Spring 2020

A Novel UV Resistance in Rad23-Depleted Tetrahymena thermophila

Emily M. Schmoll
Missouri State University, Schmoll97@live.missouristate.edu

A NOVEL UV RESISTANCE IN RAD23-DEPLETED *TETRAHYMENA THERMOPHILA*

A Master's Thesis

Presented to

The Graduate College of

Missouri State University

In Partial Fulfillment

Of the Requirements for the Degree

Master of Science, Cell and Molecular Biology

By

Emily Marie Schmoll

May 2020

A NOVEL UV RESISTANCE IN RAD23-DEPLETED *TETRAHYMENA THERMOPHILA*

Biomedical Sciences

Missouri State University, May 2020

Master of Science

Emily Marie Schmoll

ABSTRACT

Rad23 is a highly conserved cellular scaffold protein which participates in the nucleotide excision repair pathway and ubiquitin proteasome system. It is hypothesized that the contradictory roles of Rad23 within these two systems, acting to enhance stability or facilitate degradation respectively, could be regulated via post-translational modification of the ubiquitin-like domain of the protein. To this end, a Rad23 somatic knockout cell line was established in *Tetrahymena thermophila*, with the eventual goal of knocking in a mutant Rad23 protein lacking potential for UbL ubiquitylation. In contrast to the UV-sensitive phenotype observed in similar models, Rad23-depleted *Tetrahymena* cell lines displayed significantly increased resistance to UV irradiation. While the mechanism of this survivability remains poorly understood, this model may shed new light on potential compensatory DNA repair mechanisms, pathways of failed apoptotic induction, or additional novel roles of the Rad23 accessory protein that have yet to be explored.

KEYWORDS: Rad23, post-translational modification, ubiquitin, nucleotide excision repair (NER), ubiquitin proteasome system (UPS), UV Survivability, UV resistance, *Tetrahymena thermophila*

A NOVEL UV RESISTANCE IN RAD23-DEPLETED *TETRAHYMENA THERMOPHILA*

By

Emily Marie Schmoll

A Master's Thesis
Submitted to the Graduate College
Of Missouri State University
In Partial Fulfillment of the Requirements
For the Degree of Master of Science, Cell and Molecular Biology

May 2020

Approved:

Joshua Smith, Ph.D., Thesis Committee Chair

Amanda Brodeur, M.D., Ph.D., Committee Member

Randi Ulbricht, Ph.D., Committee Member

Julie Masterson, Ph.D., Dean of the Graduate College

ACKNOWLEDGEMENTS

I would like to thank the Biomedical Sciences department of Missouri State University, particularly the faculty within the Cell and Molecular Biology program, for the extensive support and guidance that I have received throughout my graduate research experience. I sincerely thank Dr. Joshua J. Smith, for recognizing my potential early on and instilling in me a passion for science, a confidence in myself, and the knowledge that progress, not perfection, will forever be the ultimate goal. I express deep gratitude towards both Dr. Amanda Brodeur and Dr. Randi Ulbricht for their advocacy, encouragement, wisdom, direction, and, perhaps most importantly, for acting as incredible models as I move on to find my place within the scientific community. To my family and friends, who acted as my support system throughout this process, I would not have made it to this point without you.

I dedicate this thesis to my grandparents. To my grandfather, who has always been my biggest fan, I will forever strive to resemble the person you believe me to be. To my grandmother, who is the strongest woman I have ever known, I hope to take on the world with a fraction of the grace, patience, and faith that you demonstrate every day. I love you both.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

An Overview of Rad23

Life on earth is astoundingly diverse. Even so, every known organism relies on the same basic cellular building blocks to support life. These processes are collectively referred to as the Central Dogma of Molecular Biology, and each stage can be linked, in some form, to the functions and activities of the protein Rad23.

Saccharomyces cerevisiae Rad23, referred to in humans as hHR23, is a radiation-sensitive scaffold protein comprised of 398 amino acids. Although it possesses no known enzymatic activity, it is closely involved in a number of cellular maintenance pathways including DNA repair, protein stability, and transcriptional regulation (Dantuma *et al.* 2009). Despite its diverse involvement, the mechanisms and pathways that govern Rad23 activity remain largely uncharacterized.

To generalize the Central Dogma, genetic information stored inside cells as DNA is transcribed into a coded RNA molecule. The mRNA is then translated into a functional protein made up of carefully folded amino acids, which is then able to carry out a number of roles within the cell. Because this process is so completely conserved, the importance of the dogma and its components are clear, and mutations to the DNA, faulty processing of the RNA, or misfolding of the final protein are all extremely detrimental to the well-being of the cell. For this reason, most organisms have developed methods of repairing and reversing these errors through the repair of DNA lesions, alteration of RNA transcripts and expression profiles, and destruction of aberrant proteins before they are able to accumulate and cause cellular defects. Rad23 is involved,

through unique roles, in all three mechanisms responsible for restoring and maintaining cellular stability across all lifeforms.

As will be discussed in later sections, Rad23 is a key player in the recognition step of nucleotide excision repair (NER), a process that removes bulky lesions from DNA. When it is not being used in NER, it can be found altering the expression levels of several DNA damage response genes. Simultaneously, the versatile protein works as an intracellular shuttle to deliver targeted proteins to various cellular destinations, including to the proteasome for degradation. Because this protein contains no enzymatic activity, its function is largely reliant on interaction with other proteins. Four primary domains, an N-terminal ubiquitin-like domain (UbL), two ubiquitin associated domains (UBA), and an internal Rad4 binding domain (Figure 1) determine the function of the protein as a whole (Xie, 2004). It is therefore possible that modification to Rad23 binding sites, or overall protein structure, could hold the key to distinguishing this complex mechanism.

To explore this hypothesis, it is first important to understand the overall structure of Rad23. Two UBA domains, one internal and one at the C-terminus, are responsible for binding polyubiquitylated protein substrates, typically as a prerequisite for transport to the proteasome for degradation. The internal UBA1 domain is responsible for most proteasome-targeted ubiquitin interaction, while the C-terminal UBA2 domain binds substrates with a slightly altered specificity and reduced degradation rate (Chen *et al.* 2001; Chen and Madura 2002; Bertolaet *et al.* 2001). Importantly, the UBA2 domain is vital for the stability of Rad23, and mutations in this region result in its own proteasomal degradation (Heessen *et al.* 2005). As such, both of these structures are necessary for activity within the ubiquitin proteasome system (UPS), and neither can be considered redundant or disposable.

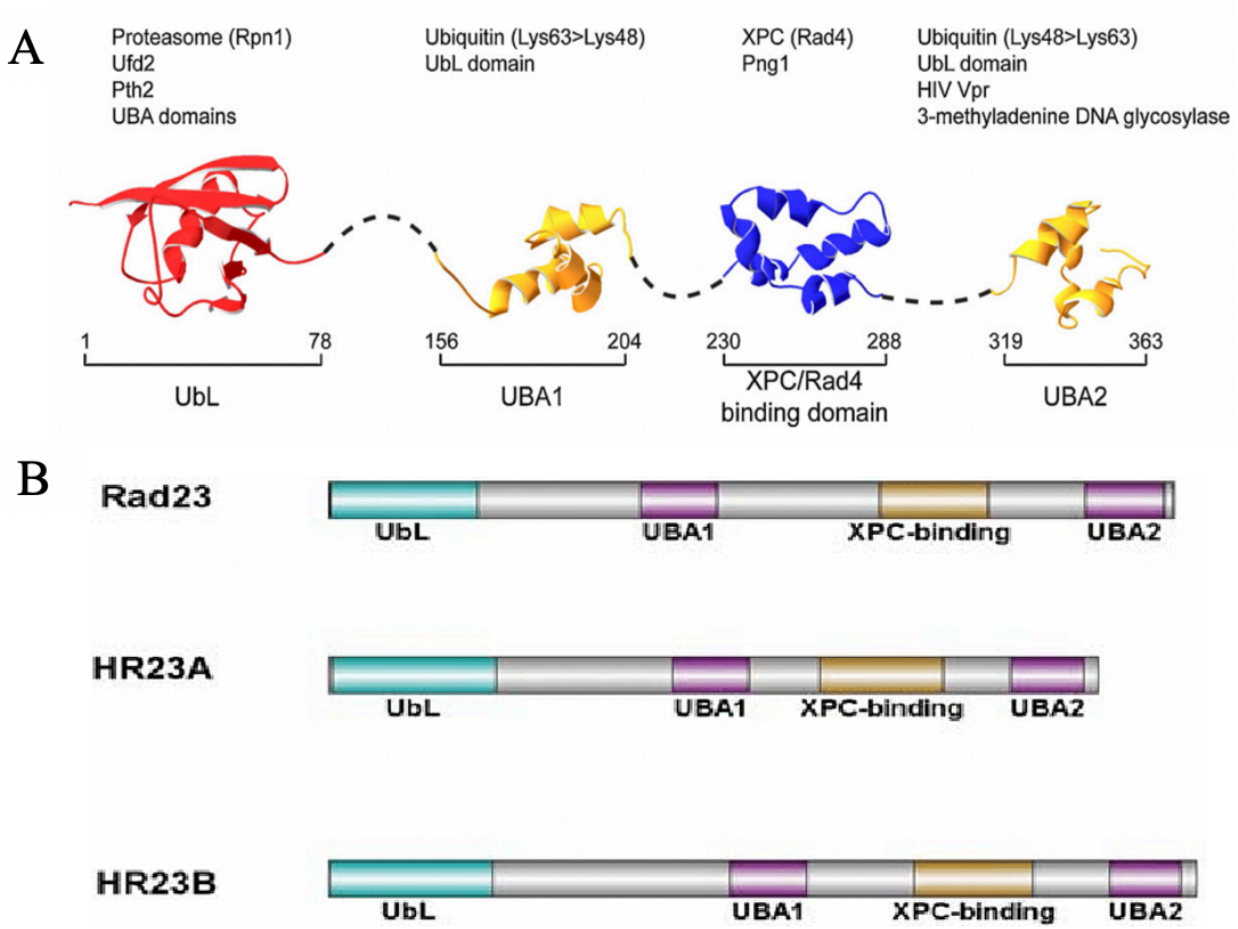


Figure 1: Functional domains in Rad23. A: Human HR23A ribbon protein structure, including the ubiquitin-like (UbL) domain, Ubiquitin Associated (UBA) domains 1 and 2, and the XPC binding domain. Interaction partners of Rad23 are shown at their respective binding locations. Ufd2 assists in ubiquitin chain assembly. Pth2 is a negative regulator of the UPS. Png1 cleaves N-glycans from ERAD substrates prior to degradation. Vpr is important for efficient infection of HIV virus. 3-methyladenine DNA glycosylases can initiate BER. **B:** The conserved domains in *S. cerevisiae* Rad23 and its human HR23 paralogues are shown for direct comparison. The four conserved functional domains are indicated. Adapted from Dantuma *et al.* 2009 and Li *et al.* 2011.

The Rad4 binding domain, located between UBA1 and UBA2, contains five helical stretches which create a hydrophobic interaction with binding partner Rad4 (Dantuma *et al.* 2009). This structure also serves as a binding site for the N-glycanase Png1 in the specialized endoplasmic reticulum associated degradation (ERAD) pathway, Apoptosis Inducing Factor (AIF) during caspase-independent programmed cell death induction, and the transcription factor E2F1 in response to DNA damage (Dantuma *et al.* 2009; Sudhakar and Chow 2014; Singh and Dagnino 2016). These variable interactions could serve as an additional mechanism by which Rad23 activity is regulated, and how the protein selectively engages in either stabilization- or degradation-centered cellular pathways.

Finally, the N-terminal UbL domain of Rad23 is capable of enhancing the efficacy of both nucleotide excision repair and the ubiquitin-proteasome system pathways (Walters *et al.* 2002). Specifically, the UbL domain facilitates the degradation of poly-ubiquitylated proteins by establishing an interaction with the 19S subunit of the 26S proteasome, allowing the polyubiquitylated substrates to be deposited into the catalytic core (Xie, 2004). Alternative interactions between the UbL and the 19S subunit inhibit the proteasome from negatively interfering with NER (Dantuma *et al.* 2009). Post-translational modification to this domain may explain the duality of these interactions, though so far, little research has been conducted on the subject.

All of these points will be expounded upon in later sections to illustrate the rationale of the experimental design. However, it is important to traverse the following sections with the context of Rad23 in mind, as to fully appreciate the progression and organization of the selected topics. Rad23 is a protein that is responsible for participating in several cellular systems, despite its simplicity and lack of enzymatic function. Its regulation is poorly understood, and a deeper

understanding of its activity could shed light on a number of vital cellular processes. This project is therefore justified, both in time and resources, due to its potential significance to both the scientific and medical communities.

DNA Repair

When DNA is damaged or altered, all downstream homeostasis mechanisms are placed at risk. Genomic order is restored by a range of DNA repair systems, each of which is specialized to respond to unique assaults to the genetic code. When functioning optimally, these pathways prevent the accumulation of deleterious modifications, but deficiencies in any system can lead to mutations, disease states, or cell death.

The natural errors incurred during DNA replication are among the most common and unavoidable threats to genome stability. DNA polymerases are capable of synthesizing DNA with high fidelity, utilizing intrinsic proofreading capabilities to reduce the number of errors introduced, but eukaryotic cells are still estimated to incorporate mutations as often as once every 10^4-10^5 bases. For a human, this means that up to one million errors are made during each cell division event (Preston *et al.* 2010).

While insertions, deletions, and mismatches are not inherently threatening to an organism, they can lead to significant changes within important genes and proteins, and therefore require a mechanism for repair (Figure 2). Mismatch Repair (MMR) is one such pathway, which restores sequences which have been improperly replicated. MMR utilizes Mut proteins to recognize and repair daughter strands of DNA that contain erroneous sequence, using the parent strand as a template. The recognition process is not entirely understood, but the daughter strand is thought to be recognized through hemimethylation patterns, residual nicks, or the eukaryotic

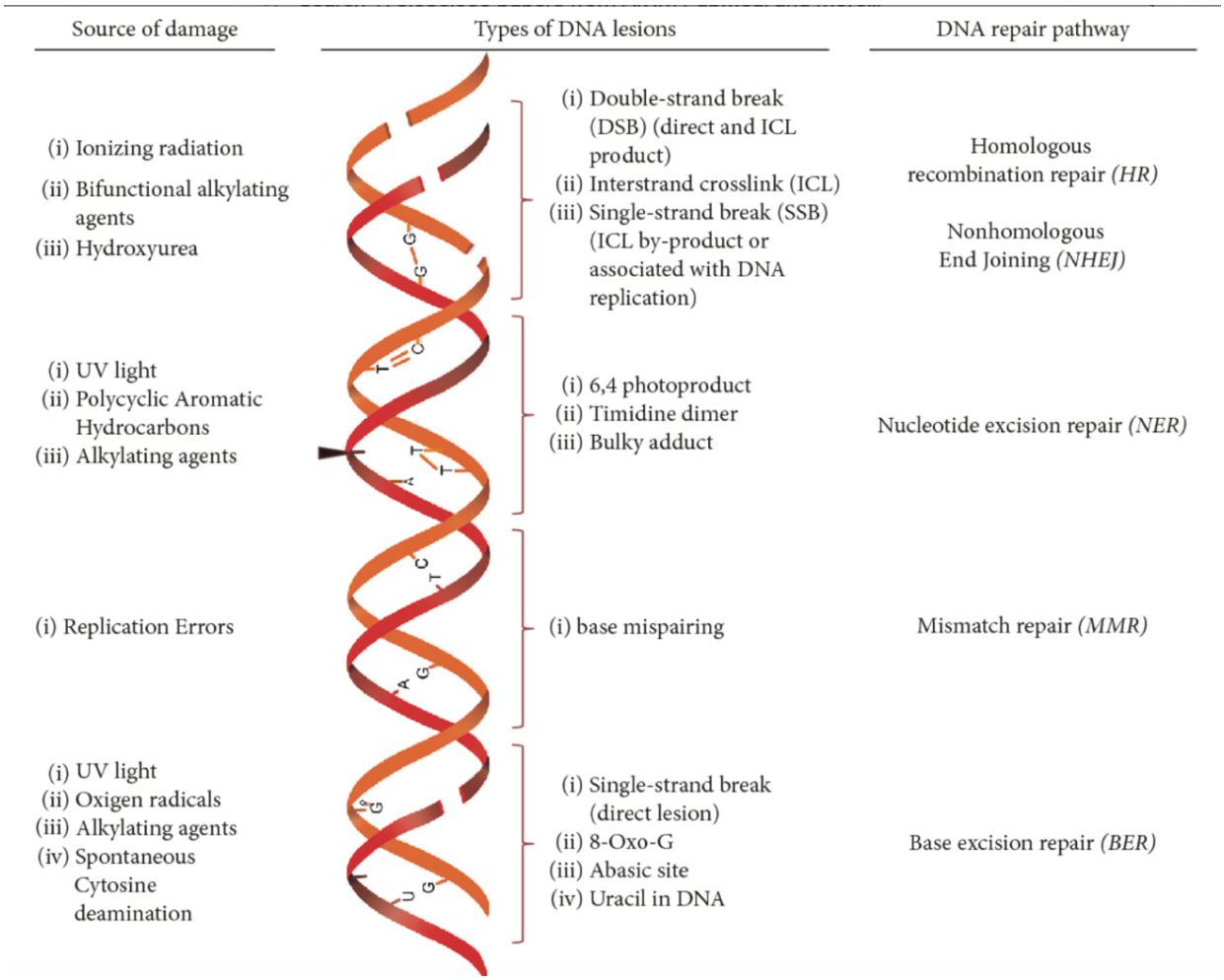


Figure 2: DNA lesions are induced by endogenous and exogenous sources. DNA damage originates from a variety of endogenous and exogenous sources, resulting in specific structural lesions. Five major DNA repair pathways have evolved in order to combat these assaults and restore genomic stability. Adapted from Garcia de Teresa *et al.* 2017.

PCNA replication accessory protein (Fleck and Nielsen 2004). Once DNA damage has been identified through MMR proteins, approximately 100 bases, up to over 1,000 bases, are removed and degraded as a part of the repair process. The daughter strand is repaired through re-synthesis of the removed bases and ligation of the DNA backbone, in a fashion similar to its original synthesis (Fleck and Nielsen 2004).

In addition to replication errors, DNA can also be endogenously or exogenously damaged through modification of the bases themselves. These lesions are corrected by the base excision repair (BER) or nucleotide excision repair (NER) pathways (Figure 2). In general, non-bulky lesions are caused by alkylation, oxidation, or deamination of individual bases, which alter the chemical and physical structure of the affected nucleotide. BER uses substrate-specific DNA glycosylases to cleave the N-glycosidic bond between damaged bases and the sugar-phosphate backbone of DNA (Fleck and Nielsen 2004). This event releases the affected nucleotide and creates an apurinic or apyrimidinic (AP) site. The single-nucleotide gap is then refilled by a polymerase using the complementary strand as a template (Fleck and Nielsen 2004). Alternatively, BER can also remove multiple nucleotides (~2-8) during long-patch BER. In this pathway, polymerases displace the affected strand and resynthesize the aberrant segment, and an endonuclease, such as FEN1, cleaves the resulting flap. In either case, the nicked DNA is sealed by a DNA ligase, restored to its undamaged state, and deleterious effects of mutation are avoided (Fleck and Nielsen 2004).

NER, which will be discussed in greater detail in a later section, is the mechanism by which cells repair helix-distorting photoproducts and bulky lesions caused by UV damage or exposure to chemicals such as Benzo[a]pyrene, a major carcinogen in tobacco smoke, or many common cancer chemotherapeutic drugs. The recognition of lesions can be accomplished by one

of two distinct pathways, but repair proceeds in a largely synonymous manner. In both cases, an NER complex comprised of transcription factor II helicase (TFIIH) and various Xeroderma Pigmentosum (XP) proteins assemble at the lesion, a 5' and 3' endonuclease create a dual incision, and a segment of ~24-32 bases is removed, resynthesized, and ligated into place. This NER pathway, along with the BER process, ensures that damage to individual bases does not threaten the state of the genomic code as a whole (Fleck and Nielsen 2004).

Finally, perhaps the most threatening form of DNA damage occurs when both strands of the double helix are broken. These lesions can result in massive loss of genetic code, which may result in cell death if left unrepaired. Homologous recombination (HR) and nonhomologous end joining (NHEJ) are two mechanisms used by cells to counteract threat of these events (Figure 2).

In HR, a homologous template, typically in the form of a sister chromosome, is used to ensure error-free repair during S and G2 phases of the cell cycle (Fleck and Nielsen 2004; Her and Bunting 2018). Each end of the double-stranded break is resected to produce a 3' overhang, the exposed single-stranded DNA invades the homologous chromosome, and new, complementary DNA is synthesized to replace the lost information. The resulting Holliday junctions, formed by base pairing between two unique strands of DNA, are resolved, and the genetic code is maintained (Fleck and Nielsen 2004).

NHEJ provides a faster, albeit less accurate, response to double-stranded breaks. This pathway is more heavily favored during G1, and relies on the ligation of two broken segments of double-stranded DNA rather than the use of a homologous template. These features facilitate quick repair of damage, but may result in the loss or translocation of some genetic code in the process (Her and Bunting 2018).

