNA termini are processed.

Based on the work presented here, future studies should investigate how the C-terminus of Ku80 influences synaptic complex formation. Clearly the DNA structure and kinase activation influence synaptic complex formation[48]. Since the effect of the C-terminus on kinase activity is dependent on DNA structure, there is likely a direct influence of the C-terminus on synaptic complex formation. In addition, several studies have shown that DNA-PK autophosphorylation is perhaps the most important downstream target of DNA-PK kinase activity[72]. How DNA-PK autophosphorylation is influenced by DNA cofactor structure has yet to be tested. Data presented here demonstrate that DNA-PK "senses" the structure of the DNA to which it is bound via interactions with Ku. It follows that DNA-PK would necessarily respond differently to distinct DNA structures, likely through specific autophosphorylation events which would allow or restrict processing events as needed. Autophosphorylation events have also been linked to repair pathway choice as discussed in Chapter 1[19]. It will also be of interest to test whether DNA structure plays a role in pathway choice. HR requires long 3' overhangs which are used to invade undamaged strands in order to find homologous template regions[85]. Current models of pathway initiation suggest that these overhangs are produced by the activity of the MRN complex. It may be that the pathway is stimulated by existing 3' overhangs or that 3' overhangs differentially stimulate DNA-PK which in turn stimulates HR activity. This work demonstrates that the structure of DNA in DSBs influences DNA-PK activation and NHEJ. Future studies should focus on measuring the existent to which DNA cofactor structure contribute to other molecular mechanisms.

6. References

1 Woods, D. & Turchi, J. J. Chemotherapy induced DNA damage response: Convergence of drugs and pathways. *Cancer biology & therapy* **14** (2013).

2 Shuck, S. C., Short, E. A. & Turchi, J. J. Eukaryotic nucleotide excision repair: from understanding mechanisms to influencing biology. *Cell research* **18**, 64-72, doi:10.1038/cr.2008.2 (2008).

3 Deans, A. J. & West, S. C. DNA interstrand crosslink repair and cancer. *Nature reviews. Cancer* **11**, 467-480, doi:10.1038/nrc3088 (2011).

4 Burma, S., Chen, B. P. & Chen, D. J. Role of non-homologous end joining (NHEJ) in maintaining genomic integrity. *DNA repair* **5**, 1042-1048, doi:10.1016/j.dnarep.2006.05.026 (2006).

5 Carr, A. M., Paek, A. L. & Weinert, T. DNA replication: failures and inverted fusions. *Seminars in cell & developmental biology* **22**, 866-874, doi:10.1016/j.semcdb.2011.10.008 (2011).

6 Radiation damage to DNA: techniques, quantitation and mechanisms. Bowness-on-Windermere, Lake District, United Kingdom, April 19-24, 1997. *Radiation research* **148**, 481-522 (1997).

7 Siddiqi, M. A. & Bothe, E. Single- and double-strand break formation in DNA irradiated in aqueous solution: dependence on dose and OH radical scavenger concentration. *Radiation research* **112**, 449-463 (1987).

8 Umezawa, H., Maeda, K., Takeuchi, T. & Okami, Y. New antibiotics, bleomycin A and B. *The Journal of antibiotics* **19**, 200-209 (1966).

9 Goodwin, K. D., Lewis, M. A., Long, E. C. & Georgiadis, M. M. Crystal structure of DNA-bound Co(III) bleomycin B2: Insights on intercalation and minor groove binding. *Proceedings of the National Academy of Sciences of the United States of America* **105**, 5052-5056, doi:10.1073/pnas.0708143105 (2008).

10 Chen, J. & Stubbe, J. Bleomycins: towards better therapeutics. *Nature reviews. Cancer* **5**, 102-112, doi:10.1038/nrc1547 (2005).

11 Chen, J., Ghorai, M. K., Kenney, G. & Stubbe, J. Mechanistic studies on bleomycin-mediated DNA damage: multiple binding modes can result in double-stranded DNA cleavage. *Nucleic acids research* **36**, 3781-3790, doi:10.1093/nar/gkn302 (2008).