These five major pathways serve as the guardians of the genome. In their presence, DNA damage is often prevented from seriously affecting the well-being of the cell, and disease states are largely avoided. However, when a key protein or process is inhibited for any number of reasons, the mutation, physical manipulation, or loss of genetic code can prove severely damaging or fatal. For the scope of this project, the nucleotide excision repair pathway, and specifically the proteins operating within the DNA damage recognition step, will be explored in the context of these cellular consequences.

Nucleotide Excision Repair

The substrates of nucleotide excision repair consist of bulky lesions which disrupt the stability of the DNA duplex, such as cyclobutene pyrimidine dimers (CPDs) and 6-4 photoproducts (6-4PPs) (Figure 3). These structures are considered to be significantly mutagenic and carcinogenic when left unrepaired (Spivak 2015). Additionally, they arise from a variety of unavoidable cellular assaults, including UV irradiation, environmental mutagens, exogenous oxidative agents, and even cancer chemotherapeutic drugs, each of which is capable of producing a chemically and structurally unique lesion (Friedberg *et al.* 2005; Gilet and Schärer 2006). This diversity requires a highly specialized and versatile system to prevent the accumulation of these potential threats.

Much of the versatility afforded to the NER system can be attributed to the recognition steps of either the transcription coupled (TC) or global genome (GG) subsets of the pathway (Figure 4). In TC-NER, RNA Polymerase II (RNAPII) translocation along the active strand of an expressed gene is blocked by the presence of a lesion, causing the protein to stall. The blocked RNAPII recruits transcription elongation factor CSB (ERCC6), which interacts directly with the

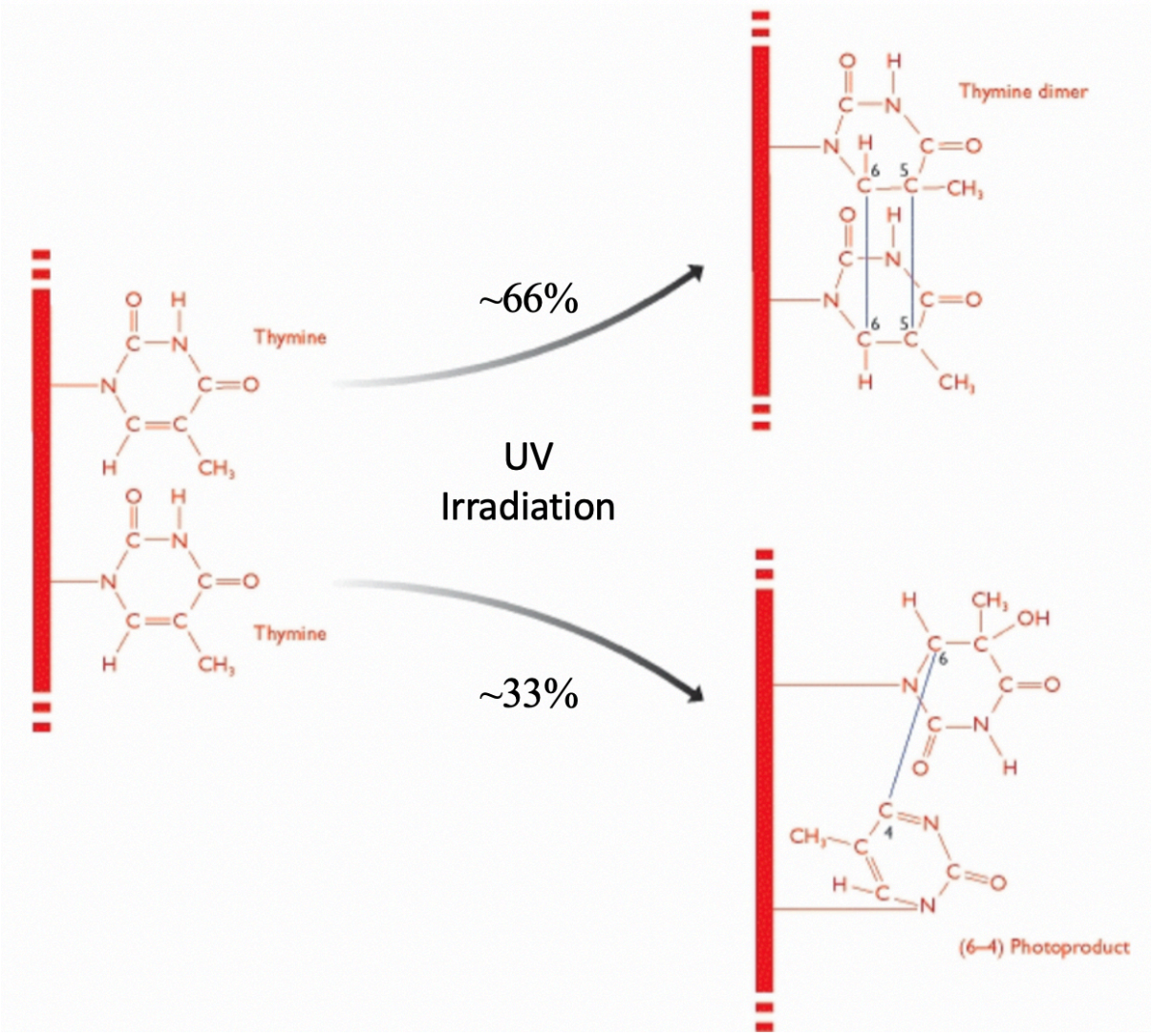


Figure 3: Formation of photoproducts by the exposure to UV rays. Adjacent thymine nucleotides are crosslinked through covalent interaction following exposure to UV irradiation. Cyclobutene pyrimidine dimers (CPDs) in which the 5th and 6th position carbons are crosslinked on each base make up approximately 2/3 of UV-induced bulky lesions. 6-4 photoproducts, in which the 4th carbon of one base is crosslinked with the 6th carbon of another, make up approximately 1/3 of UV-induced bulky lesions. Adapted from Brown 2002 with Jaspers 2001.

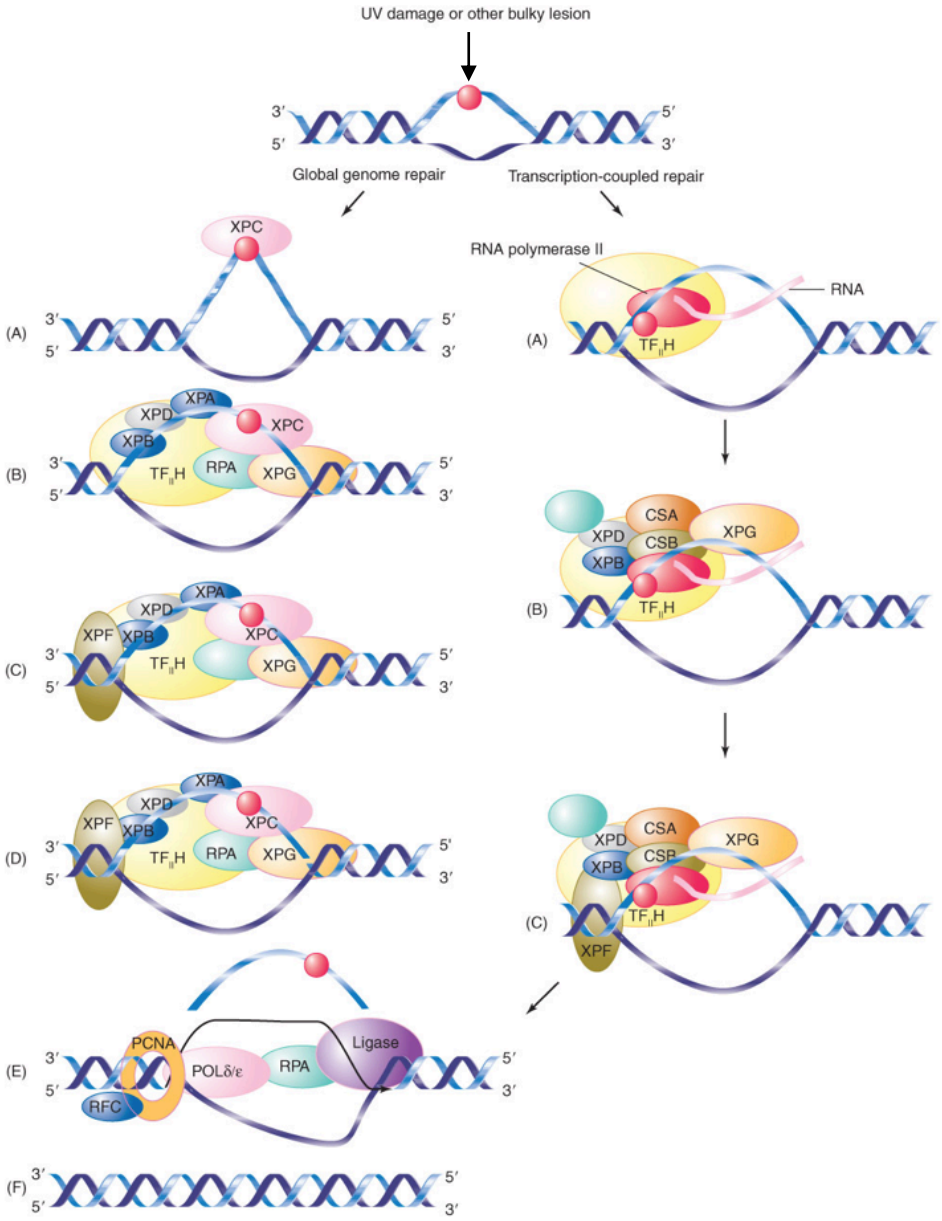


Figure 4: **Nucleotide Excision Repair Pathways.** Nucleotide excision repair proceeds via global genome and transcription-coupled repair sub-pathways, which differ primarily in the damage recognition step (**A**) and the accessory repair factors which assemble at the lesion (**B**). Both utilize XP endonucleases XPF and XPG for excision of the damaged region (**C, D**). DNA polymerase δ, ε, or κ resynthesizes the strand using the complement as a template, and DNA Ligase I or III seals the nick to repair the double helix (**E, F**). Adapted from Krebs et. al 2017.

DNA to alter the interaction between the DNA and polymerase. CSB then recruits the "master coordinator" of human TC-NER, the CSA complex (ERCC8), which in turn recruits the downstream NER factors required for excision, synthesis, and ligation (Spivak 2015). Likewise, GG-NER occurs indiscriminately across the genome, and relies on a dimer of XPC and HR23B (Rad4/Rad23) to recognize a broad spectrum of lesions and recruit the necessary repair factors. The structure of the damage, rather than the specific lesion subtype, appears to trigger detection; duplexed DNA that is thermodynamically destabilized is able to interact with a lesion- and sequence-nonspecific transglutaminase homology domain of XPC, which is then able to recruit the core NER factors to facilitate repair (Min and Pavletich 2007).

Once damage has been identified, the two sub-pathways converge upon the arrival of Transcription Factor II Helicase (TFIIH) at the lesion, a 10-subunit complex which marks the transition from recognition to repair. The TFIIH core is comprised of six proteins: XPB, p52, p8, p62, p34, and p44. The cyclin-activated kinase complex, which does not play an essential role in NER, contributes three additional players to the repair site: CDK7, cyclin H, and MAT1. A final protein, XPD, serves as a bridge between the two complexes (Compe and Egly, 2012).

Of the ten TFIIH subunits, the primary functions required for nucleotide excision repair reside within the two Xeroderma Pigmentosum (XP) family proteins, XPB and XPD. While both are classified as helicases, only the helicase activity of XPD is required for NER. This protein's function relies on the its additional ATPase activity to fulfill the role of DNA damage verification; as XPD translocates along the DNA duplex, it stalls at the lesion and therefore confirms the need for repair. Conversely, XPB uses only its ATPase activity in NER, indicating that its presence is primarily required to open the DNA and load other NER factors onto their correct positions on the DNA (Coin *et al.* 2007). Many of the other TFIIH core subunits are

involved in facilitating and supporting these assembly processes: p52 stimulates the activity of XPB, p44 stimulates XPD, and p8 assists in duplex opening and downstream recruitment of XPA (Coin *et al.* 2007; Schärer 2013; Theil *et al.* 2011; Theil *et al.* 2013). XPA is a member of the trimeric preincision complex (PIC), which assembles following lesion verification by XPD. Along with RPA and XPG, this protein cooperatively ensures that the required repair factors are present, positioned, and primed for the removal of the damaged DNA (Schärer 2013). RPA binds to the undamaged, single-stranded region of the DNA to position the endonucleases, including XPG (de Laat *et. al* 1998).

The actual excision of the lesion proceeds in a highly controlled manner. Endonuclease XPG, a member of the PIC, arrives first at the lesion following recruitment by TFIIH (Schärer 2013). The presence of XPA in the PIC facilitates the joining and incorporation of a second endonuclease, ERCC1-XPF, to trigger the dual incision at either end of the damaged oligonucleotide segment (Schärer 2013; Li *et al.* 1994; Li *et al.* 1995). A 5' incision by ERCC1-XPF produces a free 3' hydroxyl group from which new DNA can be synthesized, then a 3' incision by XPG produces a free 5' phosphate group for ligation of the new sequence (Fagbemi *et al.* 2011). With the second incision, the damaged oligonucleotide is released from the DNA duplex and degraded, and genomic stability is restored following re-synthesis and ligation (Kemp *et al.* 2012).

GG-NER Damage Recognition

The recognition step of global genome nucleotide excision repair is of particular interest for this project, and therefore must be explored in greater detail. In this pathway, the DNA damage recognition complex XPC/HR23B is largely responsible for ensuring that the genome

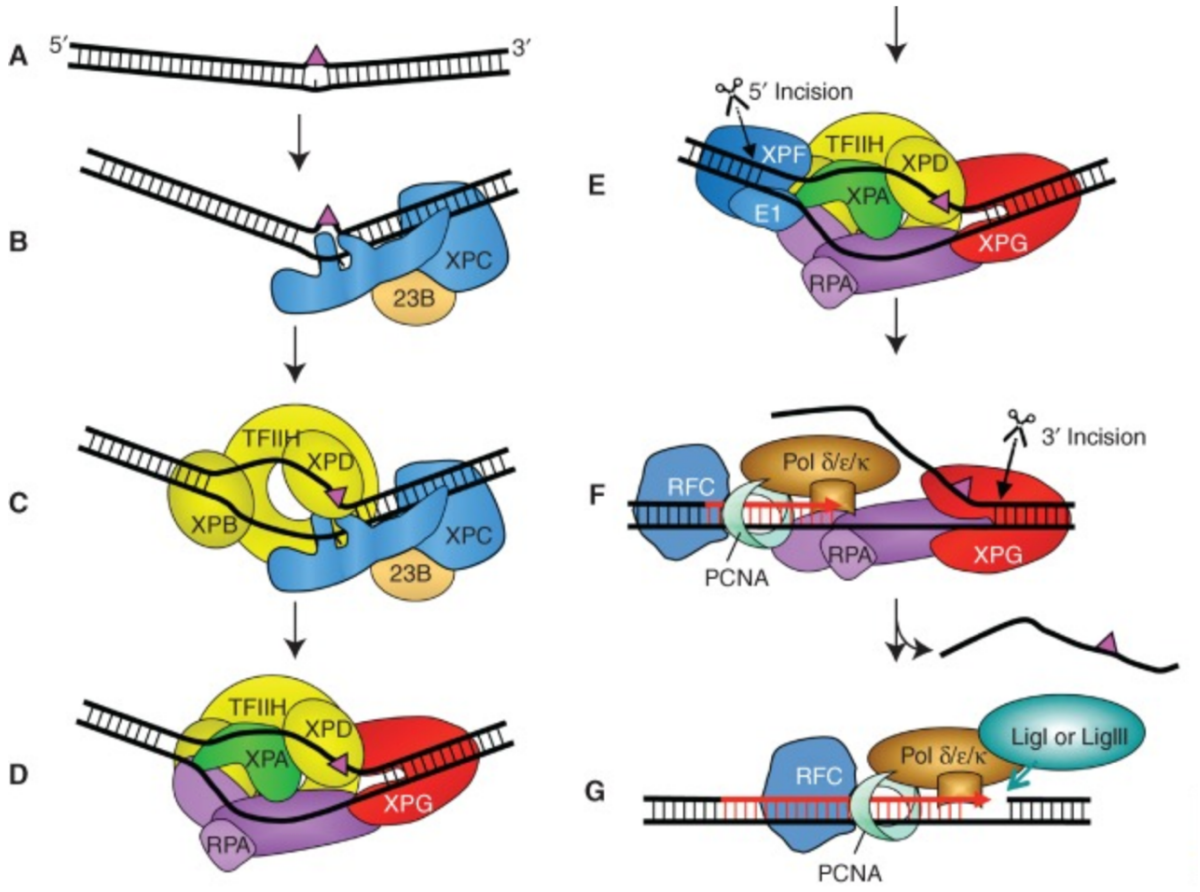


Figure 5: Model for the core GG-NER reaction. Bulky lesions (**A**) are recognized by the XPC-HR23B heterodimer ,which binds to the undamaged strand opposite the lesion and induces a kink to the DNA (**B**). XPC recruits TFIIH to open the helix and allow assembly of XPD for lesion verification and XPB for further assembly (**C**). The preincision complex of XPA, RPA, and XPG assemble at the lesion. Rad23 has dissociated (**D**). ERCC1-XPF interacts with XPA at the lesion to nick the DNA upstream (5') of the damage (**E**). DNA polymerases begin to resynthesize the strand using the undamaged complement as a template, and XPG induces the 3' nick (**F**) to release the damaged strand (**G**). The resynthesized strand is sealed by DNA ligase I or III, and genomic stability is restored (**G**). Adapted from Schärer 2013.

remains stable and free from damage (Figure 5). While TC-NER is capable only of correcting damage on actively transcribed genes, the proteins involved in GG-NER are responsible for constitutively scanning all regions of the genome, to prevent the accumulation of lesions (Schärer 2013). Strictly speaking, XPC, also known as Rad4, is solely responsible for the detection of lesions in GG-NER. A lesion- and sequence-nonspecific transglutaminase homology domain, alongside a beta-hairpin domain (BHD1), work together to anchor Rad4 at the site of damage (Min and Pavletich 2007). Two additional beta-hairpin domains (BHD2/3) bind the undamaged strand of DNA, where the lesion increases the ssDNA character of its complement (Schärer 2013). Together, the beta-hairpin domains infiltrate the helix and "flip" the damaged bases out and away, creating space for the downstream repair factors to assemble at the target site (Min and Pavletich 2007; Janićijević *et al.* 2003).

While XPC is the active member of the recognition step, it relies on the assistance of three other proteins to fulfill its role. The most significant contributor, HR23, is also known as Rad23. This protein boasts no catalytic activity, but interacts directly with XPC to provide protection against proteasomal degradation; steady-state levels of XPC/Rad4 are consistently depleted in the absence of its HR23/Rad23 interaction (Bergink *et al.* 2012). Additionally, HR23/Rad23 stimulates the binding of XPC to lesions prior to its dissociation (Bergink *et al.* 2012). The intricate relationship between these heterodimeric proteins will be explored more thoroughly in a later section.

In addition to the classic Rad4/Rad23 recognition dimer, two additional proteins contribute to the targeting, facilitation, and stimulation of GG-NER lesion recognition. Centrin-2 (CEN2), a member of the calcium-binding calmodulin family of proteins, stimulates nucleotide excision repair when in complex with XPC and HR23 (Nishi *et al.* 2005). DNA damage binding

protein DDB2 helps target XPC/HR23 to sites of difficult-to-identify lesions, and to increase the access of the proteins to inaccessible areas of the genome such as chromatin (Schärer 2013).

While it has been established that NER is responsible for the removal of CPDs, these structures, which crosslink two neighboring thymine or cytosine bases, do not dramatically distort the DNA duplex (Schärer 2013). As discussed previously, helix destabilization is responsible for XPC recognition. XPC is therefore less likely to identify the lesion independently and relies on DDB2 to recruit the recognition complex to the site of UV damage (Fitch *et al.* 2003).

DDB2 has a high affinity for UV damaged DNA, and possesses a shallow, hydrophobic binding site for interaction with CPDs and 6-4PPs (Scrima *et al.* 2008; Fischer *et al.* 2011; Yeh *et al.* 2012). The wedge-like structure of DDB2 inserts into the helix at the site of the lesion and extrudes the two damaged bases, compressing the phosphate groups of the nucleotides and creating a 40% "kink" in the DNA capable of triggering recognition by XPC (Scrima *et al.* 2008). In chromatin, where the DNA is condensed and largely inaccessible to DNA repair factors, DDB2 acts in complex with CUL4-ROC1 (RBX1) to ubiquitylate XPC, histones, and DDB2 itself (Schärer 2013). These modifications lead to destabilization of histones and DDB2, but increase XPC binding to target DNA (Sugasawa *et al.* 2005; Schärer 2013). This process, commonly known as the "access-repair-restore" model, allows damaged DNA to be relaxed from its supercoiled state, accessed by repair factors, then returned to its native conformation, in an elegant and efficient cycle of genome maintenance (Smerdon and Lieberman 1978).

In-Depth View of the Rad4/Rad23 Complex in NER

The XPC/HR23 heterodimer, henceforth referred to solely as Rad4/Rad23, is responsible for recognizing DNA damage within the GG-NER pathway (Bergink *et al.* 2012). Under normal physiological conditions, Rad23 is expressed in 10-fold excess to Rad4 (Ghaemmaghami *et al.* 2003), and therefore the latter protein is rarely observed in the absence of its accessory. The exception to this rule occurs directly following Rad4 binding to bulky lesions, which causes Rad23 to dissociate from the complex and return to nuclear or cytoplasmic protein pools. Rad23 therefore has no direct role in the excision or repair of the DNA lesions beyond that of the recognition step (Bergink *et al.* 2012).

The Rad4/Rad23 complex scans the eukaryotic genome for structural distortions, transiently binding and releasing the undamaged DNA of the genome. Upon recognition of a lesion, Rad4 binds the undamaged complement strand and further distorts the double helix through the insertion of a beta hairpin domain. The hairpin structure "flips" the damaged bases out of the helical structure and creates space for a helicase to complex and further melt the DNA (Min and Pavletich 2007; Janićijević *et al.* 2003). This process occurs selectively in helix-distorting lesions (Sugasawa *et al.* 2001), and it is hypothesized that the recognition of the distortion itself, rather than the recognition of specific crosslinked nucleotides, is the catalyst (Maillard *et al.* 2007).

While Rad23 remains in complex as Rad4 scans the genome, three major hypotheses exist regarding the sequence of events occurring following lesion identification. First, it is possible that Rad4 tightly binds the lesion while still complexed with Rad23, which results in a conformational change that lowers its binding affinity for the accessory and results in Rad23 dissociation (Bergink *et al.* 2012). Second, Rad23 may dissociate immediately after Rad4

damage recognition, which could facilitate tight binding of Rad4 to the lesion and confirm CPD identification (Bergink *et al.* 2012). Finally, it has been proposed that TFIIH, the factor immediately downstream of Rad4 (Sugasawa *et al.* 2005), may disrupt the Rad4/Rad23 complex due to its translocation activity. This disruption would directly result in the release of Rad23 from the complex (Bergink *et al.* 2012). In any case, upon later dissociation of Rad4 from the newly repaired DNA, it rapidly re-associates with Rad23 and continues scanning the genome until another lesion is identified and the process repeats. Through this cycle, Rad4 and Rad23 provides a level of resistance to UV irradiation and other environmental assaults.

The role of Rad23 in the recognition complex has been deduced from a number of experiments in recent years. Rad4 knockout cells present with greater UV sensitivity than Rad23 knockouts, though steady-state levels of Rad4 are consistently depleted in the absence of its accessory (Xie, 2004). The reduced Rad4 levels, in turn, lead to a slower initiation of NER and a reduced rate of cyclobutene pyrimidine dimer (CPD) removal in damaged cells, indicating the importance of Rad23 in stabilizing Rad4. *In vitro*, however, Rad4 is maintained even in the absence of Rad23, and NER is adequately and successfully initiated where the necessary downstream proteins are present (Xie, 2004). The introduction of exogenous Rad4 into a Rad23-depleted cell line which *also* lacks Rad4 ubiquitylation proteins (Ubc4/5) also restores UV resistance to wild-type levels (Ortolan 2004).

In the absence of Rad23, Rad4 is likely labeled as aberrant or misfolded *in vivo*, and is therefore shuttled to the proteasome for degradation (Sherman and Goldberg 2001; Ng 2003). The Rad4 binding domain of Rad23 alone is capable of preventing this process (Ortolan 2004). Rad23, despite a complete lack of enzymatic activity or DNA binding capabilities, is therefore an

integral part of maintaining UV survivability through the protection of Rad4, potentially by means of preventing Rad4 ubiquitylation.

In addition to the protective effects of Rad23, its presence in complex with Rad4 has been shown to stimulate transient DNA interactions, as well as tight binding of NER-susceptible DNA lesions (Xie, 2004). The heterodimer has been shown to display a higher affinity for damage than Rad4 alone, which also increases the rate of CPD removal and DNA repair *in vivo* (Xie, 2004).

Rad23 knockout models maintain some resistance to UV irradiation, but are deficient in NER, increasing mutation rate, and more sensitive to UV than their wild-type counterparts. Overexpression of Rad4 protein in these models does not readily rescue NER-sufficient phenotypes, which indicates that the role of Rad23 extends further than simply stabilizing Rad4 in the DNA damage recognition complex (Xie, 2004). Instead, Rad23 appears to be an important regulator, and even stimulator, of Rad4 activity within the GG-NER pathway (Sugasawa *et al.* 1996).

A deficiency in either the Rad4 or Rad23 proteins significantly increases sensitivity to UV irradiation, leading to a number of adverse effects and phenotypes (Bergink *et al.* 2012). With that said, Rad4 and its homologues have been far more extensively studied due to their direct implications with the XP disease state discussed later. It is important, however, to understand precisely how this complex functions and is regulated *in vivo*, and therefore greater attention must be paid to the roles of Rad23 if the pathway is ever to be fully understood and characterized.

Deficiencies in NER and Clinical Significance

Nucleotide excision repair is responsible for reversing damage caused primarily by exogenous assaults to the genome, which can result in genetic mutations and disease states if impaired. Both CPDs and 6-4PPs are crosslinks formed between pyrimidine bases on a shared DNA strand, and arise due to exposure to harmful UV irradiation from sun exposure (de Boer and Hoeijmakers 2000). Chemicals in cigarette smoke, such as Benzo[a]pyrene, can lead to the accumulation of polycyclic aromatic hydrocarbon adducts on DNA, and even cancer chemotherapeutic drugs, such as Cisplatin and Bleomycin, can cause inter-strand crosslinks that prevent the proper replication and transcription of DNA (de Boer and Hoeijmakers 2000, Laczmanska *et al.* 2007). Because these assaults are often unavoidable in everyday life, and can lead to syndromes as severe as cancer in even the healthiest of individuals, the nucleotide excision repair pathway is a relevant and clinically significant area of research.

In addition to the risk of carcinogenesis in healthy individuals, three rare, genetically recessive conditions are linked to deficiencies in nucleotide excision repair: Cockayne Syndrome (CS), Trichothiodystrophy (TTD), and Xeroderma pigmentosum (XP) (Table 1). CS and TTD are both linked to transcriptional defects, and present with UV-sensitive phenotypes without a marked increase in cancer risk. Conversely, XP is characterized by an extreme sensitivity to UV light, and a 1,000-2,000-fold increase in overall lifetime risk of cancer (Schärer 2008, Velez-Cruz *et al.* 2013).

Cockayne syndrome is caused by a mutation in either the CSA or CSB protein (see section: Nucleotide Excision Repair). The malfunction of these proteins results in severe neurological abnormalities and physical deficiencies, including photosensitivity, mental retardation, skeletal abnormalities, neurological degeneration, delayed mental or psychomotor

Table 1: Clinical features of patients with XP, TTD, CS, or XP/CS.

*Genetic Syndrome	Clinical Features	Mutated Genes
XP	Sun hypersensitivity, >1,000-fold cancer susceptibility, abnormal skin pigmentation	*XPA-G and XPV*
TTD	Brittle hair, mental retardation, ataxia, tremors	*TTDA, XPD*
CS	Sun hypersensitivity, cachectic dwarfism, severe mental retardation, skeletal and retinal abnormalities, segmental progeria	*CSA, CSB*
XP/CS	Extreme sun hypersensitivity, hair thinning and freckling, neurological development abnormalities, severe mental retardation, abnormal skin pigmentation, high cancer proneness, segmental progeria	*XPB, XPD, XPG*

*Syndromes caused by NER protein mutations or deficiencies in humans: Xeroderma Pigmentosum (XP), Trichothiodystrophy (TTD), and Cockayne Syndrome (CS). Adapted from Velez-Cruz *et al.* 2013.

development, myelination abnormalities, hearing loss, short stature, thin hair, cataracts, and tooth decay (Schärer 2008; de Boer and Hoeijmakers 2000). While symptoms may vary from patient to patient, the average lifespan of affected individuals is only 12.5 years of age, with the primary causes of death often attributed to secondary respiratory infections (Nance and Berry 1992).

Trichothiodystrophy, like Cockayne syndrome, is not characterized by an increased risk of carcinogenesis, but comes with its own slew of clinical identifiers. TTD patients suffer from a spectrum of photosensitivity, mental retardation, physical abnormalities, ichthyotic skin, neurodevelopmental delay, and, most notoriously, sulfur-deficient brittle hair (Schärer 2008; de Boer and Hoeijmaker 2000).

Xeroderma pigmentosum is the condition most relevant to GG-NER, and specifically the Rad4/Rad23 recognition complex. Colloquially referred to as “children of the moon disease”, patients afflicted with XP present as early as two years of age with extreme sensitivity to the sun, and are required to take extensive precautions when venturing outdoors in the daylight (Kraemer 1997). These patients are prone to freckling and blistering; what would be considered tolerable UV doses for a healthy individual significantly increases their risk of malignant tumor development. Affected individuals also commonly experience progressive neurological aberrations (~18%), delayed cognitive development, and premature aging, as well as optical tumors and general photophobia (Berneburg and Lehmann 2001; Bootsma 2001).

As is the case for CS and TTD, XP can be caused by mutations in one of several proteins within the NER pathway, all of which are classified specifically as XP proteins. Patients that express mutations in XP proteins A through G, including damage recognition protein XPC, experience a strong predisposition to the development of basal cell and squamous cell

carcinomas, with a slightly increased rate of melanoma development (de Boer and Hoeijmaker 2000). These lesions appear most commonly on the sun-exposed skin of the face, neck, head, and even tongue. At only eight years old, the average age of first onset is approximately fifty years earlier than that of the general population (Kraemer 1997). Many XP patients do ultimately succumb to either these cutaneous carcinomas or later development of internal cancers, which reduces their lifespan by an average of 30 years (de Boer and Hoeijmaker 2000).

The conditions associated with NER deficiency are severe and debilitating, and even genetically unaffected individuals are put at risk by everyday exposure to UV irradiation, chemical carcinogens, and even FDA-approved chemotherapeutic drugs. It is therefore important to pursue further research efforts in this area, and to look deeper into the Rad4/Rad23 recognition complex as a potential target for the milder forms of XP and UV sensitivity. By fully mapping these relationships and pathways, and through a fuller understanding of the regulatory checkpoints that governs them, better preventative care and treatment can be developed.

The Ubiquitin Proteasome System

The ubiquitin proteasome system (UPS) is a constitutive eukaryotic regulatory pathway responsible for degrading and recycling the majority of all endogenous intracellular proteins (Hershko and Ciechanover 1998; Rock *et al.* 1994). Just like DNA repair, protein degradation pathways ensure that aberrant or mutagenic biomolecules do not accumulate and cause damage within the cell, maintaining genetic and physiological stability. The UPS in particular is essential in maintaining regulation of cell cycle progression, programmed cell death, differentiation, intracellular signaling, stress and immune response, gene transcription, protein quality control, organelle biogenesis, and DNA damage recognition and repair (Sherman and Goldberg 2001;

Lecker *et al.* 2006; Glickman and Ciechanover 2002). In short, this pathway ensures that the cell is capable of quickly and effectively acclimating to physiological and environmental changes.

As the name suggests, there are two key players in the ubiquitin proteasome system. Ubiquitin, discussed in greater detail in a later section, is a highly conserved protein modifier that is covalently linked to proteasomal substrates to target them for degradation (Lecker *et al.* 2006). The proteasome, a multi-subunit complex (26S), is divided into two major functional units: the 19S regulatory particle and the 20S catalytic core (Lecker *et al.* 2006). These two major components, ubiquitin and the 26S proteasome, participate in a highly regulated, stepwise pathway to selectively remove target substrates from circulation within the cell (Glickman and Ciechanover 2002).

The first step in UPS is the tagging of the target substrate with a single, or more often a chain, of ubiquitin protein moieties (Figure 6), which allows the substrate to be recognized by the 26S proteasome (Glickman and Ciechanover 2002; Lecker *et al.* 2006). The linkage of these monomers is processed by a team of three classes of enzymes: a ubiquitin activating enzyme (E1), a ubiquitin conjugating enzyme (E2), and a ubiquitin ligase (E3). In short, a nonspecific E1 utilizes ATP to form a high-energy thioester bond with ubiquitin, which is then transferred to one of several cellular E2 molecules. The E2, which is specific for the substrate protein, forms its own thioester bond with ubiquitin, then transfers the bond to the E3-bound target substrate (Glickman and Ciechanover 2002). In this way, E2s are largely responsible for the regulation of the ubiquitylation process, as they bridge nonspecific ubiquitin priming with the specific patterns of fate-determining chain formation (Ye and Rape 2009). The E3 adaptor molecule creates the interactions between substrates and the ubiquitin molecules, facilitating covalent bond formation.

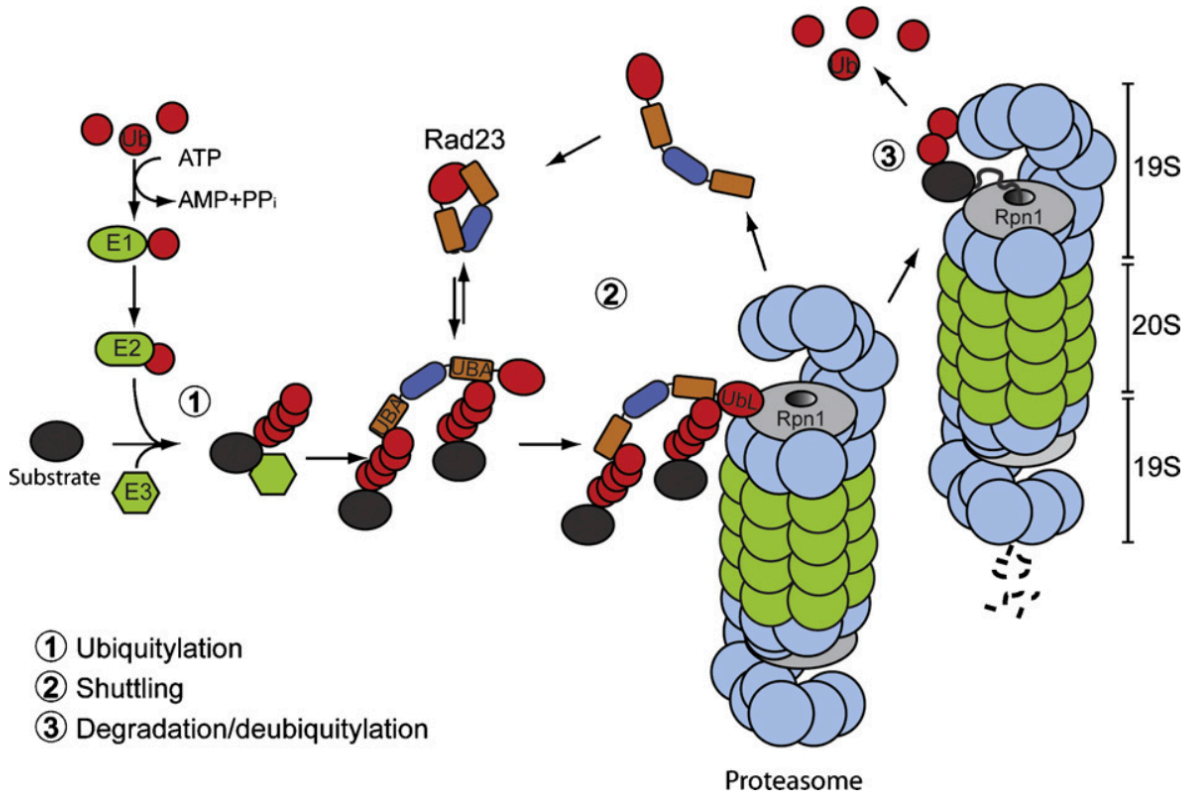


Figure 6: The role of Rad23 in ubiquitin-dependent proteasomal degradation. The UPS can be divided into two stages: ubiquitylation (**left**) and proteasomal degradation (**right**). Substrates targeted for degradation are ubiquitylated in a cascade involving an E1 ubiquitin activating enzyme, and E2 ubiquitin conjugating enzyme, and an E3 ubiquitin ligase, which results in a chain of 4-6 K48 linked ubiquitin conjugates (**1**). Ubiquitin shuttle proteins, such as Rad23, DDI1, or Dsk2, associate with the ubiquitin chains through ubiquitin associated domains. After translocating to the proteasome with the cargo, the ubiquitin-like domain of Rad23 interacts with Rpn1, a member of the 19S regulatory subunit of the proteasome (**2**). Following deposition of the substrate into the 20S core of the proteasome, Rad23 and free ubiquitin molecules are released and the deubiquitylated substrate is unfolded and degraded into its amino acid components (**3**). Adapted from Dantuma *et al.* 2009.

A specific E3 subset, known as E4, covalently links additional ubiquitin molecules to form the final poly-ubiquitin chain (Glickman and Ciechanover 2002).

Importantly, the substrates susceptible to ubiquitin targeting are widely diverse in structure and function, and their half-lives range from only minutes, to days, to even weeks. (Lecker *et al.* 2006). As such, E2 and E3 molecules must identify and interact with a wide variety of substrate proteins, and the regulation of the pathway must be tightly controlled to prevent inadvertent and non-specific degradation of cellular proteins. When a target protein is identified, the specific ubiquitylation cascade presents a clear signal to shuttling proteins that the substrate is fated for relocalization to the proteasome. This redirection constitutes the second major step in the UPS. When the substrate arrives, the third major step in the UPS, proteasomal degradation, is initiated.

Polyubiquitylated proteins first interact with the 26S proteasome via the 19S regulatory particle, which caps the barrel-shaped 20S catalytic core (Figure 6). Once the regulatory particle recognizes the polyubiquitin degradation signal, ATPases in the base of the subunit begin to unfold the secondary and tertiary structures of the target protein to facilitate its translocation into the catalytic core (Benaroudj *et al.* 2003). These ATPases are also thought to act as a "key" in opening a gated channel into the 20S core particle, allowing the substrate to enter (Smith *et al.* 2005).

The 20S catalytic core is composed of four stacked rings with a hollow center: two inner beta rings capped by two outer alpha rings. Once the substrate enters into the core, six catalytic protease sites along the inner face of the chamber cleave the substrate protein into fragments 3 to 25 amino acids long, which are then expelled from the proteasome (Kisselev *et al.* 1999). The resulting peptides are quickly degraded into individual amino acids by cytosolic peptidases,

which can be utilized for further protein synthesis. Ubiquitin and the shuttles are released and recycled for further use in the cell, and the life cycle of intracellular proteins is completed.

The Ubiquitin Code

Ubiquitin is a stable, highly conserved protein that is added post-translationally to thousands of protein substrates within the cell. Depending on the context and organization of the molecules added, ubiquitylation can alter protein stability, localization, expression, interaction, and activity (Figure 7). This diversity illustrates the need for a comprehensive ubiquitin code system (Kommander and Rape 2012). This code is highly relevant to this project, as it ties in closely with both the interaction partners of Rad23, as well as the regulation of Rad23's activity as a whole.

Ubiquitin is composed of seventy-six amino acids, nine of which are involved in the formation of ubiquitin chains. When a substrate is targeted for ubiquitylation, enzymes involved in the ubiquitylation cascade (see The Ubiquitin Proteasome System) covalently link the C-terminal glycine residue of a ubiquitin molecule to an internal lysine amino group in the substrate molecule, resulting in a monoubiquitylated protein (Glickman and Ciechanover 2002). If the substrate contains additional exposed lysine residues, mono-ubiquitylation can occur at multiple locations, resulting in a multi-monoubiquitylated protein (Kommander and Rape 2012).