12 D'Andrea, A. D. & Haseltine, W. A. Sequence specific cleavage of DNA by the antitumor antibiotics neocarzinostatin and bleomycin. *Proceedings of the National Academy of Sciences of the United States of America* **75**, 3608-3612 (1978).

13 Wu, J. C., Kozarich, J. W. & Stubbe, J. Mechanism of bleomycin: evidence for a rate-determining 4'-hydrogen abstraction from poly(dA-dU) associated with the formation of both free base and base propenal. *Biochemistry* **24**, 7562-7568 (1985).

14 Kunimoto, T., Hori, M. & Umezawa, H. Modes of action of phleomycin, bleomycin and formycin on HeLa S3 cells in synchronized culture. *The Journal of antibiotics* **20**, 277-281 (1967).

15 Pommier, Y., Leo, E., Zhang, H. & Marchand, C. DNA topoisomerases and their poisoning by anticancer and antibacterial drugs. *Chemistry & biology* **17**, 421-433, doi:10.1016/j.chembiol.2010.04.012 (2010).
16 Brandsma, I. & Gent, D. C. Pathway choice in DNA double strand break repair: observations of a balancing act. *Genome integrity* **3**, 9, doi:10.1186/2041-9414-3-9 (2012).
17 Pawelczak, K. S., Bennett, S. M. & Turchi, J. J. Coordination of DNA-PK activation and nuclease processing of DNA termini in NHEJ. *Antioxidants & redox signaling* **14**, 2531-2543, doi:10.1089/ars.2010.3368 (2011).
18 Convery, E. *et al.* Inhibition of homologous recombination by variants of the catalytic subunit of the DNA-dependent protein kinase (DNA-PKcs). *Proceedings of the National Academy of Sciences of the United States of America* **102**, 1345-1350, doi:10.1073/pnas.0406466102 (2005).
19 Neal, J. A. *et al.* Inhibition of homologous recombination by DNA-dependent protein kinase requires kinase activity, is titratable, and is modulated by autophosphorylation. *Molecular and cellular biology* **31**, 1719-1733, doi:10.1128/MCB.01298-10 (2011).
20 Neal, J. A. & Meek, K. Choosing the right path: does DNA-PK help make the decision? *Mutation research* **711**, 73-86, doi:10.1016/j.mrfmmm.2011.02.010 (2011).
21 Downs, J. A. & Jackson, S. P. A means to a DNA end: the many roles of Ku. *Nature reviews. Molecular cell biology* **5**, 367-378, doi:10.1038/nrm1367 (2004).
22 Mimori, T. *et al.* Characterization of a high molecular weight acidic nuclear protein recognized by autoantibodies in sera from patients with polymyositis-scleroderma overlap. *The Journal of clinical investigation* **68**, 611-620 (1981).
23 Singleton, B. K. *et al.* Molecular and biochemical characterization of xrs mutants defective in Ku80. *Molecular and cellular biology* **17**, 1264-1273 (1997).
24 Gu, Y., Jin, S., Gao, Y., Weaver, D. T. & Alt, F. W. Ku70-deficient embryonic stem cells have increased ionizing radiosensitivity, defective DNA end-binding activity, and inability to support V(D)J recombination. *Proceedings of the National Academy of Sciences of the United States of America* **94**, 8076-8081 (1997).
25 Errami, A. *et al.* Ku86 defines the genetic defect and restores X-ray resistance and V(D)J recombination to complementation group 5 hamster cell mutants. *Molecular and cellular biology* **16**, 1519-1526 (1996).
26 Gu, J., Lu, H., Tsai, A. G., Schwarz, K. & Lieber, M. R. Single-stranded DNA ligation and XLF-stimulated incompatible DNA end ligation by the XRCC4-DNA ligase IV complex: influence of terminal DNA sequence. *Nucleic acids research* **35**, 5755-5762, doi:10.1093/nar/gkm579 (2007).
27 Andres, S. N. *et al.* A human XRCC4-XLF complex bridges DNA. *Nucleic acids research* **40**, 1868-1878, doi:10.1093/nar/gks022 (2012).
28 Pang, D., Yoo, S., Dynan, W. S., Jung, M. & Dritschilo, A. Ku proteins join DNA fragments as shown by atomic force microscopy. *Cancer research* **57**, 1412-1415 (1997).