In addition to the C-terminal glycine, ubiquitin contains an N-terminal methionine and seven internal lysines that can be targeted for ubiquitin chain formation. Polyubiquitylation most commonly occurs through the linkage of four to six ubiquitin molecules in either a linear or branched configuration (Dantuma *et al.* 2009). The identity of the lysine residue through which the chain is elongated (K11, K48, K63, etc.) is largely responsible for dictating the cellular

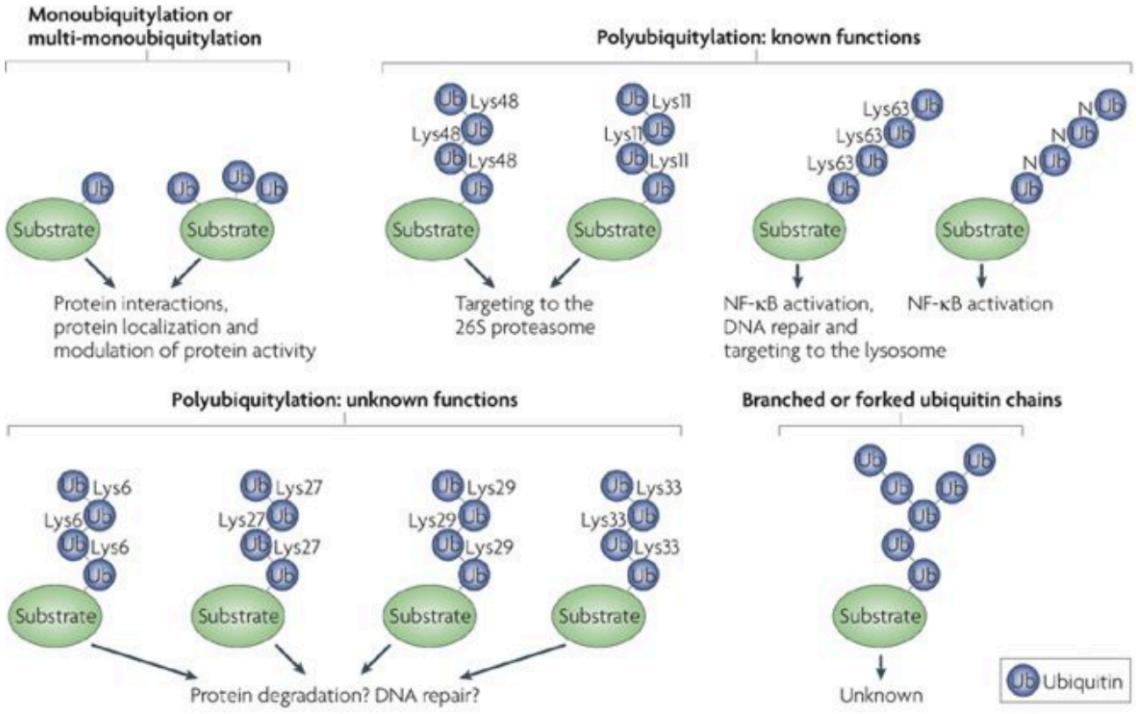


Figure 7: Ubiquitylation Patterns and Cellular Outcomes. The number and linkage pattern of ubiquitin molecules attached to a protein substrate can largely influence its fate in a cell. Monoubiquitylation and multi-monoubiquitylation are often associated with altering protein localization, activity, or interactions. K48- and K11-linked chains are often associated with proteasomal degradation, where K63- and N-terminal linkages are primarily linked to NF-κB activation. K63-linked chains can also mediate the recruitment of substrate interaction partners, which can initiate other downstream pathways. Some of the additional observed chain types, including K6, K27, K29, K33, and branched chains, are yet to be fully characterized, and may contribute to a number of additional pathways and activities. Adapted from Ye and Rape 2009.

outcome of the substrate (Kommander and Rape, 2012).

Monoubiquitylation and multi-monoubiquitylation have been shown to alter cellular transcriptomes, trigger internalization and degradation of cell surface proteins, and alter the localization and activity of nuclear and cytosolic substrates (Hicke and Dunn 2003; Sigismund *et al.* 2004; Ye and Rape 2009; Kommander and Rape 2012). Ubiquitin chains formed through Lys6, Lys27, Lys29, and Lys33 are poorly characterized, as are branched or forked ubiquitin chains. However, the effects Lys11, Lys 48, and Lys63 have been extensively studied and characterized over the past several years (Ye and Rape 2009).

The Lys11- and Lys48-linked ubiquitin chains have both been associated with proteasomal targeting and substrate degradation, whereas Lys63-linked substrates are often targeted to the lysosome for degradation (Ye and Rape 2009). Lys63-linked chains are also capable of regulating protein-protein interactions and protein activity, as observed in the activation pathway of NF_KB immune response (Kommander and Rape 2012). In short, the pattern of ubiquitin linkage is a major determinant in altering cellular outcomes.

The counterpart to the ubiquitin code is the set of proteins that recognize, translate, and implement the changes dictated by the encryption. Proteins that contain ubiquitin binding domains (UBDs), such as ubiquitin associated domains (UBAs) or ubiquitin interaction motifs (UIMs), are believed to recognize features such as chain flexibility, molecular spacing, and other structural signatures that provide context into the meaning of the codes (Kommander and Rape 2012). For example, Lys11- and Lys48-linked ubiquitin chains tend to adopt “compact” formations, in which adjacent ubiquitin molecules interact closely with one another; Lys63-linked chains tend to form more “open” conformations which interact solely through the linkage site (Kommander and Rape 2012). The structures of the UBD-containing ubiquitin decoder

proteins are believed to facilitate interaction with only a particular subset of substrates, such as the interaction between Lys48-linked modified proteins and the ubiquitin shuttle protein Rad23. This specificity provides a basis for purposeful interaction and deliberate signaling, which results in an efficiency for instituting cellular response.

Rad23: A Ubiquitin Shuttle

Substrates of the ubiquitin proteasome system are only able to interact directly with the ubiquitin receptors within the 19S proteasome when they are in close physical proximity (Madura 2004). However, many cellular proteins do not naturally reside in close proximity to the proteasome, and must still be cleared in order to maintain cellular stability and homeostasis. To resolve this conflict, ubiquitin shuttling proteins, which have the ability to interact directly with ubiquitin chains and proteasomal ubiquitin receptors, are used to relocate targeted proteins and deposit them into the 26S proteasome (Zientara-Rytter and Subramani 2019)

The properties and composition of Rad23's structural domains (See An Overview of Rad23) allow it to claim membership within a small family of proteasomal shuttle proteins, alongside siblings Ddi1 and Dsk2 (Godderz *et al.* 2017). The three scaffolds contain common N-terminal ubiquitin-like (UbL) domains and C-terminal ubiquitin associated (UBA) domains, which allow them to interact with both polyubiquitylated substrates and the ubiquitin receptors of the proteasome (Nowicka *et al.* 2015). While these three proteins share similar molecular structures, they show distinct intermolecular interactions and substrate preferences within the UPS (Godderz *et al.* 2017).

All three ubiquitin shuttles bind K48-linked polyubiquitylated proteins to their C-terminal ubiquitin-associated (UBA) domains. Once substrates are bound, the shuttles translocate to and

interact with the 19S regulatory subunit of the proteasome via their N-terminal ubiquitin-like (UbL) domains, and deposit their binding partners into the 20S catalytic core to initiate proteolysis (Nowicka *et al.* 2015; Heesen *et al.* 2005). The UbL domain is simultaneously responsible for protecting the shuttle from degradation itself, allowing it to be released safely back into the cytosol to continue the cycle (Bertolaet *et al.* 2001; Xie, 2004).

Rad23 has significant structural and functional homology with Dsk2, and the two scaffolds have been shown to compete for interaction with the ubiquitin receptor Rpn10 and proteasomal subunit Rpn1 to deposit substrates. Ddi1 is more distinct in its preferences, and more distantly related structurally (Nowicka *et al.* 2015). Importantly, neither Ddi1 nor Dsk2 have been associated with DNA repair or stabilization of proteins, as Rad23 has been implicated through its partnership with Rad4. This represents a novel and distinct function of Rad23, apart from other shuttles in its class.

To further distance Rad23 from its siblings, the UBA domains of Rad23 has functions apart from the transport of polyubiquitylated substrates. Studies suggest that Rad23 has the capacity to act as an inhibitory regulator of the UPS by preventing polyubiquitylation of proteins already bound to its multiple interaction domains (Verma *et al.* 2004; Elsasser *et al.* 2004; Chen *et al.* 2001; Ortolan *et al.* 2000; Raasi and Pickart 2003). This interaction actually stabilizes substrates, despite often utilizing identical interaction sites. Rad4, which is the binding partner of Rad23 within the nucleotide excision repair pathway, is one such substrate; when bound to Rad23, Rad4 is protected from polyubiquitylation and thus is shielded from degradation by the proteasome (Zientara-Rytter and Subramani 2019). This duality highlights the novelty and exceptionality of Rad23, and illustrates the reason its regulation is of such interest.

Rad23 and Transcription

In addition to its roles in DNA repair and protein degradation, Rad23 also plays a role in the regulation of gene transcription. This is especially apparent in the context of UV damage response genes. While this area of study is still in its infancy, recent research has tied Rad23 to a number of regulatory pathways that correspond to hundreds of discrete gene products.

In one such pathway, Rad4 and Rad23 cooperatively bind the promotors of DNA damage response genes in the absence of UV-induced damage (Zhou *et al.* 2015). The complex binds to the Dun1 kinase promoter, inhibiting the expression of a downstream ribonuclease reductase protein responsible for cellular deoxynucleotide triphosphate (dNTP) pool regulation. As a large store of nucleotides is unnecessary without the context of cell division or DNA repair, the complex remains associated until it detects DNA damage. When the associated stress signals are received, the heterodimer dissociates and transcription proceeds to accommodate the changing cellular requirements (Zhou *et al.* 2015).

In the presence of UV stress, Rad23 may cooperate with other factors to further the damage response cascade. Following irradiation, the Rad4/Rad23 complex releases from its perch on gene promoters, and instead spends time scanning the genome for DNA damage. When lesions are detected, Rad23 is displaced from its partner and can carry out additional functions within the nucleus (Bergink *et al.* 2012; Sugasawa *et al.* 2005). Recent data shows that Rad23 is at least partially responsible for the regulation of two-thirds of the genes altered by UV exposure, and that over one-third of all *Saccharomyces cerevisiae* genes are dysregulated when Rad23 is absent (Wade *et al.* 2009; Wade and Auble 2010).

These effects, governed in part by Rad23, are closely interwoven with Snf1 kinase. These two proteins jointly facilitate the displacement of a Mig3 transcriptional suppressor from the

HUG1 gene promoter, which controls the expression of numerous DNA damage response genes (Figure 8). With the promoter free, HUG1 is bound by RNA polymerase, and DNA damage response genes are expressed (Wade *et al.* 2009).

In a third pathway, Rad23 interacts with yet another protein, E2F1 transcription factor, in a manner similar to that of Rad4. Rad23 shields both unmodified and polyubiquitylated E2F1 from proteasomal degradation, most likely through its interaction with Rad23's Rad4 binding domain. Together as a complex, the duo enhances cellular response to UV damage repair through a symbiotic relationship: E2F1 is protected from degradation by Rad23, and Rad23 subcellular localization into the nucleus is modulated by E2F1 (Singh and Dagnino 2016).

While these roles of Rad23 are more poorly understood than that of the major pathways discussed previously, they serve to illustrate the diverse and intricate relationships in which Rad23 participates. This protein interacts with a myriad of cellular substrates, and each interaction must be uniquely governed to ensure that the cargo is properly delivered in the appropriate state. The role of Rad23, when each pathway is taken on its own, can be explained in a fairly straightforward manner. In the context of total cellular processing, however, the current understanding of Rad23 activity is underwhelming at best.

Rad23 at the Crossroad

From the topics discussed, it is clear that Rad23 is far more complex and imperative than a simple cellular taxi service. In some circumstances, Rad23 binds to and protects Rad4 from damage, stimulates its binding to damaged DNA, and intuitively dissociates once lesions are detected to aid in repair complex assembly. In others, it interacts with targeted substrates translocate them to the proteasome, where it is able to safely interact with the proteolytic

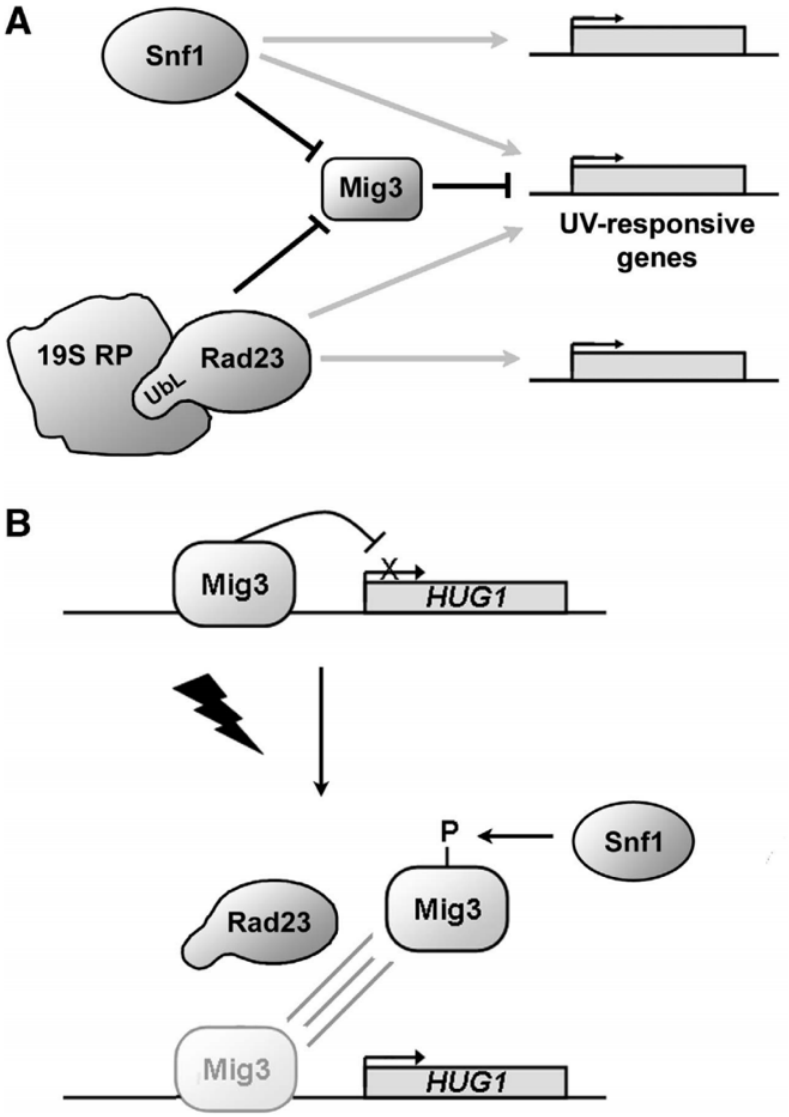


Figure 8: Model for transcriptional regulation by Snf1 and Rad23. A: Snf1 kinase and Rad23 cooperatively enhance UV-responsive gene expression through the inhibition of transcriptional repressor Mig3. This effect may rely on the physical and functional relationship between the UbL domain of Rad23 and the 19S regulatory subunit of the proteasome following UV damage, though this relationship is still poorly characterized. **B:** In the absence of damage, the transcription factor Mig3 represses the HUG1 promoter. Following irradiation, however, Mig3 is phosphorylated through a Snf1- and Rad23-dependent mechanism, which de-represses the HUG1 promoter and activates the transcription of UV-responsive under its control. Adapted from Wade *et al.* 2009.

structure to deposit its unlucky partners. In other conditions still, Rad23 interacts with a number of nuclear proteins, not to alter their stability, but to regulate the transcription of hundreds of gene products.

Despite the intricate activities of Rad23 *in vivo*, very little research has been afforded to the understanding of this protein. This is due, in part, because there are a number of safeguards in each pathway that prevent the loss of Rad23 from becoming severely debilitating. In NER, Rad4 is capable of acting in the absence of Rad23, albeit less efficaciously, and transcription-coupled NER is considered entirely Rad23 independent. In the UPS, other shuttles, such as Ddi1 and Dsk2, are capable of shuttling several overlapping substrates. In transcription regulation, there are other factors, such as Snf1, which can aid in the activation and suppression of Rad23-reliant genes. In short, loss of Rad23 does not pose a significant risk to human pathogenesis. That is not to say, however, that Rad23 is not closely intertwined with human molecular biology and overall well-being. For these reasons, the primary focus of this investigation is the regulation of Rad23.

Preliminary Data

The current project is derived from an ongoing investigation of Rad23 function in nucleotide excision repair. Prior analysis of *Saccharomyces cerevisiae* Rad23 following UV damage revealed an increase in ubiquitylation following exposure, which was abolished in strains lacking the Rad23 UbL domain (Figure 9, Smith Unpublished Data). These findings begged the question as to whether ubiquitylation of Rad23, in the presence or absence of DNA damage, could be responsible for the diversity of its cellular response.

To further investigate the potential relationship between post-translational modification and Rad23 activity, several Rad23 mutant cell lines were established in which various

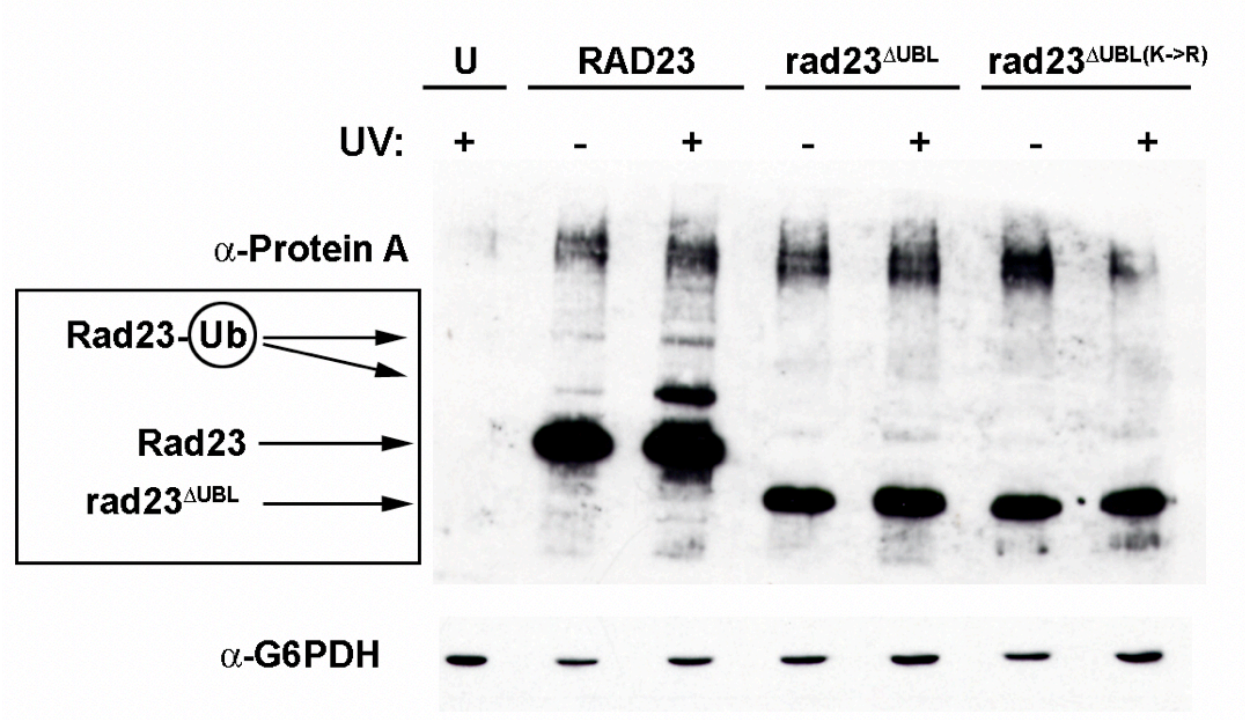


Figure 9: *Saccharomyces cerevisiae* Rad23 UbL is ubiquitylated following UV damage. Western blot analysis of three forms of TAP-tagged Rad23: Wild-type (Rad23), a UBL-less mutant (*rad23*$^{\Delta UBL}$), and a UbL- and "lysine-less" mutant in which the UbL domain has been deleted and the remaining lysine residue in UBA2 has been converted to arginine (*rad23*$^{\Delta UBL(K->R)}$). Yeast expressing these forms of Rad23 were treated with (+) or without (-) a dose of 100 J/m^2 of UV light. Samples were isolated 2 hours post-treatment. Note the ubiquitin-attributed shift in migration distance occurs only in WT Rad23 cells. Anti-Protein A was used to detect Rad23 forms and anti-G6PDH was used as a loading control (Adapted from Smith, Unpublished Data).

combinations of UbL lysine residues were substituted to the unmodifiable amino acid arginine. Survivability assays, utilizing UV irradiation as a source of cellular stress, compared wild-type *S. cerevisiae* cell lines to the strains lacking Rad23, its UbL domain, or modifiable UbL lysines (Figure 10; Figure 11). These studies illustrated not only that the UbL domain of Rad23 is required for efficient repair (Smith unpublished), but also that the absence of ubiquitylation targets (lysine residues) directly correlates to a decreased percent survival rate post-UV (Shrestha Thesis, 2011).

Despite uncovering an important relationship between UV damage repair and post-translational modification of Rad23, these data did not definitively establish which amino acid positions, if any, were critical for ubiquitylation to facilitate efficient NER. The mechanism by which these modifications may regulate Rad23 also remains unclear. As such, this project seeks to further the understanding of this phenomenon and make progress towards a comprehensive pathway for Rad23-mediated GG-NER.

Tetrahymena thermophila as a Model Organism

Most Rad23/HR23 studies to date have utilized *S. cerevisiae* or *M. musculus* as primary model organisms. This is generally due to similarities to the human model, and ease of manipulation and maintenance, respectively. Generally, mammalian models more closely resemble the systems and states of humans, and therefore have significant benefit over more distantly related animals, plants, or fungi in clinical and laboratory studies. However, mammalian models often require time and resources beyond that of simpler subjects, which is why *S. cerevisiae* became a popular choice, especially in basic research endeavors.

While *S. cerevisiae* does employ a GG-NER pathway similar to that of humans, and *S.*

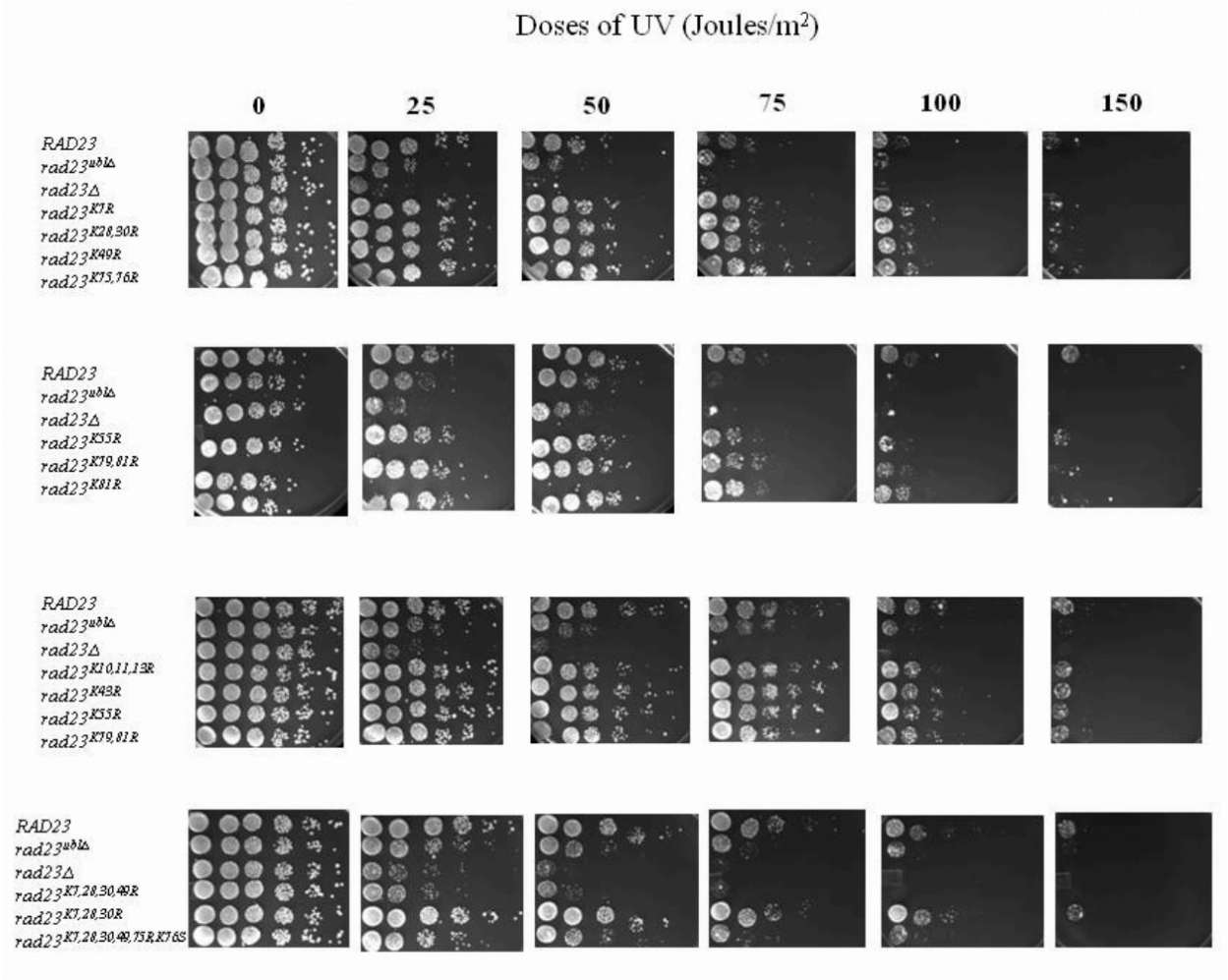


Figure 10: UV Survival Assay of Rad23 Mutants in *Saccharomyces cerevisiae*. Serially diluted yeast cells were spotted on YPD plates and treated with increasing doses of UV. Wild type *Saccharomyces cerevisiae*, used as a positive control, is indicated as RAD23. The mutant lacking the UbL domain is listed as $rad23^{UbL\Delta}$. Cells lacking Rad23 completely are shown as *rad23*Δ. Single point mutants made to lysine residues within the Rad23 UbL domain ($rad23^{K\#R}$) were comparable to wild type lines in response to UV treatment. Combination lysine-to-arginine mutants $rad23^{K7,28,30,49,75R,\ K76S}$ and $rad23^{K7,28,30,49R}$ were UV sensitive compared to wild type, and combination mutant $Rad23^{K28,30,49R}$ was slightly UV sensitive compared to wild type. Adapted from Shrestha Thesis, 2011.

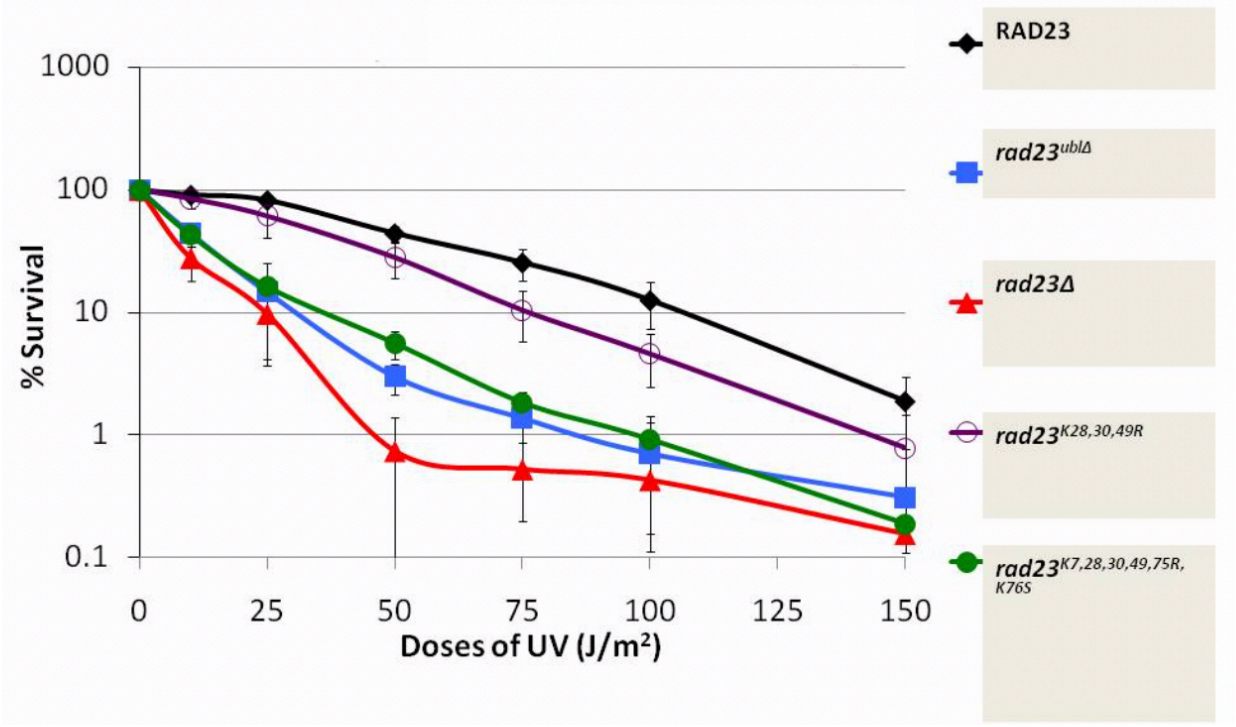


Figure 11: UV Survivability of Rad23 UbL Mutants. Percent survival of various Rad23 mutant *S. cerevisiae* cell lines were quantified relative to untreated controls following treatment with increasing doses of UV irradiation. Wild-type Rad23 is shown in black diamonds. Rad23 lacking the entire UbL domain is shown in blue squares. A total Rad23 deletion is shown in red triangles. Rad23 expressing lysine to arginine point mutations at positions 28, 30, and 49 is shown in purple circles; lysine to arginine point mutations at positions 7, 28, 30, 49, and 75, and a lysine to serine point mutation at position 76, is shown in green circles. The values plotted are the mean of three independent experiments, each performed in triplicate. Standard errors of the mean are indicated by error bars. Adapted from Shrestha Thesis, 2011.

cerevisiae Rad23 is largely comparable in structure and function to the human hHR23A/B orthologues, this model, too, has its limitations. In particular, the *S. cerevisiae* genome contains very few transcriptionally silent loci; the mating type loci (HML and HMR) and telomeric regions are among the only genomic locations that rely solely on the GG-NER (Gartenberg and Smith 2016). This means that any data obtained from yeast UV survivability assays is a summation of TC-NER and GG-NER, and therefore cannot be attributed solely to the manipulated state of Rad23. Another affordable but more complex model organism is the protist *Tetrahymena thermophila*.

Tetrahymena thermophila is a microscopic, unicellular, ciliated protist found primarily in freshwater environments. From a laboratory standpoint, these 30μm-50μm pond scum are quick growing, easy to maintain, and readily manipulatable. The cells double in culture every 2-5 hours, can be stored in stock tubes for several months without attention, or cryopreserved in liquid nitrogen indefinitely. Genetic mutations are readily induced through biolistic transformation of plasmids followed by drug selection, and phenotypes can be observed through a variety of assays.

The unique genome of *Tetrahymena thermophila* affords additional value to research projects such as this one. Each cell exhibits nuclear dimorphism, and splits its impressive genome between a transcriptionally silent germline micronucleus and a transcriptionally active somatic macronucleus (Collins and Gorovsky 2005). In contrast with the diploid human genome, *Tetrahymena* possess 45 copies each of ~200 chromosomes in its macronucleus, and pairs of 5 additional chromosomes (containing all macronuclear genetic information) within its micronucleus (Ruehle *et al.* 2016). *Tetrahymena thermophila* therefore possess an entire nuclear structure that relies exclusively on GG-NER, likely mediated by Rad23.

Hypothesis

Fifteen of the twenty-seven modifiable lysine residues in *T. thermophila* Rad23 are located within its N-terminal UbL domain (Figure 12). This domain, in turn, is at least partially responsible for the protein's role in NER. It is therefore possible that modification to one or more of these residues induces conformational changes to the protein structure, which then result in the distinct protein-protein interactions discussed. Ubiquitin modification may also allow for the formation and disruption of intramolecular interactions between the UbL and one or both of the UBA domains, which has been shown to occur *in vitro* (Walters *et al.* 2002). This restructuring would result in distinct open and closed protein conformations (Ramsey *et al.* 2004), which would greatly enhance substrate binding diversity (Figure 13). This hypothesis forms the basis for the current project.

The "Jekyll and Hyde Hypothesis", as it is currently named, describes the idea that the ubiquitylation patterns of the Rad23 UbL domain dictates the role of the protein under differing cellular conditions. In the absence of damage, Rad23 is likely unmodified, and therefore exists in a state of moderate affinity for Rad4 (Figure 13). Upon induction of UV irradiation, however, the Rad23 UbL domain is ubiquitylated on one or more conserved lysine residues, which facilitates the formation of intramolecular interactions between the UbL and one or both UBA domains (Figure 13). This conformational change may alter the Rad4 binding site to a degree sufficient to increase affinity between the two proteins, resulting in a dramatic increase in Rad4 stabilizing capacity and DNA binding stimulation. Once the damaged bases have been excised and repaired, and structural stability restored, Rad23 is deubiquitylated, adopts the original "open" conformation, and can resume its roles in proteolytic destruction within the UPS.

>TTHERM_00013290(protein)
MKINI**K**TLKGTDFFDVNLEETATVAEL**K**E**K**IATEKQKEKDTIKLVH**K**G**K**QLTEDSKTLGELGIKDN
DFVILMFFQ**KK**AEKEDAPQQAQSDTTSTTSAASTTATNPTTVPKPAVSQPATTQQTGSQGTGSDLL
QGPELEAKIKEIESMGFERPKVLQALKAAYYNPERAVDYLLSGNIPKEPSQQQSPLQGLQGPGVEQ
LAQLAQNPQFQHIAQAIRQNPALLQPVMQQLAQTNPDVARLLQQNPQAFLQLLLAASENEGGQTLP
PNAIQVTPEEKADIDDIISMGFDKNDALEAYITCDKNKELAINYLFEARESGTLLSEHIQKEELEA
AQQQSQNQNQQQQGNNPNQQQQQGGQGEGGDNDDEDDDDFMYQ

Figure 12: *Tetrahymena thermophila* Rad23 protein sequence. The protein sequence of *Tetrahymena thermophila* Rad23 (TTHERM_0013290) was obtained from the Tetrahymena Genome Database. Domain boundaries are highlighted as follows: Ubiquitin-Like domain (UBL) - pink; Ubiquitin Associated Domains (UBA) - yellow; XPC Binding Domain - blue. Ubiquitin-modifiable lysine residues are shown in red text. Lysine residues that have been identified as potential targets for mutagenesis are bolded and underlined within the UBA domain, including residues 6, 28, 30, 47, 49, 76, and 77.

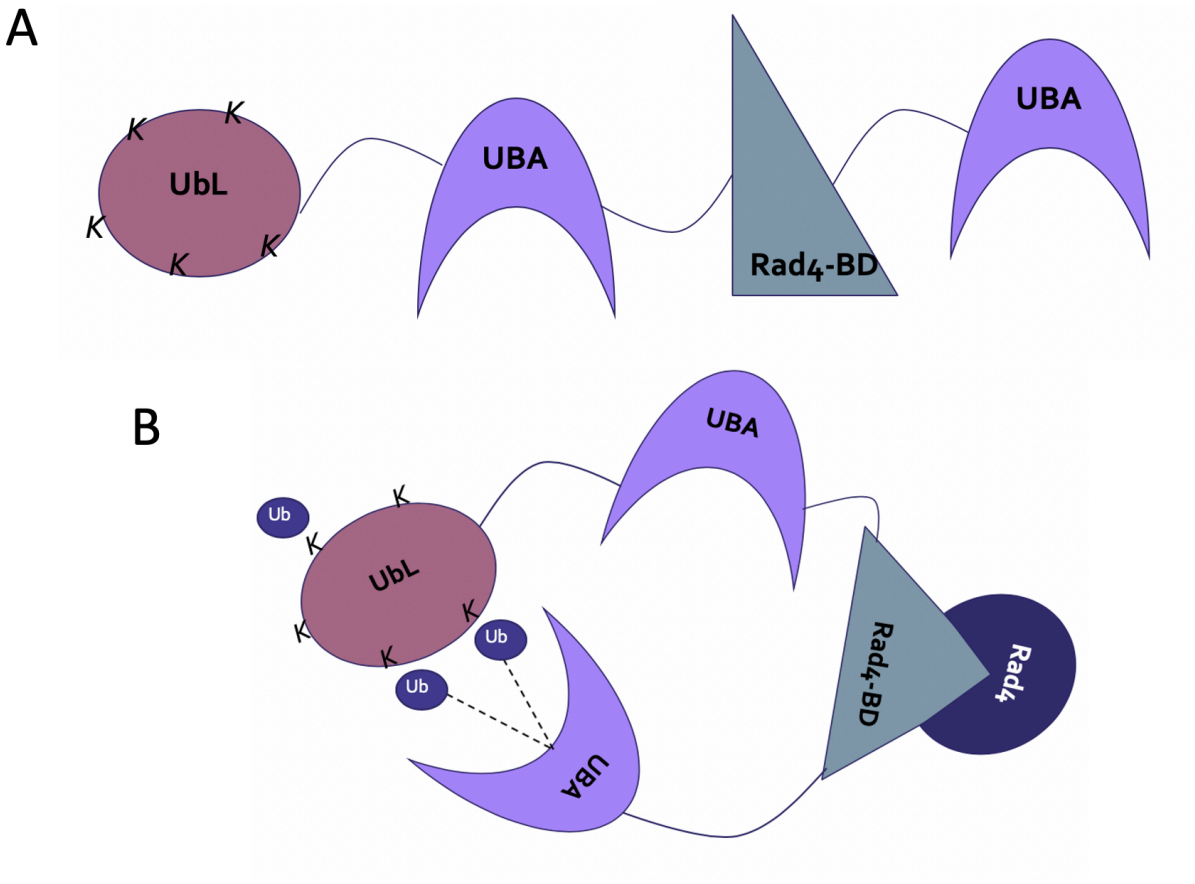


Figure 13: Schematic of the "Jekyll and Hyde" Hypothesis of Rad23. A: In the absence of damage, Rad23 exists in a conformation unconducive to high-affinity Rad4 interaction, and most cellular Rad23 does not participate in Rad4 binding. **B:** UV-induced stress results in ubiquitylation of one or more conserved lysine residues within the Rad23 UbL domain. Modifications allow intramolecular interaction between the UbL and one or both of the UBA domains. This circularized conformation increases affinity for Rad4 and increases Rad23 capacity for stabilization and stimulation of Rad4 DNA binding activity.
"open" conformation, and can resume its roles in other maintenance and regulatory pathways within the cell.

Experimental Rationale

While the role of Rad23 in NER has been illustrated across numerous model organisms, the same phenotypes of increased sensitivity to UV irradiation and deficiency of NER in the absence of Rad23 have not yet been observed in *Tetrahymena thermophila*. To first implicate *T. thermophila* Rad23 in efficient GG-NER, a somatic knockout of the protein was established; the micronuclear copy of Rad23 is transcriptionally silent, and therefore was not of interest when addressing expression. With Rad23 effectively depleted from the cells, the phenotype of Rad23 deficiency could be assessed.

To accomplish the genomic silencing, a previously designed Rad23 knockout plasmid (Appendix A1) was biolistically transformed into *Tetrahymena* cells. This plasmid utilized the 3' and 5' non-transcribed spacers (NTS) of the Rad23 gene to target an MTT1 promoter controlled Neo4 cassette to the Rad23 gene locus. Upon recombination, the Rad23 gene was permanently lost, and the Neo4 gene was incorporated in its place. When the MTT1 promoter is induced through the addition of $CdCl_2$ into the system, Neo4 confers resistance to the drug Paromomycin, which provides a mechanism for transformant selection.

Three *Tetrahymena* cell lines were utilized to examine the effects of Rad23 depletion, the first being a wild-type strain known as CU428. The two additional cell backgrounds, CU522 and CU725, were chosen due to a unique, advantageous mutation in one of the β-tubulin genes. These mutants, when untransformed, are susceptible to the tubulin stabilizing drug Paclitaxel, due to their overexpression of the cytoskeletal protein. However, when an exogenous sequence, such as a mutant or tagged Rad23 gene, is knocked into the mutant β-tubulin region, the cells become resistant to Paclitaxel and sensitive to Vinblastine, a tubulin destabilizing drug. This characteristic provides an extremely useful mechanism of transformant selection following re-

transformation. Therefore, while CU428 provides a look at how a typical, wild-type *Tetrahymena* cell responds to Rad23 loss, CU522 and CU725 provide supporting data, as well as a method for analyzing the effects of mutant Rad23.

Characterization of these mutants utilized both UV survivability assays and qRT-PCR gene expression analysis in order to assess the validity of the previously established phenotype. Due to the development of novel, largely unanticipated results, this became a major focus of the project, and a great deal of effort was put towards characterization of these *rad23*Δ models.

To establish the mutant Rad23 plasmids for future analysis, the Kunkle and Phusion methods of site-directed mutagenesis were attempted. Mutagenic primers were designed to support either method, by annealing to each strand of a 2HA-Rad23 containing plasmid and induce mutations to one or two of seven modifiable UbL lysine residues. Restriction enzyme sites were incorporated into each primer to facilitate screening following mutagenesis. Once confirmed, plasmids can later be transformed into the established CU522 and CU725 *rad23*Δ lines to further assess the effects of Rad23 post-translational modification on nucleotide excision repair, providing a long list of potential future directions and research projects.

Purpose Statement

This thesis aims to explore the role of *Tetrahymena thermophila* Rad23, and its modifications, in global genome nucleotide excision repair. Several cell lines will be established to examine the effects of Rad23 deficiency, as well as mutant Rad23 plasmids to later examine the effect of Rad23 ubiquitylation on cellular survivability following UV damage. This project requires a solid foundation in bioinformatic analysis, primer design, *in vitro* recombinant DNA techniques, gene expression analysis, and whole-cell and molecular phenotypic assays. When complete, this

project will provide greater insight into the inner workings of a highly conserved eukaryotic repair mechanism, as well as contribute understanding to the relationship between NER and the UPS pathways, especially in the context of *Tetrahymena thermophila*. The products that result from this research will provide a starting point for a broad array of spin-off projects for both undergraduate and graduate students in years to come.

MATERIALS AND METHODS

Bioinformatics

In order to confirm the validity of *Tetrahymena thermophila* as a model organism, the relationship between Rad23/HR23 orthologues was evaluated through the use of a variety of bioinformatic tools and databases. The Tetrahymena Genome Database (TGD) was used to gather preliminary data on the sequence, structure, and expression profile of *T. thermophila* Rad23. NCBI protein BLAST was then used to further identify related orthologues (Table 2). The seventeen orthologues, representing proteins from twelve different models, were analyzed against the *T. thermophila* protein sequence through TCOFFEE alignment and MEGA-X phylogenic tree construction.

Cell Culturing

Multiple strains both of *Escherichia coli* and *Tetrahymena thermophila* were utilized over the course of this project. Each cell line was cultured under conditions specifically tailored to support optimal growth while fulfilling experimental parameters. While general culturing protocol was kept consistent across strains, variables including type and concentration of drug selection, incubation protocols, and culturing timeline were adapted to individual experimental requirements. These parameters are described within these specific protocol entries.