29 Cary, R. B. *et al.* DNA looping by Ku and the DNA-dependent protein kinase. *Proceedings of the National Academy of Sciences of the United States of America* **94**, 4267-4272 (1997).
30 Roberts, S. A. *et al.* Ku is a 5'-dRP/AP lyase that excises nucleotide damage near broken ends. *Nature* **464**, 1214-1217, doi:10.1038/nature08926 (2010).
31 Strande, N., Roberts, S. A., Oh, S., Hendrickson, E. A. & Ramsden, D. A. Specificity of the dRP/AP lyase of Ku promotes nonhomologous end joining (NHEJ) fidelity at damaged ends. *The Journal of biological chemistry* **287**, 13686-13693, doi:10.1074/jbc.M111.329730 (2012).
32 Sun, J., Lee, K. J., Davis, A. J. & Chen, D. J. Human Ku70/80 protein blocks exonuclease 1-mediated DNA resection in the presence of human Mre11 or Mre11/Rad50 protein complex. *The Journal of biological chemistry* **287**, 4936-4945, doi:10.1074/jbc.M111.306167 (2012).
33 Falzon, M., Fewell, J. W. & Kuff, E. L. EBP-80, a transcription factor closely resembling the human autoantigen Ku, recognizes single- to double-strand transitions in DNA. *The Journal of biological chemistry* **268**, 10546-10552 (1993).
34 Blier, P. R., Griffith, A. J., Craft, J. & Hardin, J. A. Binding of Ku protein to DNA. Measurement of affinity for ends and demonstration of binding to nicks. *The Journal of biological chemistry* **268**, 7594-7601 (1993).
35 Walker, J. R., Corpina, R. A. & Goldberg, J. Structure of the Ku heterodimer bound to DNA and its implications for double-strand break repair. *Nature* **412**, 607-614, doi:10.1038/35088000 (2001).
36 Harris, R. *et al.* The 3D solution structure of the C-terminal region of Ku86 (Ku86CTR). *Journal of molecular biology* **335**, 573-582 (2004).
37 Lehman, J. A., Hoelz, D. J. & Turchi, J. J. DNA-dependent conformational changes in the Ku heterodimer. *Biochemistry* **47**, 4359-4368, doi:10.1021/bi702284c (2008).
38 Yoo, S., Kimzey, A. & Dynan, W. S. Photocross-linking of an oriented DNA repair complex. Ku bound at a single DNA end. *The Journal of biological chemistry* **274**, 20034-20039 (1999).
39 Zhang, Z. *et al.* The three-dimensional structure of the C-terminal DNA-binding domain of human Ku70. *The Journal of biological chemistry* **276**, 38231-38236, doi:10.1074/jbc.M105238200 (2001).
40 Aravind, L. & Koonin, E. V. Prokaryotic homologs of the eukaryotic DNA-end-binding protein Ku, novel domains in the Ku protein and prediction of a prokaryotic double-strand break repair system. *Genome research* **11**, 1365-1374, doi:10.1101/gr.181001 (2001).
41 Laskowski, R. A., Rullmannn, J. A., MacArthur, M. W., Kaptein, R. & Thornton, J. M. AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. *Journal of biomolecular NMR* **8**, 477-486 (1996).
42 Hu, S., Pluth, J. M. & Cucinotta, F. A. Putative binding modes of Ku70-SAP domain with double strand DNA: a molecular modeling study. *Journal of molecular modeling* **18**, 2163-2174, doi:10.1007/s00894-011-1234-x (2012).

43 Gell, D. & Jackson, S. P. Mapping of protein-protein interactions within the DNA-dependent protein kinase complex. *Nucleic acids research* **27**, 3494-3502 (1999).
44 Sibanda, B. L., Chirgadze, D. Y. & Blundell, T. L. Crystal structure of DNA-PKcs reveals a large open-ring cradle comprised of HEAT repeats. *Nature* **463**, 118-121, doi:10.1038/nature08648 (2010).
45 Singleton, B. K., Torres-Arzayus, M. I., Rottinghaus, S. T., Taccioli, G. E. & Jeggo, P. A. The C terminus of Ku80 activates the DNA-dependent protein kinase catalytic subunit. *Molecular and cellular biology* **19**, 3267-3277 (1999).
46 Falck, J., Coates, J. & Jackson, S. P. Conserved modes of recruitment of ATM, ATR and DNA-PKcs to sites of DNA damage. *Nature* **434**, 605-611, doi:10.1038/nature03442 (2005).
47 Weterings, E. *et al.* The Ku80 carboxy terminus stimulates joining and artemis-mediated processing of DNA ends. *Molecular and cellular biology* **29**, 1134-1142, doi:10.1128/MCB.00971-08 (2009).
48 Hammel, M. *et al.* Ku and DNA-dependent protein kinase dynamic conformations and assembly regulate DNA binding and the initial non-homologous end joining complex. *The Journal of biological chemistry* **285**, 1414-1423, doi:10.1074/jbc.M109.065615 (2010).
49 Morris, E. P. *et al.* Evidence for a remodelling of DNA-PK upon autophosphorylation from electron microscopy studies. *Nucleic acids research* **39**, 5757-5767, doi:10.1093/nar/gkr146 (2011).
50 Llorca, O. & Pearl, L. H. Electron microscopy studies on DNA recognition by DNA-PK. *Micron* **35**, 625-633, doi:10.1016/j.micron.2004.05.004 (2004).
51 Jovanovic, M. & Dynan, W. S. Terminal DNA structure and ATP influence binding parameters of the DNA-dependent protein kinase at an early step prior to DNA synapsis. *Nucleic acids research* **34**, 1112-1120, doi:10.1093/nar/gkj504 (2006).
52 Pawelczak, K. S., Andrews, B. J. & Turchi, J. J. Differential activation of DNA-PK based on DNA strand orientation and sequence bias. *Nucleic acids research* **33**, 152-161, doi:10.1093/nar/gki157 (2005).
53 Pawelczak, K. S. & Turchi, J. J. A mechanism for DNA-PK activation requiring unique contributions from each strand of a DNA terminus and implications for microhomology-mediated nonhomologous DNA end joining. *Nucleic acids research* **36**, 4022-4031, doi:10.1093/nar/gkn344 (2008).
54 Meek, K., Dang, V. & Lees-Miller, S. P. DNA-PK: the means to justify the ends? *Advances in immunology* **99**, 33-58, doi:10.1016/S0065-2776(08)00602-0 (2008).
55 Liu, S. *et al.* Distinct roles for DNA-PK, ATM and ATR in RPA phosphorylation and checkpoint activation in response to replication stress. *Nucleic acids research* **40**, 10780-10794, doi:10.1093/nar/gks849 (2012).
56 Douglas, P., Gupta, S., Morrice, N., Meek, K. & Lees-Miller, S. P. DNA-PK-dependent phosphorylation of Ku70/80 is not required for non-homologous end joining. *DNA repair* **4**, 1006-1018, doi:10.1016/j.dnarep.2005.05.003 (2005).
57 Lou, Z. *et al.* MDC1 regulates DNA-PK autophosphorylation in response to DNA damage. *The Journal of biological chemistry* **279**, 46359-46362, doi:10.1074/jbc.C400375200 (2004).