E. coli. *E. coli* strains (Table 3) were cultured in LB media comprised of 10 g/L bacto-tryptone (Thermofischer #211705), 5 g/L yeast extract (Thermofischer #212750), 10 g/L NaCl (Thermofischer BP358-1), or on LB plates containing an added 15g/L bacto-agar (Amresco #J637-1KG). Selection was accomplished through the addition 100mg/mL Ampicillin, or

Table 2: Rad23 Orthologue Identifiers

Organism	Official Identifier	Common Name	Shorthand Identifier
†*Tetrahymena thermophila*	TTHERM_00013290	Rad23	TtRad23
†*Dictyostelium discoideum*	XP_637729.1	repC-binding protein A	DdRad23
Saccharomyces cerevisiae	AJU48317.1	Rad23p	ScRad23
Schizosaccharomyces pombe	NP_596231.1	UV excision repair protein rhp23	SpRph23
§*Arabidopsis thaliana*	OAO95837.1	Rad23D	AtRad23D
§*Zea mays*	NP_001345371.1	uncharacterized protein LOC100191313	ZmRad23Un
§*Zea mays*	NP_001142065.1	DNA repair protein RAD23	ZmRad23
‡*Homo sapiens*	AAH88364.1	RAD23 homolog A (S. cerevisiae)	HsRad23A
‡*Homo sapiens*	NP_002865.1	UV excision repair protein RAD23 homolog B isoform 1	HsRad23B
‡*Mus musculus*	NP_033037.2	UV excision repair protein RAD23 homolog B	MmRad23B
‡*Mus musculus*	NP_033036.2	UV excision repair protein RAD23 homolog A isoform 2	MmRad23A
‡*Danio rerio*	NP_001003739.1	UV excision repair protein RAD23 homolog A	DrRad23A
‡*Danio rerio*	NP_956858.2	UV excision repair protein RAD23 homolog B	DrRad23B
‡*Xenopus laevis*	NP_001082494.1	RAD23 homolog B, nucleotide excision repair protein S homeolog	XlRad23B
‡*Caenorhabditis elegans*	NP_496488.2	Uncharacterized protein CELE_ZK20.3	CeRad23
‡*Drosophila melanogaster*	NP_651212.1	uncharacterized protein Dmel_CG10694	DmRad23Un
‡*Drosophila melanogaster*	NP_001259052.1	Rad23, isoform C	DmRad23C

†Kingdom Protista *Kingdom Fungi §Kingdom Plantae ‡Kingdom Animalia

Table 3: Cell Lines

Organism	Cell Line	Genotype	Obtained From
E. coli	RZ1032	lacZbd-279::Tn10 Hfr dut-1 ung-1	D. Romero University of Minnesota
E. coli	DH10B	F⁻ *mcr*A Δ(*mrr-hsd*RMS-*mcr*BC) φ80*lac*ZΔM15 Δ*lac*X74 *rec*A1 *end*A1 *ara*D139 Δ(*ara-leu*)7697 *gal*U *gal*K λ⁻ *rps*L(Str^R) *nup*G	D. Romero University of Minnesota
T. thermophila	CU428.2 (SD00178)	*mpr1-1/mpr1-1* (*MPR1*; mp-s, VII)	Tetrahymena Stock Center
T. thermophila	CU522 (SD00271)	*mpr1-1/mpr1-1* (*mpr1-1*; mp-r, ory-r, tax-s, IV)	Tetrahymena Stock Center
T. thermophila	CU725 (SD00481)	*chx1-1/chx1-1* (*chx1-1*; cy-r, ory-r, tax-s, VII)	Tetrahymena Stock Center

50mg/mL Kanamycin (Thermofischer #BP906-5). Cells were generally allowed to grow at 37°C for 12-16 hours prior to experimental analysis. Liquid cultures were placed in shaking incubators at 200-220 RPM for this purpose.

T. thermophila. *T. thermophila* strains (Table 3) were cultured in 2% PPYS with 1X PSF (HyClone #SV30079.01) to prevent contamination. PPYS 20% stock was prepared using 200g Proteose peptone #3 (Thermofischer #211693), 20g yeast extract (Thermofischer #212750), and 0.3g Sequestrine (BioWORLD #30630115-3) per liter. Liquid cultures were grown overnight at 30°C and 100 RPM; plates were kept at room temperature or grown in humidity chambers at 30°C. Drug selection using 0.1 mg/mL -1 mg/mL Paromomycin (Sigma #P5057-5G) or 20μM-40μM Paclitaxel (LKT Laboratories #P0092) was used to select for transformant clonal colonies when indicated. Selection of cells involved regular passaging through 24-well, 48-well, and 96 well plates as needed. Confirmed transformants were moved into stock tubes containing 10mL 1% PPYS composed of equal parts 2%PPYS and 10mM Tris (Thermofischer #BP152-1) with 1X PSF. These stocks were remade approximately every 6 months to ensure survival and preservation of the strains.

Plasmid DNA Isolation: Midiprep

Plasmids were designed and synthesized as components of earlier research projects (Appendix A). In order to isolate plasmid DNA from storage vectors, a lysosome-boil midiprep method was employed.

Prior to isolation, 25mL liquid cultures of Ampicillin-selective LB media were inoculated with single colonies of plasmid-containing *E. coli.* Following incubation, cells and media were transferred to sterile conical vials and centrifuged for 10 minutes at 6,000 RPM to pellet the cells. With the supernatant removed, cells were resuspended in 3.5mL sucrose lysis buffer

composed of 8% sucrose, 0.5% Triton X-100 (Sigma #X11-500mL), 50 mM EDTA (Thermofischer #BP118-500), 10 mM Tris pH 8.0 (Thermofischer #BP152-1), followed by 250µL of 10mg/mL lysozyme (Thermofischer BP535-1). This volume was divided amongst four lock-cap 1.5-mL Eppendorf tubes per sample, and allowed to incubate at room temperature for 5 minutes to facilitate cell lysis.

Following incubation, samples were heated in a boiling water bath for 60 seconds to inactivate the enzyme, then centrifuged at maximum speed for 15 minutes to precipitate proteins, genomic DNA, and other unwanted cellular components. Pellets were removed and discarded. The plasmid DNA for each sample, which remained in the supernatant of the four tubes, was combined into a single 15-mL conical vial per strain and treated with 400µL 3M sodium acetate (Thermofischer #S209-500) and 2.2mL isopropyl alcohol to precipitate the plasmid DNA. Samples were again divided into 2-3 1.5-mL Eppendorf tubes per strain and incubated at room temperature for 5 minutes to facilitate precipitation. Following incubation, cells were centrifuged at maximum speed for 5 minutes and the supernatant discarded. The resulting pellet was washed with 1mL of 70% ethanol, centrifuged for 5 minutes at max, and supernatant discarded. After a series of quick spins and pipette removal of residual ethanol, followed by complete air-drying of the pellet, the DNA was submerged in 50µL of 1X TE pH 8, allowed to incubate for 5 minutes at room temperature, then completely resuspended by pipetting. Each strain was re-combined into a single 1.5-mL Eppendorf tube, and brought to a final volume of 300µL with 1X TE (10mM Tris pH 8.0 and 1mM EDTA pH 8.0). Samples were frozen and stored at -20°C until purified (see DNA Purification) and quantified (see Nucleic Acid Quantification).

Genomic DNA Isolation

Genomic DNA (gDNA) was isolated from *T. thermophila* from 3mL or 4mL liquid cultures. Cells were then transferred to a 1.5-mL Eppendorf tube and spun at maximum speed for 60 seconds to pellet the cells. Supernatant was discarded and any remaining culture was added, spun, and decanted once more. The tube was flicked gently to loosen the cells in the residual media, then cells resuspended in 700µL of urea lysis buffer. This mixture was inverted until viscous, then RNase treated and phenol:chloroform extracted as described in the next section.

DNA Purification: RNase and Phenol:Chloroform

To remove residual RNA contamination, DNA was treated with 10mg/mL RNase A (Thermofischer #50824384). One microliter RNase A was added per 100µL of sample, and allowed to incubate at 37°C for 15 minutes (gDNA) or room temperature for 3 hours (plasmid DNA). Following incubation, samples were stored at -20°C or used immediately for subsequent procedures.

Phenol:Chloroform extraction was performed through the addition of 25:24:1 mixture of Phenol:Chloroform:Isoamyl Alcohol in a volume equal to that of the sample being extracted. These samples were inverted or vortexed to mix thoroughly, then centrifuged for 5 minutes at maximum speed to produce a distinct organic and aqueous layer. The top, aqueous layer, containing the DNA, was transferred into a fresh 1.5-mL Eppendorf tube, and treated with 1/10 volume of 3M NaOAc and 2.5 volumes 100% ethanol or isopropyl alcohol (gDNA and plasmid DNA respectively). The samples were allowed to precipitate; gDNA was precipitated for at least 10 minutes at room temperature, while plasmid DNA was allowed to precipitate overnight at -20°C. Following precipitation, samples were centrifuged for max speed for 10 minutes, and

supernatant removed. Pellets were washed in 0.5mL-1mL of 70% ethanol, respun, and dried completely. Resuspension of gDNA was accomplished in 50µL 1X TE (pH 8), plasmid DNA in 150µL of sterile water.

Nucleic Acid Quantification

A Nanodrop Spectrophotometer was utilized to quantify nucleic acid concentration and purity. Prior to quantification, 1µL-2µL of 1X TE (pH 8) or sterile water, whichever was used to resuspend samples, was used to blank the equipment, then each sample was loaded in equal quantity, read, and removed using a Kim wipe. Values recorded included concentration (µg/µL) and purity ($A_{260}/A_{280)}$ values.

Sequencing

DNA sequencing was performed through the MU DNA Core Facility, according to published guidelines. Samples to be sequenced, in the form of plasmid DNA, were isolated and purified according to previously described procedures. Reactions were prepared using 1µL of primer, 1µg/µL plasmid DNA, and sterile water to a final volume of 16µL. Primers used to sequence 2HA-Rad23 GTW plasmids included MTT1 -151F and BTU2 154R (Table 4). Primers used to sequence Rad23 KO pUC118 plasmids included M13F and M13R (Table 4). Data files were analyzed using Codon Code Aligner software.

Restriction Enzyme Digestion

All restriction enzymes and buffers were obtained from New England Biolabs (NEB), and prepared as 20µL reactions. For each digest, a master mix was prepared for n+1 reactions,

Table 4: Plasmid Confirmation and Transformant Screening Primer List

Primer Name	Primer Sequence (5'-3')
Rad23 -132F	CTG TCT ACA TAC ATC AAC AGA AAA
MTT1 -815R	TAA AAG GAG TTA TTC AAA ACC CTT
BTU2 154R	GTT AAG TTA TAG GGA AGA ATG AAT AAG
MTT1 -151F	CGT GAT TCA CGA TTT ATG CAA TGA TC
M13 -20F	GTA AAA CGA CGG CCA GT
M13 -27R	CAG GAA ACA GCT ATG AC
Rad23 E3 R	GCT CTT TCA GGA TTG TAA TAA
Rad23 204R	TAG CAA TCT TCT CCT TCA ATT CAG
DDI1 383 R	AGATGATGGGATTTAACAACAATT

each consisting of 0.5µL high-fidelity restriction enzyme(s), a final concentration of 1X Cutsmart Buffer (NEB #B70204S), and sterile water to volume. Master mix was aliquoted 19µL per tube, and 1µL of appropriate DNA was added. Samples were incubated at 37°C for 3.5 hours to 16 hours. Samples were visualized via agarose gel electrophoresis and ethidium bromide staining.

Agarose Gel Electrophoresis

Agarose gel electrophoresis was prepared using ISC BioExpress Agarose (E-3120-500) and 1µL of 10 mg/mL Ethidium Bromide (Sigma #E7637-1G) per 1mL of TAE. Gel concentrations ranged from 1.0% to 1.5% agarose in 1X TAE (1L: 242.0g Tris Base, 57.1mL Glacial Acetic Acid, 100mL 0.5M EDTA pH 8.0). For all procedures excluding single-stranded DNA isolation visualization, gels were prepared through pre-staining of the gel: EtBr was added in volume (µL) equal to the mass (g) of agarose used. For single-stranded DNA isolation visualization, the gel was stained post-electrophoresis, through soaking of the gel in 10µL EtBr in 100mL water. All gel images were obtained using the Kodak Gel Logic 200 imaging system.

Electroporation of E. coli

Electroporation was conducted using electro-competent DH10B *E. coli*. Cells were removed from storage at -80°C and thawed slowly on ice. Plasmid was diluted to ~10ng/µL in sterile water, and 10ng DNA was gently mixed into 50µL of competent cells. The cells were transferred into 2mm cuvettes, and the Bio-Rad Gene Pulser II set to 25 capacitance and 2.5 kV. Voltage was initiated, and time constant was recorded.

Following electroporation, a pasture pipette was used to resuspend and transfer the cells into 1mL LB media. Cells were allowed to recover for 1 hour at 37°C in the shaking incubator, then plated on LB+Amp plates for overnight incubation and colony counting.

Chemical transformation of *E. coli*

Chemically competent *E. coli* cells were thawed on ice then distributed into 100µL aliquots into Falcon 14-mL polypropylene round-bottom "snap-cap" tubes (352059). One to two microliters of plasmid DNA or ddH_2O (negative control) was gently mixed into the cells. Cells were incubated for 20 minutes on ice, heat shocked for 2 minutes at 37°C, then chilled on ice for 2 additional minutes. Cells were mixed with 900µL of LB media, then incubated for 1 hour at 37°C, 200 RPM. Following recovery, cells were plated on LB+Amp and/or LB+Kan plates and incubated overnight at 37°C for colony counting and isolation.

Plasmid Preparation for Biolistic Transformation

Biolistic transformation of *Tetrahymena thermophila* relies on homologous recombination of exogenous DNA into the host genome. As a consequence, plasmid DNA must be linearized, without compromising the desired sequences and plasmid elements. Prior to biolistic transformation, the Rad23 knockout plasmid was digested with SacI-HF (NEB #R3156S) and SphI-HF (NEB #R3182S), and the 2HA-Rad23 plasmid was digested with SpeI-HF (NEB R3142S).

Restriction enzyme digestion was set up as described in the corresponding section above. Samples were allowed to incubate for 16 hours at 37°C, then phenol:chloroform extracted according to the protocol described in the DNA Purification section. Following purification,

samples were quantified and concentrations were adjusted to 2µg/µL. A 1.0% or 1.5% agarose gel was used to visualize results of the digestion and assess suitability for transformation.

Biolistic Transformation of *Tetrahymena thermophila*

Preparation of *Tetrahymena.* *Tetrahymena* cell lines CU428, CU522, and CU725 were transferred from stock to 24 well plates in triplicate and treated with 20µM-40µM Vinblastine and 20µM-40µM Paclitaxel (LKT Laboratories #P0092) to ensure proper starting phenotype. Cells were incubated in a humidity chamber for three days, then used to start a 25-mL starter culture in 2% PPYS+1X PSF. CU522 and CU725 were taken from Vinblastine drugged wells, CU428 from un-drugged wells.

Following overnight incubation, cells were counted using a hemocytometer, and each cell line was transferred to a 150-mL culture to achieve a confluency of $3x10^5$ cells/mL following 24 hour incubation (CU428 doubles approximately every 3 hours, CU522 and CU728 double approximately every 4 hours). After this time period, cells were again counted using a hemocytometer, transferred into sterile oil flasks, and spun at 3,000 RPM for 3 minutes to concentrate the cells. Supernatant was discarded, and cells were resuspended in 100mL of 10mM Tris, pH 7.5. This step was repeated once to remove any residual PPYS media. The cells were transferred to clean 500-mL flasks. PSF was added to each strain to a final concentration of 1X, and flasks were allowed to incubate overnight.

On the day of transformation, cells were quantified using the hemocytometer to ensure a concentration of at least $1x10^7$ cells/mL (1mL cells per shot). Cells were again transferred into sterile oil flasks and spun as described. Supernatant was removed, and cells were resuspended in 10mM HEPES (Thermofischer #BP310-500) to a final concentration of $1x10^7$ cells/mL.

Preparation of DNA. For each desired shot, one tube of gold beads (25µL) was prepared at 4°C. To each tube, 2.5µL of prepared DNA (2µg/µL), 25µL 2.5M $CaCl_2$, and 10µL 100mM spermidine (Thermofischer #AC13274-0010) was added. Tubes were then vortexed continuously for 30 minutes to facilitate DNA binding to gold beads (BioRad #1652263). Following vortexing, samples were centrifuged for 2-3 seconds, supernatant removed, and 100µL of 70% ethanol was added to wash cells. This step was repeated with 100µL of 100% ethanol, then 20µL of 100% ethanol. Samples were kept on ice until transformation.

Preparation of Equipment. The hood was turned on and allowed to run at least 15 minutes prior to equipment set-up. All surfaces were sterilized using 70% ethanol. Steel stopping screens (BioRad #1652336) were autoclaved prior to use. Steel macrocarrier holders (BioRad #1652322) were soaked in 70% ethanol and thoroughly dried prior to assembly.

Macrocarriers (BioRad #1652335) were assembled into their holders and gold-bound DNA was applied directly to the center of the disk. These samples were allowed to dry completely prior to transformation. Flasks of pre-warmed 2% PPYS+1XPSF were prepared and labeled for cell recovery following transformation.

In the lid of a sterile 10mm petri dish (Thermofischer #FB0875713)) , 3mL 10mM HEPES (Thermofischer #BP310-500) was allowed to thoroughly soak a 9mm filter paper (Whatman #1003-090). One milliliter of appropriate cell culture was added to each filter. The gas tank was turned on and the valve tightened the to achieve a pressure of 1100 PSI. The vacuum pump was turned on and the rupture disk tightened into its holder within the gene gun apparatus. A stopping screen was placed into the appropriate groove in the platform, and the DNA-coated microcarrier was inverted and positioned just above the stopping screen. The cell-

containing petri dish was placed in the farthest position from the stopping screen to maximize dispersion of the DNA.

With the apparatus assembled, the door was closed and sealed securely. Pressure was increased to and held at 25mm Hg. The gun was fired until the rupture disk burst at approximately 900 PSI. Pressure was released and the cells were removed. For negative control lines, the same procedure was performed without the inclusion of macrocarriers or rupture disks.

Cell lines were resuspended in a few milliliters of liquid media from pre-warmed flasks, which were swirled across the filter paper to lift the cells. Liquid was returned to the corresponding flask, and sterilized tweezers were used to carefully fold the filter paper and transfer it into the flask to be submerged.

Cell Recovery and Plating. Flasks were transferred to a shaking incubator for 6 hours. Following incubation, cells were transferred to 50-mL conical vials and drugged with 20μM Paclitaxel (2HA-Rad23) OR 100 μg/mL Paromomycin and 1 μg/mL cadmium chloride (Rad23 KO). The volume was then distributed to three 96-well plates and one 48-well plate per strain (two and one respectively for negative controls) and incubated in humidity chambers for 4 days. After this time, cells were passaged with increasing dosages of drug to encourage reassortment of recombinant DNA. Transformant cells were screened using GoTaq PCR.

GoTaq PCR Confirmation of Transformants

GoTaq PCR screenings were conducted on biolistically transformed *Tetrahymena thermophila* to confirm presence of recombinant DNA. Confirmation of *rad23*Δ cell lines was accomplished using Rad23 -132F and MTT -815R primers. Confirmation of 2HA-Rad23 cell

lines was accomplished using MTT -151F and either Rad23 E3R or Rad23 204R primers (Table 4).

A GoTaq master mix was created for each screening and primer optimization reaction, consisting of a final concentration of 1X GoTaq polymerase (Promega #M712B), 0.4 pmol/µL forward and reverse primer, template DNA, and sterile water to a total volume of 25µL/reaction. The thermocycler was programmed as follows: 1) 95°C for 2 minutes. 2) 95°C for 45 seconds. 3) Calculated annealing temperature for 45 seconds. 4) 72°C for 1 minute, 45 seconds. 5) repeat steps 2-4 for 32 additional cycles. 6) 72°C for 5 minutes. 7) Hold at 4°C forever.

Primer optimization was accomplished using 1µL of a 1:100 dilution of pure plasmid corresponding to the amplification target. A 6X reaction was prepared, and a temperature gradient protocol was used to amplify the DNA using annealing temperatures from 52°C to 62°C. Products were run on a 1.0% agarose gel and temperatures were evaluated for band cleanliness and intensity. The optimal temperature was used for subsequent sample screenings.

Colony PCR was accomplished through the screening of 2µL of transformant cells. Additionally, 2µL-3µL of untransformed wild-type cells were screened as a negative control, and 1µL of the dilute plasmid DNA was screened as a positive control. Following amplification, these samples were run on a 1.0% agarose gel and assessed for expected banding patterns.

Genomic DNA screening was also performed. Template DNA was isolated and purified as previously described, and two microliters was used for screening reactions. Again, 1µL of dilute plasmid DNA was used as a positive control.

Single Stranded DNA Isolation

The Kunkel method of site-directed mutagenesis requires the synthesis of single-stranded plasmid DNA. The 2HA-Rad23 plasmid isolated from the RZ1032 *E. coli* cell line of was therefore prepared using an M13K07 helper phage system.

Prior to isolation, plasmid-containing cells were streaked on LB+Amp plates and grown overnight into single colonies. Four colonies were used to inoculate glass tubes containing 2mL 2X YT media (16.0g Bacto Tryptone, 10.0g Bacto Yeast Extract, 5.0g NaCl in 1 L H_2O), 20µL of 0.1% Vitamin B1, 0.5µL M13K07 helper phage (NEB #N0315S), 4µL 50mg/mL Ampicillin. Tubes were incubated for 2-4 hours at 37°C and 200 RPM, until a "silkiness" was observed in the media, representing early stages of confluency. Two microliters of 25mg/mL Kanamycin were added to select for infected cells, and allowed to continue incubating, shaking, overnight.