58 Goodarzi, A. A. *et al.* DNA-PK autophosphorylation facilitates Artemis endonuclease activity. *The EMBO journal* **25**, 3880-3889, doi:10.1038/sj.emboj.7601255 (2006).
59 Serrano, M. A. *et al.* DNA-PK, ATM and ATR collaboratively regulate p53-RPA interaction to facilitate homologous recombination DNA repair. *Oncogene* **32**, 2452-2462, doi:10.1038/onc.2012.257 (2013).
60 Liaw, H., Lee, D. & Myung, K. DNA-PK-dependent RPA2 hyperphosphorylation facilitates DNA repair and suppresses sister chromatid exchange. *PloS one* **6**, e21424, doi:10.1371/journal.pone.0021424 (2011).
61 Lees-Miller, S. P., Chen, Y. R. & Anderson, C. W. Human cells contain a DNA-activated protein kinase that phosphorylates simian virus 40 T antigen, mouse p53, and the human Ku autoantigen. *Molecular and cellular biology* **10**, 6472-6481 (1990).
62 Bennett, S. M., Woods, D. S., Pawelczak, K. S. & Turchi, J. J. Multiple protein-protein interactions within the DNA-PK complex are mediated by the C-terminus of Ku 80. *International journal of biochemistry and molecular biology* **3**, 36-45 (2012).
63 Sears, C. R. & Turchi, J. J. Complex cisplatin-double strand break (DSB) lesions directly impair cellular non-homologous end-joining (NHEJ) independent of downstream damage response (DDR) pathways. *The Journal of biological chemistry* **287**, 24263-24272, doi:10.1074/jbc.M112.344911 (2012).
64 Johnson, D. E. *et al.* High-throughput characterization of intrinsic disorder in proteins from the Protein Structure Initiative. *Journal of structural biology* **180**, 201-215, doi:10.1016/j.jsb.2012.05.013 (2012).
65 Yoo, S. & Dynan, W. S. Geometry of a complex formed by double strand break repair proteins at a single DNA end: recruitment of DNA-PKcs induces inward translocation of Ku protein. *Nucleic acids research* **27**, 4679-4686 (1999).
66 Meek, K., Lees-Miller, S. P. & Modesti, M. N-terminal constraint activates the catalytic subunit of the DNA-dependent protein kinase in the absence of DNA or Ku. *Nucleic acids research* **40**, 2964-2973, doi:10.1093/nar/gkr1211 (2012).
67 Ferretti, L. P., Lafranchi, L. & Sartori, A. A. Controlling DNA-end resection: a new task for CDKs. *Frontiers in genetics* **4**, 99, doi:10.3389/fgene.2013.00099 (2013).
68 Budman, J. & Chu, G. Processing of DNA for nonhomologous end-joining by cell-free extract. *The EMBO journal* **24**, 849-860, doi:10.1038/sj.emboj.7600563 (2005).
69 Nick McElhinny, S. A. & Ramsden, D. A. Polymerase mu is a DNA-directed DNA/RNA polymerase. *Molecular and cellular biology* **23**, 2309-2315 (2003).
70 Gu, J. *et al.* XRCC4:DNA ligase IV can ligate incompatible DNA ends and can ligate across gaps. *The EMBO journal* **26**, 1010-1023, doi:10.1038/sj.emboj.7601559 (2007).
71 Williams, D. R., Lee, K. J., Shi, J., Chen, D. J. & Stewart, P. L. Cryo-EM structure of the DNA-dependent protein kinase catalytic subunit at subnanometer resolution reveals alpha helices and insight into DNA binding. *Structure* **16**, 468-477, doi:10.1016/j.str.2007.12.014 (2008).