Following incubation, cells were transferred to microcentrifuge tubes and spun at max speed for 10 minutes. Supernatant was collected and respun, and the top 1mL of supernatant was placed into fresh Eppendorf tubes. To each sample, 300µL of 20% PEG (PEG 8000, Sigma #202452-250G) in 2.5M NaCl (Thermofischer #BP358-1) was added and inverted to mix. This solution was incubated for one hour, at room temperature or on ice, then spun at max speed for 10 minutes to pellet the viral particles. Following complete removal of supernatant, pellets were resuspended in 200µL of TE.

An equal volume of phenol was added to the sample (200µL), thoroughly mixed, then centrifuged. The aqueous layer was transferred into a fresh Eppendorf tube and phenol:chloroform extracted in the same manner previously described. To this aqueous layer, a volume of chloroform was added and the sample extracted. From this sample, up to 162µL of aqueous layer was removed and transferred to a fresh tube for ethanol precipitation. Samples

were allowed to precipitate at room temperature or on ice for 30 minutes, then ethanol was removed and DNA resuspended in 15µL of TE. A 1.0% agarose gel test of isolates was used to visualize the presence of ssDNA in the sample.

Phusion Mutagenesis

Single-stranded DNA isolation of the 2HA-Rad23 plasmid proved inefficient; Phusion mutagenesis was therefore selected as the method of lysine to arginine mutagenesis. Unlike the Kunkel method, Phusion mutagenesis requires mutagenic primer pairs to be added individually, rather than in a large batch reaction. While this procedure is more time-consuming and effort-intensive, it does not require isolation of a single-stranded DNA template, and can use RZ1032 or DH10B isolated plasmid in its native form. A ThermoScientific Phusion Site-Directed Mutagenesis kit with DH10B cells (#F542) was used for mutagenesis.

Phosphorylation of Mutagenic Primers. Mutagenic primers (Table 5) were phosphorylated at the 5' end to facilitate ligation into a circularized plasmid following the mutagenesis reaction. A 50µL reaction mixture was prepared for each oligonucleotide, consisting of 250pmol of the respective primer, 5µL of 10X reaction buffer A for T4 polynucleotide kinase (NEB #B0201S), 5µL of 10 mM ATP, and 2µL of T4 polynucleotide kinase 10 U/µL (NEB # M0201S). The reaction was incubated at 37°C for 30 minutes, then heat inactivated at 75°C for 10 minutes. Products were stored at -20°C while awaiting mutagenesis.

Site-Directed Mutagenesis (PCR). The mutagenic reaction mix was assembled as follows: 10µL 5X Phusion HF Buffer (Thermofischer #F-518); 1µL 10mM dNTPs (Thermofischer #F-560KS); 5µL of each phosphorylated mutagenic primer (one forward and one reverse); 1µL 2HA-Rad23 template plasmid DNA; 0.5µL of 2U/µL Phusion or Phusion Hot

Table 5: Mutagenic Primer List

Rad23 Mutation Position	RE Screen	Primer Sequence (5'-3')
K6R+	BsiWI	CGAGAAGATCAACATTCGTACGTTAAAGGGCACTGAT
K6R+	BsiWI	ATCAGTGCCCTTTAACGTACGAATGTTGATCTTCTCG
K28R+	SacI	CTTGTTAGGTTGCTGAGCTCAGGGAGAAGATTGCTACTG
K28R-	SacI	CAGTAGCAATCTTCTCCCTGAGCTCAGCAACCTAACAAG
K28,30R+	SacI	CTTGTTAGGTTGCTGAGCTCAGGGAGAGGATTGCTACTGAAAAG
K28,30R-	SacI	CCTTTCAGTAGCAATCCTCTCCCTGAGCTCAGCAACCTAACAAG
K47R+	SacII	CTATTAAGTTAGTTCACCGCGGAAAATAATTGACCGAAG
K47R-	SacII	CTTCGGTCAATTATTTTCCGCGGTGAACTAACTTAATAG
K49R+	ZraI	GTTAGTTCATAAAGGACGTCAATTGACCGAAGACT
K49R-	ZraI	AGTCTTCGGTCAATTGACGTCCTTTATGAACTAAC
K47,49R+	SacII	CTATTAAGTTAGTTCACCGCGGAAGATAATTGACCGAAGAC
K47,49R-	SacII	GTCTTCGGTCAATTATCTTCCGCGGTGAACTAACTTAATAG
K76R+	XbaI	GTTATTCTCATGTTCTTCTAGAGGGTAATATTATCTTAGT
K76R-	XbaI	ACTAAGATAATATTACCCTCTAGAAGAACATGAGAATAAC
K77R+	XbaI	CTAAATATAACATTATTCTAGAGAGCAGAAAAAGAAGATG
K77R-	XbaI	CATCTTCTTTTTCTGCTCTCTAGAATAATGTTATATTTAG

Start II DNA Polymerase (Thermofischer #F-549 or #F-530S respectively); ddH_2O to 50µL. The thermocycler was programmed as follows: 1) 98°C for 30 seconds. 2) 98°C for 10 seconds. 3) 54°C for 30 seconds. 4) 72°C for 2.5 minutes. 5) Repeat steps 2-4 for 24 additional cycles. 6) 72°C for 10 minutes. 7) Hold at 4°C forever. To verify proper amplification of the desired product, 5µL was used to run a gel test. The remaining product was stored at -20°C until ligation.

Degradation of Template and Ligation of Products. Following amplification, the unmutagenized parent template was removed prior to transformation. Dpn1 restriction enzyme nicks DNA at Dam methylated sequences, and therefore serves to selectively degrade DNA synthesized by and isolated from *E. coli* hosts. Samples were therefore incubated with 1µL of FastDigest DpnI (ThermoScientific, #F-541) for 15 minutes at 37°C prior to ligation.

A ligation mix was prepared consisting of 10-20ng of DnpI digested PCR product, 2µL 5X Rapid Ligation buffer (Thermoscientific #F-545), and sterile water to a volume of 9.5µL. Half a microliter of T4 DNA ligase (Thermoscientific #F-544) was added to bring the final volume to 10µL. Samples were mixed thoroughly, then centrifuged (quick spin) and incubated for 5 minutes at room temperature. Tubes were chilled on ice, then transformed or stored at -20°C.

Transformation into *E. coli*. Mutagenized plasmids were transformed into chemically competent DH10B *E. coli* according to the protocol outlined in the section Chemical Transformation of *E. coli*. Template volumes used ranged from 1µL-5µL, and 50µL-100µL chemically competent cells were used per reaction. Cells were plated on LB-AMP plates and incubated overnight at 37°C.

Isolation and Screening. Plasmid DNA was isolated from single transformant colonies according to the procedure described in the section Plasmid DNA Isolation. Each mutagenic

primer set was designed to incorporate a specific restriction enzyme cut site to facilitate screening. Following isolation, plasmid DNA was digested with SacI and the restriction enzyme dictated in Table 5 (SacI only for designated primers), according to the protocol described in the section Restriction Enzyme Digestion. The digested products were run on an agarose gel for visualization, and samples producing expected banding patterns were used as templates in subsequent mutagenesis reactions.

UV Survivability Assay

UV survivability assays were performed on drug selected, PCR-confirmed cells, as well as wild-type controls. All three CU backgrounds (Table 3) were represented. To prepare cells for the assay, 25-mL cultures were allowed to incubate at overnight, to a confluency of at least $1x10^5$ cells/mL. If confluency exceeded $2.5x10^5$ cells/mL, the cultures were diluted to reach the desired range, then returned to the incubator for one hour before proceeding. Once at the desired concentration, cells were transferred to 50-mL conical tubes and centrifuged for 3 minutes at 3,000 RPM to pellet the cells and remove media. Pelleted cells were resuspended in 20mL of 10mM TRIS, counted again, and diluted to a final concentration of $1x10^5$ cells/mL.

Each cell line was divided into two identical petri dishes, 10mL each, and any extra volume discarded. One of these two dishes was treated, via the CL-1000 Ultraviolet Crosslinker, with 75 J/m^2 UV unless otherwise indicated. The other dish was left untreated as a control. From both sample populations, 1mL was taken and used for two sequential 1:10 dilutions in 10mM Tris. From the most dilute tube, 600µL was transferred into 29.4mL of 2% PPYS+PSF. Each cell line was distributed into three 96-well plates each and allowed to incubate in a darkened humidity chamber for seven days. Following this recovery period, live wells were counted and survivability was determined relative to the untreated WT control average for each cell line.

RNA Isolation (RNeasy)

An RNeasy kit from Qiagen (#74104) was used to isolate total RNA from wild-type and *rad23*Δ cell lines, both before and after UV treatment. To prepare the cells for isolation, a 25-mL starter culture of each cell line was established and allowed to grow overnight (~16 hours) to a concentration between $1x10^5$ to $3x10^5$ cells/mL. A hemocytometer was used to quantify cell density.

Cells were moved to sterile 50-mL conical vials and centrifuged for 3 minutes at 3,000 RPM to pellet the cells and remove supernatant. Pellets were resuspended to 20mL in 10mM TRIS, and all cell lines were quantified and diluted with TRIS to a final concentration of $1x10^5$ cells/mL. Cell lines were divided equally (10mL each) into two petri dishes, one subjected to 100 J/m^2 UV and the other left untreated. Samples were then transferred into separate 15-mL conical tubes, spun to remove supernatant, and transferred to flasks containing 10mL of 2% PPYS. Cells were allowed to recover for 1 hour in the shaking incubator, then transferred back to 15-mL conical tubes, spun, and resuspended in 10mL of 10mM TRIS to remove residual media. Cells were spun again, media was removed, and tubes were flicked to resuspend cells in the residual TRIS.

To each tube of cells, 600µL RLT buffer with 10µl BME (Thermofischer #VP176-100) was added and vortexed for 1 minute. One volume of 70% ethanol was added to lysate and mixed by pipetting. The sample, including any precipitate, was transferred to the spin column in 700µL fractions. Samples were centrifuged at max speed for 15 seconds and flow-through discarded. This step was repeated until all sample had been transferred, at which point 350µL

RW1 Buffer was added to the column and spun for 15 seconds and max speed. Again, flow-through was discarded.

To eliminate any DNA contamination, all samples were treated with 800µL DNase (Qiagen #174104) in RDD buffer, applied directly to the column membrane. This was incubated at room-temperature for 15 minutes, then washed with another 350µL RW1 Buffer. Two washes with 500µL RPE Buffer were then conducted, the first spun for 15 seconds and flow-through discarded, the second for 2 minutes and flow-through discarded. The spin column was then transferred to a fresh collection tube and spun again at max speed for 1 minute to completely remove any potential reagent contaminants. The column was transferred to a fresh 1.5-mL Eppendorf tube for final collection, and 50µL RNase-free water was added directly to the column membrane and spun for 1 minute and max speed. Two microliters were used for nucleic acid quantification using the nanodrop spectrophotometer, and the remaining volume was diluted to a concentration of 1µg/µL and stored at -80°C.

Reverse Transcriptase PCR (RT-PCR)

To accommodate the large number of RNA samples, a 30X Reverse Transcriptase cocktail was prepared as follows: 120µL 5X AMV Buffer (Thermofischer #M515A); 120µL 25mM $MgCl_2$ (Thermofischer #M87-100); 60µL 10mM dNTPs (Thermofischer #F-560KS); 30µL RNasin RNase inhibitor (NEB #M0307S); 30µL 7.5 U/µL AMV RT (Thermofischer #M510A); 60µL 50µM oligo dTVN; 120µL RNase-free water. This cocktail was distributed into aliquots of 18µL/tube, and 2µL (2µg) of the appropriate total RNA template was added. Samples were amplified via reverse transcriptase PCR using the following thermocycler program: 1) 42°C for 25 minutes. 2) 99°C for 5 minutes. 3) 4°C for 5 minutes. Samples were immediately removed

from the thermocycler, and equal volumes of RNase-free water was added to each tube to make concentrated cDNA. Samples were then stored at -20°C for use in qPCR -80°C for long-term storage.

Quantitative PCR (qPCR)

The qPCR cocktail was prepared on ice and consisted of a final concentration of 1X SsoFast Evagreen supermix (BioRad #1725200), 500nM forward and reverse primers (Table 6), and nuclease free water to a total volume of 19µL. One microliter of the reverse transcribed cDNA was used as a template. The cycling conditions were set as follows: 1) 98°C for 2 minutes. 2) 95°C for 5 seconds. 3) 54°C for 30 seconds. 4) Plate Read. 5) Repeat steps 2-4 for 39 additional cycles. 6) Melt curve from 54°C to 95 °C with increments of 0.5°C for 10 seconds each, followed by plate reads. Following the first amplification, samples were used for a gel test to ensure the absence of additional products. Subsequent experiments using identical primer pairs were not subjected to further testing.

Tetrahymena Protein Isolation

Evaluation of the expression of 2HA-tagged Rad23 was achieved through the isolation and analysis of protein from transformant *T. thermophila*. For this assay, the cell lines CU725 and CU725 *rad23*Δ (#13) were grown in 10mL cultures to a density of $1x10^5$ to $3x10^5$ cells per milliliter. Half of each culture was induced with 1µg/µL (5µL per 5mL fraction) $CdCl_2$ (Thermofischer #AC29633-0050) for 3 hours. Each sample was spun at 3000 RPM for 3 minutes to pellet cells and remove supernatant. Pellets were resuspended in 5mL of 10mM Tris (pH 7.4-7.5) and spun again. Supernatant was discarded, and cells were resuspended in 1mL 10mM Tris

Table 6: qPCR Primer List

Primer Name	Target	Primer Sequence (5'-3')
TtRad23qRTPCR F	Rad23	CAA GGC TTA CAA GGT CCT GGT GTT G
TtRad23qRTPCR R	Rad23	GAC TTG GAT GGC ATT TGG AGG TAA G
ACT1 677F	Actin	TGA ATT AAA GGC TTA CAA GGA ATC
ACT1 859R	Actin	CAC ACT TCA TGA TAG AGT TGA AGG
HHP1 76F	HHP1	GGA AAG GCT ACT CTC TCA AAT AAG
HHP1 420R	HHP1	TTT AGG AGT AGA TTT AGG ATT AGA TG
APN2 427F	APN2	GCA TGT TGG AAT GTG AAT GG
APN2 927R	APN2	CTT CGT TCT CGT GCT CTG C
Rad6 144F	Rad6	GAA GGT GCA ACA TTC AGA TTA AC
Rad6 337R	Rad6	GGA GTG AAC GTA TAG ATG TGA GG
Rad5E1F	Rad51	AAA CAG GCT CTC TCA CTG
Rad51E2R	Rad51	GAC TTA GGA CAT AAC TTG GTT G
TKU80 403F	TKU80	GAA AAA CCA GCT ACA AAG AAA AAT
TKU80 965R	TKU80	TAT CTG GTA TTT TAT GCA GCA TCG

(pH 7.4-7.5). The volume was transferred to 1.5-mL Eppendorf tubes and spun at 5000 RPM for 2 minutes. The supernatant was removed, save ~100µL, and the pellet resuspended.

To extract the protein from the cell, 350µL of Breaking Buffer with protease inhibitor was added to each sample; tubes were kept on ice or in the cold room from this point on. Cells were vortexed for one minute until the suspension cleared and spun again at max speed for 5 minutes at 4°C. Supernatant was transferred to a fresh tube, spun for 15 minutes, and supernatant removed again. This fraction was stored at -20°C awaiting quantification.

Protein Quantification

A BSA standard curve was constructed to quantify individual protein samples. Each standard was mixed with 200µL Bradford protein assay Reagent (Bio-RAD #500-0006) and quantified in a cuvette at 595nm using a nanodrop spectrophotometer. These five standards were used to construct a standard curve.

To quantify the samples, 5µL of each sample was diluted in 20µL of breaking buffer with protease inhibitor. Five microliters of this solution were mixed with 795µL of sterile water, then 200µL of the Bradford reagent. Samples were read at 595nm and the protein concentration calculated according to the standard curve.

Western Blot

Gel and Sample Preparation, and Electrophoresis. The running and stacking polyacrylamide gels were prepared according to Table 7. Two liters of running buffer was prepared using 28.067g glycine (Thermoscientific #BP381-1), 6g Tris Base (Thermoscientific #BP152-1), 6mL 10% SDS (Thermoscientific #BP166-500). Protein samples were prepared by

combining 25µL of a 4X SDS Loading Dye, 200µg isolated protein, and breaking buffer with protease inhibitor to a final volume of 100µL. Eighty micrograms of protein (40µl) were loaded onto the gel, alongside 5µL of PageRuler Plus Prestained protein ladder (ThermoScientific, #26619) or Full Range Rainbow Recombinant Protein Molecular Weight Marker (GE Healthcare Life Sciences #RPN800E). The gel was run at 220V until samples entered the running gel, at which point voltage was increased to 250V for 1.5 hours.

Sample Transfer, Ponceau Staining, and Detection. The polyacrylamide gel was removed from the plates and trimmed. The gel, a nitrocellulose membrane (GE Water and Process Technologies), and two blotting papers were soaked in transfer buffer (48 mM Tris, 39 mM glycine, 20% methanol, 1.3 mM SDS, pH 9.2) for 5 minutes. The BIO-RAD Trans-Blot SD Semi-Dry Transfer Cell apparatus was assembled and allowed to run at 25V for 50 minutes.

In order to confirm protein transfer to the membrane, Ponceaus stain was applied to membranes for 1 minute. Excess stain was removed, and the membrane was rinsed and imaged. A razor blade was used to trim the membrane, which was blocked in 5% milk in 1X TBS (50 mM Tris, 150 mM NaCl) for 48 hours. Following incubation, milk was removed and replaced with fresh 5% milk in TBS containing 1:2000 monoclonal anti-HA mouse antibody (Sigma-Aldrich H9658) overnight. The next morning, milk was removed and the membrane was washed four times with 1X TBS. The membrane was exposed to 5% milk in TBS containing 1:5000 Anti-Mouse antibody for 1 hour at room temperature. Following incubation, the membrane was washed four times with 1X TBS and coated with SuperSignal West Dura working solution and 400µL of Luminol/Enhancer solution with 400µL stabile Peroxide solution. Membranes were carefully covered in plastic wrap for imaging and storage. Visualization was performed using light sensitive Blue Ultra Autorad film (ISC BioExpress #F-9029-8X10). The AGFA

Table 7: PAGE Recipes

Reagent	12% Running Gel	Stacking Gel (mL)
Bis Acrylamide	12 mL	1.7 mL
1 M Tris HCL	11.4 mL (pH 8.8)	1.25 mL (pH 6.8)
H_2O	6.2 mL	6.8 mL
10% SDS	0.3 mL	100 µL
10% APS	0.3 mL	100 µL
1% BPB	---	80 µL
TEMED	30 µL	10 µL

CP1000 was used to process the films following various durations of exposure (1 second to 15 minutes).

Statistical Analysis

Statistical analysis was performed with the support of the Missouri State University *R*Stats office utilizing JASP. UV survivability assays were analyzed using non-parametric Mann-Whitney T-tests. P-values and rank-biserial correlation was obtained and reported for differences between wild-type and mutant survivability within each cell line. Gene expression data was analyzed via two-way ANOVA, in order to analyze the effects of both Rad23 loss and UV treatment on gene expression within each cell line. When necessary assumptions were fulfilled, a parametric ANOVA was used to report F-values, p-values, and η_p^2. Post-hoc analysis utilized independent T-tests with a Bonferroni correction. When assumptions were not met, a nonparametric ANOVA was used to report p-values and x^2 values. Post-hoc analysis provided means and standard errors for mutant vs. wild-type and UV vs. no UV groups. For all reported statistics, p-values above the threshold of 0.05 were considered to be non-significant (ns). Significant p-values are classified and denoted as follows: * $p \leq 0.05$; ** $p \leq 0.01$; *** $p \leq 0.001$.

RESULTS

Bioinformatics

The selection of *Tetrahymena thermophila* as a model organism required extensive bioinformatic analysis to ensure that the data obtained from *T. thermophila* Rad23 (TTHERM_00013290) could be compared and applied to that of other models. Preliminary information regarding *T. thermophila* Rad23 was obtained from the Tetrahymena Genome Database (TGD), including protein sequence (Figure 12), expression analysis (Figure 14), and gene structure (Figure 15).

Gene expression analysis revealed that the Rad23 gene is expressed most significantly during early stages of starvation, with a sharp decrease in expression at the start of conjugation (Figure 14). RNA Sequencing data confirmed the gene is composed of five exons and four introns, and provided a decorated FASTA gene sequence that was used for plasmid mapping, primer design, and general genetic analysis (Figure 15, data not shown).

The *Tetrahymena thermophila* Rad23 protein sequence, at 373 amino acids in length, was used to identify 17 additional orthologues for comparative bioinformatic analysis using the NCBI protein Basic Local Alignment Search Tool (BLAST) (Table 2). For each organism searched, the top orthologue with a published reference was selected, and the associated amino acid sequence collected. For organisms that had more than one isoform with comparable homology, both protein sequences were collected and differentiated using the isoform title given. Each of these sequences was given a working, shorthand identifier to be used for subsequent data collection (Table 2).