72 Dobbs, T. A., Tainer, J. A. & Lees-Miller, S. P. A structural model for regulation of NHEJ by DNA-PKcs autophosphorylation. *DNA repair* **9**, 1307-1314, doi:10.1016/j.dnarep.2010.09.019 (2010).
73 Jones, J. M., Gellert, M. & Yang, W. A Ku bridge over broken DNA. *Structure* **9**, 881-884 (2001).
74 Ma, Y. & Lieber, M. R. DNA length-dependent cooperative interactions in the binding of Ku to DNA. *Biochemistry* **40**, 9638-9646 (2001).
75 Rivera-Calzada, A., Spagnolo, L., Pearl, L. H. & Llorca, O. Structural model of full-length human Ku70-Ku80 heterodimer and its recognition of DNA and DNA-PKcs. *EMBO reports* **8**, 56-62, doi:10.1038/sj.embor.7400847 (2007).
76 DeFazio, L. G., Stansel, R. M., Griffith, J. D. & Chu, G. Synapsis of DNA ends by DNA-dependent protein kinase. *The EMBO journal* **21**, 3192-3200, doi:10.1093/emboj/cdf299 (2002).
77 Weterings, E., Verkaik, N. S., Bruggenwirth, H. T., Hoeijmakers, J. H. & van Gent, D. C. The role of DNA dependent protein kinase in synapsis of DNA ends. *Nucleic acids research* **31**, 7238-7246 (2003).
78 Orren, D. K. Werner syndrome: molecular insights into the relationships between defective DNA metabolism, genomic instability, cancer and aging. *Front Biosci* **11**, 2657-2671, doi:Doi 10.2741/1999 (2006).
79 Friedl, A. A. Ku and the Stability of the Genome. *Journal of biomedicine & biotechnology* **2**, 61-65, doi:10.1155/S1110724302201035 (2002).
80 Cui, X. P. *et al.* Autophosphorylation of DNA-dependent protein kinase regulates DNA end processing and may also alter double-strand break repair pathway choice. *Molecular and cellular biology* **25**, 10842-10852, doi:Doi 10.1128/Mcb.25.24.10842-10852.2005 (2005).
81 Carr, A. M. & Lambert, S. Replication Stress-Induced Genome Instability: The Dark Side of Replication Maintenance by Homologous Recombination. *Journal of molecular biology*, doi:10.1016/j.jmb.2013.04.023 (2013).
82 Langerak, P., Mejia-Ramirez, E., Limbo, O. & Russell, P. Release of Ku and MRN from DNA Ends by Mre11 Nuclease Activity and Ctp1 Is Required for Homologous Recombination Repair of Double-Strand Breaks. *Plos Genet* **7**, doi:ARTN e1002271DOI 10.1371/journal.pgen.1002271 (2011).
83 Rass, E. *et al.* Role of Mre11 in chromosomal nonhomologous end joining in mammalian cells. *Nature structural & molecular biology* **16**, 819-824, doi:10.1038/nsmb.1641 (2009).
84 Rahal, E. A. *et al.* ATM regulates Mre11-dependent DNA end-degradation and microhomology-mediated end joining. *Cell cycle* **9**, 2866-2877, doi:DOI 10.4161/cc.9.14.12408 (2010).
85 Belmaaza, A. & Chartrand, P. One-sided invasion events in homologous recombination at double-strand breaks. *Mutation research* **314**, 199-208 (1994).

Curriculum Vitae

Derek S. Woods

EDUCATION

Indiana University, Indianapolis IN 2008-2013
Ph.D. in Biochemistry & Molecular Biology
Minor in Cancer Biology
Advisor – John J. Turchi PhD.

Wabash College, Crawfordsville, IN 2004-2008
B.A. in Biology with Minors in Chemistry and Classics
Graduated Cum Laude

PUBLICATIONS (Peer-Reviewed)

1) Bennett SM*, **Woods DS***, Pawelczak KS, and Turchi JJ. Multiple protein-protein interactions within the DNA-PK complex are mediated by the C-terminus of Ku 80. *Int J Biochem Mol Biol.* 2012 Feb;15(1):36-45. *Authors contributed equally

2) **Woods, D** and Turchi JJ. Chemotherapy Induced DNA Damage Response: Convergence of Drugs and Pathways. *Cancer Biol and Ther*. 2013 Feb 4; 14(5) [Epub ahead of Print]

3) **Woods DS**, Sears CR, and Turchi JJ. The Mechanism of DNA Double Strand Break Differentiation by the DNA-Dependent Protein Kinase. *In Preparation.*

4) DeWalt RI, **Woods DS**, and Jalal, Sl. Mus81/Eme1: A Nuclease Complex with Implications in Cancer. *In Preparation*.

ABSTRACTS (NON-PEER REVIEWED) (*Presenting author)

International Meetings

1) **Woods DS*,** Sears CR, and Turchi JJ. Different DNA Cofactors Influence the Activation of the DNA-Dependent Protein Kinase and DNA Repair. Genomic Instability and DNA Repair: Keystone Symposia on Molecular and Cellular Biology. Banff, Alberta, Canada.

National Meetings

1) **Woods DS*,** Bennett SM, Turchi JJ. (2010) The Carboxy-Terminus of Ku80 Interactions within DNA-PK Reveals Homodimerization. The 12th Annual Midwest DNA Repair Symposium. Louisville, KY.

2) **Woods DS*,** Bennett SM, Pawelczak KS, and Turchi JJ. (2011) The Carboxy-Terminus of Ku80 Interacts with DNA-PKcs, Forms Homodimers, and Tethers DNA Termini. The 13th Annual Midwest DNA Repair Symposium. Toledo, OH.

3) **Woods DS*** and Turchi JJ. (2012) The Roles of Separate Ku80 Protein Regions in Stimulating DNA-PKcs Activity are Dependent on DNA Substrate Structure. The 14th Annual Midwest DNA Repair Symposium. Cincinnati, OH.

University Affiliated

1) **Woods DS** and Turchi JJ. (2010) DNA-PKcs/Ku80CTR Interactions in Nonhomologous End Joining. Biochemistry Research Day. Indianapolis, IN.

2) **Woods DS*,** Bennett SM, Pawelczak KS, and Turchi JJ. (2011) The Carboxy-Terminus of Ku80 Interacts with DNA-PKcs, Forms Homodimers, and Tethers DNA Termini. IU School of Medicine Cancer Research Day. Indianapolis, IN.

3) **Woods DS** and Turchi JJ. (2011) The Ku80 Carboxy Terminal Region Mediates Multiple Protein/Protein Interactions within the DNA-dependent Protein Kinase Complex: Relating Structure and Function. Department of Biochemistry and Molecular Biology Research in Progress. Indianapolis, IN.

4) **Woods DS** and Turchi JJ. (2012) The Roles of Separate Ku80 Protein Regions in Stimulating DNA-PKcs Activity are Dependent on DNA Substrate Structure. Biochemistry Research Day. Indianapolis, IN.

5) **Woods DS** and Turchi JJ. (2012) The Roles of Separate Ku80 Protein Regions in Stimulating DNA-PKcs Activity are Dependent on DNA Substrate Structure. IU School of Medicine Cancer Research Day. Indianapolis, IN.

RESEARCH ORAL PRESENTATIONS (^Selected from Abstracts)

1) **Woods DS** and Turchi JJ. (2011) The Ku80 Carboxy Terminal Region Mediates Multiple Protein/Protein Interactions within the DNA-dependent Protein Kinase Complex: Relating Structure and Function. Department of Biochemistry and Molecular Biology Research in Progress. Indianapolis, IN.

2) **Woods DS** and Turchi JJ. (2012) The Roles of Separate Ku80 Protein Regions in Stimulating DNA-PKcs Activity are Dependent on DNA Substrate Structure. Biochemistry Research Day. Indianapolis, IN.

Professional Development/Fellowships

Central Dogma Journal Club Director 2011-2013

- Organized journal club meetings between graduate students and professors from multiple departments
- Led discussions on cutting edge research from a variety of disciplines focusing on the central dogma
- Reviewed and approved papers prior to presentations

Department of Biochemistry Student Representative 2011-2012

- I attended and participated in the Department of Biochemistry faculty meetings to represent the opinion of the graduate students.
- Scheduled, organized, and executed Biochemistry Research Day in 2012
- Established and organized a Department of Biochemistry graduate student Research in Progress Lecture Series
- Invited and hosted the Department of Biochemistry Student Invited Speaker
  2012 Thomas Stossel, Professor, Harvard University
  Can Actin Cytoskeletal Research Improve Inflammation Control?

Indiana University Preparing Future Faculty Training Program
2011-2013

- Training entailed grant writing, acquiring internal funding, applying for tenure, and effective teaching strategies for undergraduate and graduate students

SpIN UP Fellowship 2012

- Fellowship provided an accelerated course on biotech companies
- Program foci included patents, licensing, startup organization, startup financing, and SBIR/STTR funding