The first step in the assessment of *T. thermophila* as a model required the total protein

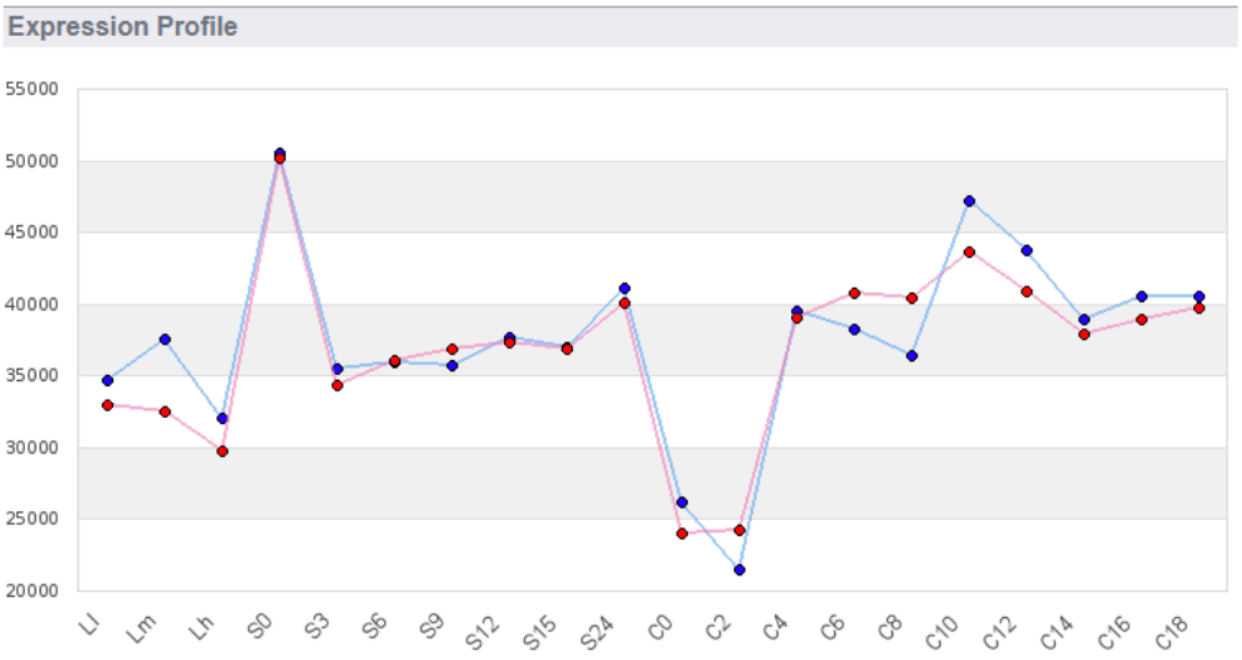


Figure 14: Gene Expression Profile of *T. thermophila* Rad23. Blue Line, data from 50 samples completed by Gorovsky and Miao labs; Growth (three time points), triplicates; Starvation (seven time points), triplicates; Conjugation (ten time points), duplicates. Red Line, data from 50 samples completed by Gorovsky and Miao lab, and 10 conjugation samples completed by Pearlman lab; all time points. For growing cells, L-l, L-m and L-h correspond respectively to ~1x10^5 cells/mL, ~3.5x10^{5}cells/mL and ~1x10^6 cells/mL. For starvation, ~2x10^{5}cells/mL were collected at 0, 3, 6, 9, 12, 15 and 24 hours, referred to as S-0, S-3, S-6, S-9, S-12, S-15 and S-24. For conjugation, equal volumes of B2086 and CU428 cells were mixed, and samples were collected at 0, 2, 4, 6, 8, 10, 12, 14, 16 and 18 hours after mixing, referred to as C-0, C-2, C-4, C-6, C-8, C-10, C-12, C-14, C-16 and C-18). Blue and red lines represent the expression values normalized by two different methods. Data obtained from the Tetrahymena Genome Database.

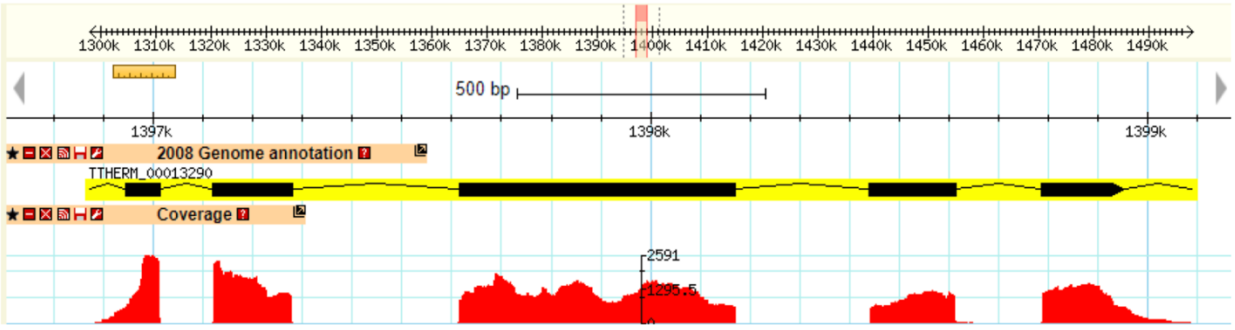


Figure 15: *T. thermophila* Rad23 Gene Structure and RNA Seq Data. This figure is a graphical representation of the structure of the 5'-3' *Rad23* gene in *Tetrahymena thermophila*. The upper black and yellow schematic represents the predicted structure of the Rad23 gene, obtained from 2008 genomic annotation data. The thick black lines represent the predicted exons, connected by the introns in thin black lines. Below, in red, is the coverage data obtained from RNA sequencing. Increased expression corresponds to larger peaks, and drop-offs in expression indicate intronic regions of the gene, removed during mRNA splicing events. Data obtained from the Tetrahymena Genome Database.

sequences to be broken down into individual functional domains. Literature dictates that the Rad23 protein orthologues contain four major functional domains: an N-terminal ubiquitin-like domain (UbL), two ubiquitin-associated domains (UBA), and an XPC-binding domain. To verify that the orthologues identified through NCBI BLAST shared the same composition, ExPasy Prosite and Interpro online databases were used to identify relevant functional structures in each protein. For organization, these domains are shown, with associated amino acid boundaries, in two categories: Kingdom Animalia orthologues (Table 8), and Kingdom Protista, Fungi, and Plantae orthologues (Table 9). The domain analysis tools collectively confirmed that each of the eighteen proteins (including *T. thermophila* Rad23) contain all four characteristic functional domains.

Armed with the total protein sequences and exact residue boundaries for each domain, comparative analysis could be attempted both between whole proteins, and the UbL domains exclusively. This distinction was important, as it provided an opportunity to focus in on the conservation of the potential ubiquitylation targets (lysine residues) within the ubiquitin-like domain, without potential skew of data from lesser conserved domains or linking regions. To ensure that *Tetrahymena thermophila* Rad23 was a suitable model from both perspectives, a neighbor-joining phylogenetic tree was generated for both the complete and selected sequences (Figure 16).

Both the whole-protein and UbL domain divergence patterns indicates that *T. thermophila* is middle-of-the-pack in regard to homology. In the whole-protein tree, *Tetrahymena* branches off with *Dictyostelium discoideum*, not far from the divergence of the plantae kingdom organisms (Figure 16A). When only the UbL domain was compared, *Tetrahymena* actually branched with *Saccharomyces cerevisiae*, though its divergence

Table 8: Kingdom Animalia Rad23 Orthologues.

Identifier	Amino Acids	ExPasy Prosite Domains	ExPasy Boundaries	Interpro Boundaries
HsRad23 A	362		3-81: UbL 161-201: UBA 317-357: UBA	3-81: UbL 161-201, 313-357: UBA 231-286: XPC
HsRad23 B	409		1-79: UbL 188-228: UBA 364-404: UBA	1-79: UbL 188-228, 364-404: UBA 274-317: XPC
MmRad23 B	416		1-79: UbL 188-228: UBA 371-411: UBA	1-79: UbL 188-288, 371-411: UBA 274-317: XPC
MmRad23 A	362		3-81: UbL 161-201: UBA 317-357: UBA	3-81: UbL 161-201, 317-357: UBA 231-286: XPC
DrRad23 A	362		1-79: UbL 159-199: UBA 317-357: UBA	1-79: UbL 159-199, 317-357: UBA 228-283: XPC
DrRad23 B	381		1-79: UbL 158-198: UBA 336-376: UBA	1-79: UbL 158-198, 336-376: UBA 251-294: XPC
XlRad23 B	412		1-79: UbL 191-231: UBA 367-407: UBA	1-79: UbL 191-231, 367-407: UBA 279-334: XPC
CeRad23	323		3-80: UbL 127-167: UBA	3-80: UbL 127-167, 278-318: UBA 196-251: XPC
DmRad23 Un	290		1-76: UbL 109-149: UBA 243-284: UBA	1-76: UbL 109-149, 243-284: UBA 172-227: XPC
DmRad23 C	414		1-78: UbL 155-197: UBA 370-410: UBA	1-78: UbL 155-197, 370-410: UBA 242-296: XPC

Table 9: Kingdom Protista, Fungi, and Plantae Rad23 Orthologues

Identifier	Amino Acids	ExPasy Prosite Domains	ExPasy Domain Boundaries	Interpro Domain Boundaries
†TtRad23	373		1-78: UbL 135-175: UBA 271-312: UBA	1-78: UbL 135-175: UBA 199-252: XPC 271-312: UBA
†DdRad23	342		1-76: UbL 161-201: UBA	1-76: UbL 161-201: UBA 227-281: XPC 298-337: UBA
*ScRad23	398		2-77: UbL 146-186: UBA 355-395: UBA	2-77: UbL 146-186: UBA 264-308: XPC 355-395: UBA
*SpRad23	369		1-77: UbL 135-185: UBA 320-360: UBA	1-77: UbL 135-185: UBA 247-302: XPC 320-360: UBA
§ZmRad23 Un	386		1-79: UbL 154-197: UBA 331-371: UBA	1-79: UbL 154-197: UBA 251-306: XPC 331-371: UBA
§ZmRad23	405		1-79: UbL 158-201: UBA 355-396: UBA	1-79: UbL 158-201: UBA 276-331: XPC 355-396: UBA
§AtRad23 D	378		1-78: UbL 146-189: UBA 330-371: UBA	1-78: UbL 146-189: UBA 253-308: XPC 330-371: UBA
§VvRad23 Un	399		1-79: UbL 174-217: UBA 353-393: UBA	1-79: UbL 174-217: UBA 274-329: XPC 353-393: UBA

†Kingdom Protista *Kingdom Fungi §Kingdom Plantae

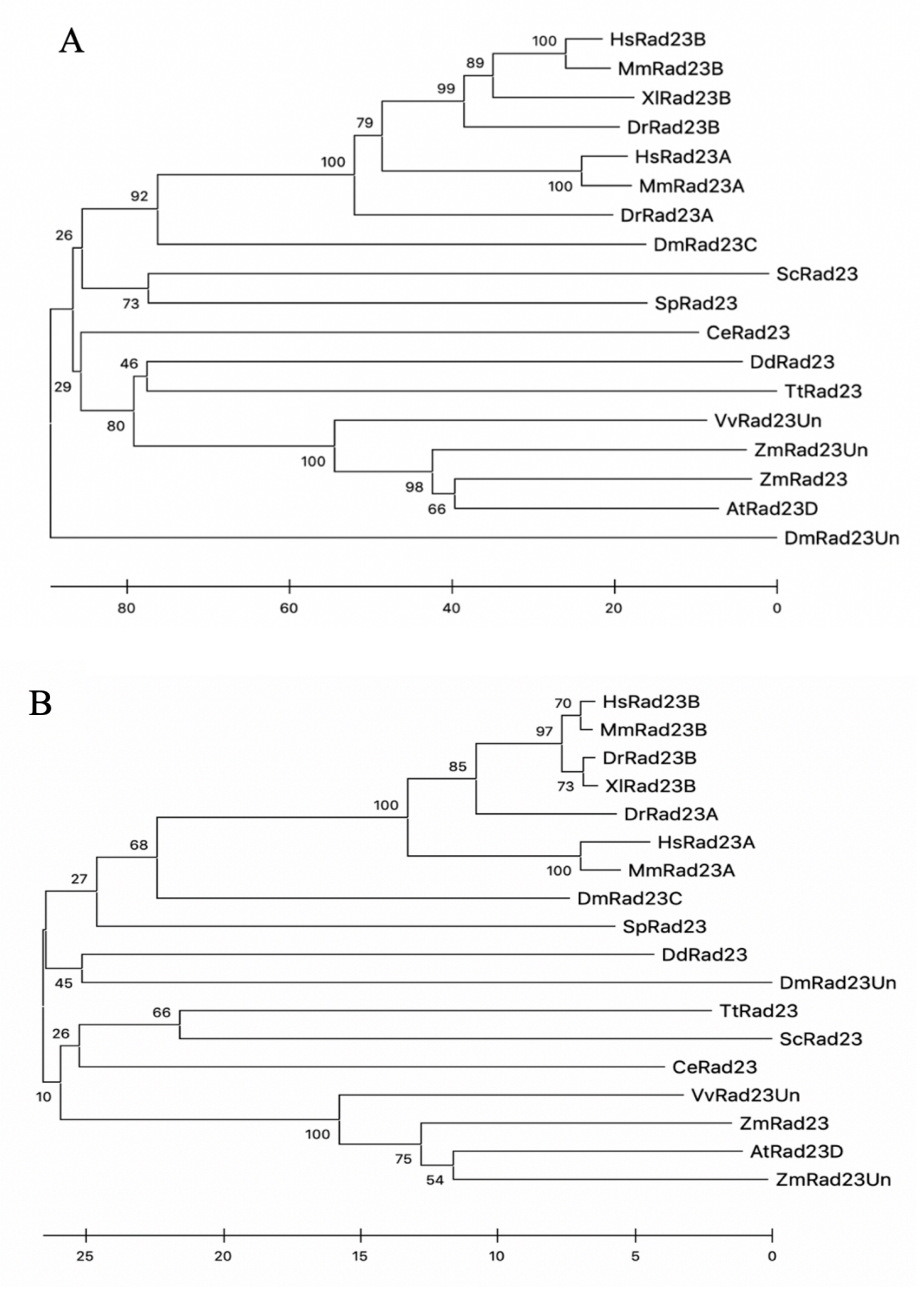


Figure 16: Neighbor-Joining Phylogeny of the Rad23 Protein. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) are shown next to the branches. Branch lengths are drawn to scale to correspond with the evolutionary distances between orthologues. Evolutionary analyses were conducted using MEGA X software. **A:** Divergence of the Rad23 protein in its entirety across models. **B:** Divergence of the ubiquitin like (UbL) domain of the Rad23 protein across models.

from *Homo sapiens* hHR23 did not shift (Figure 16B). TCOFFEE alignment of the orthologous UbL domains was used to evaluate the overall level of conservation, as well as identify potential targets for lysine to arginine mutagenesis (Figure 17). Typically, areas of higher conservation correspond to greater functional importance, as the residues involved have been maintained over the course of evolutionary divergence and selection.

In a line-by-line comparison of the eighteen sequences, only one lysine residue was registered as completely conserved (*), though two more were marked as highly conserved (:) across organisms (Figure 17A). Gaps and minor differences in amino acid positions can serve to skew the alignment, however, and may not indicate actual divergence. Six lysine residues in total were identified and manually marked as conserved, both using the software recommendations and human discretion.

In order to directly compare the *T. thermophila* Rad23 UbL with that of the two *Homo sapiens* isoforms, these three lines were isolated from the total alignment and compared side-by-side (Figure 17B). All six lysine residues previously identified were present in the three samples, which could further implicate them in protein function. In *Tetrahymena*, these residues correspond to positions 6, 28, 30, 49, 76, and 77. A seventh residue, at *T. thermophila* position 47, was also marked as a target despite its lack of conservation. Though it is less likely to be integral for protein function, mutagenesis can easily be designed to facilitate modification of this additional site in conjunction with K49, so its location has been marked as well (Figure 17B).

Generating a *rad23*Δ Cell Line

To test the effects of Rad23 loss in *Tetrahymena thermophila,* a *rad23*Δ cell line was established for phenotypic characterization. A recombinant plasmid had already been designed

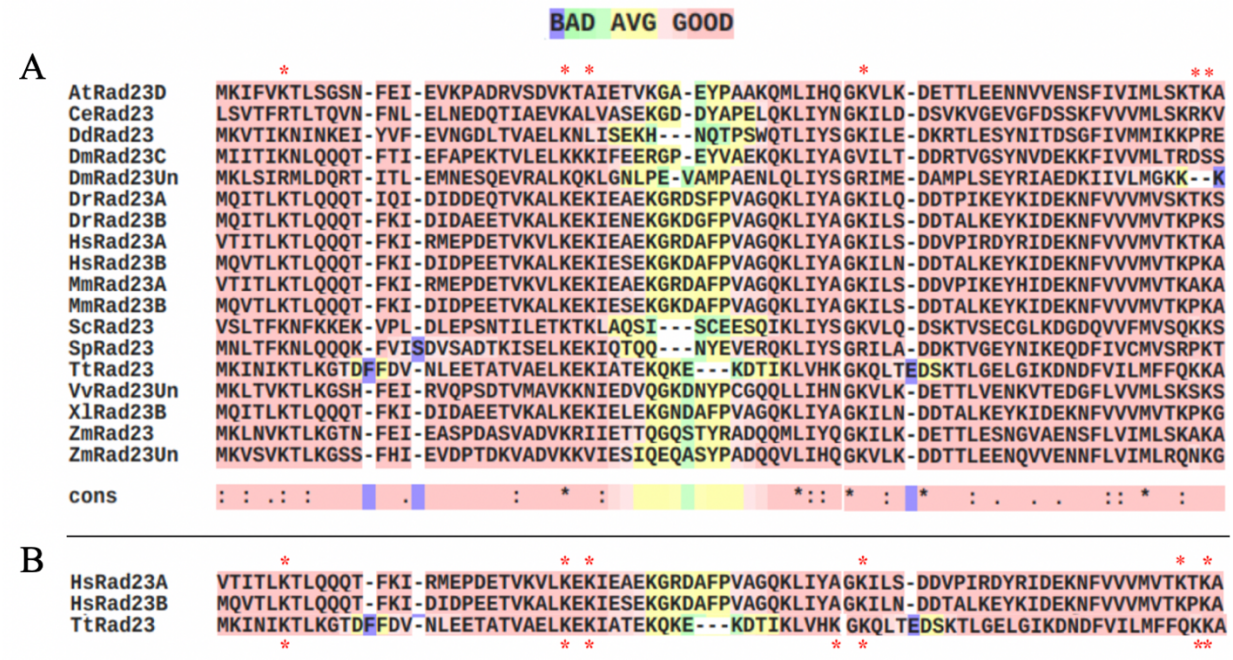


Figure 17: TCOFFEE Alignment of Rad23 Orthologue UbL Domains. The UBL domains of Rad23 orthologues were compared via TCOFFEE alignment. Homology strength is indicated by the color code shown at the top of the figure. Within the "cons" section, asterisks (*) indicate complete conservation at the given residue. Colons (:) and periods (.) indicate very strong and strong homology, respectively. **A:** Total alignment of UBL domains among Rad23 orthologues. Red asterisks along the top of the figure indicate the relative positions of the most highly conserved lysine residues relative to the *T. thermophila* sequence. These residues appear at positions 6, 28, 30, 49, 76, and 77. **B:** The alignments corresponding to the two *H. sapiens* orthologue isoforms and the original *T. thermophila* sequence were formatted for direct comparison. Red asterisks along the top indicate the conserved lysine residues identified in part A, as they appear in the two *H. sapiens* isoforms. Red asterisks along the bottom indicate the conserved residues as they appear in *T. thermophila*, with the addition of the lysine residue at position 47. These seven residues were selected as targets for mutagenesis.

and synthesized for this purpose (Appendix A1), but required confirmation and preparation prior to transformation into live cells. Based on sequencing data, the "Rad23KO 1.5Kb" plasmid was chosen for biolistic transformation into *Tetrahymena thermophila*.

Tetrahymena utilize homologous recombination for the incorporation of exogenous DNA sequences. DNA that is to be transformed must therefore be linearized via restriction enzyme digestion, to separate the desired gene from the bacterial plasmid backbone. For the Rad23 knockout plasmid, SacI and SphI were selected for this purpose. Though not all plasmid was successfully digested, resulting in a band at 7,863 base pairs, the digested fragments expected at 5,048 base pairs and 3,123 base pairs were also observed (Figure 18). The digested sample was therefore used in the transformation of CU428, CU522, and CU725 cell lines (Table 3), with the goal of producing transformants expressing reduced or abolished levels of Rad23 protein. Cells that survived the biolistic transformation were selected in the microtubule stabilizing drug Paclitaxel.

Cells were passaged in increasing concentrations of drug and decreasing concentrations of cadmium chloride ($CdCl_2$: used to induce expression of the Neo4 resistance marker under the inducible MTT promoter) for several weeks to encourage reassortment of the knockout construct and to eliminate false positive colonies. The remaining colonies were arbitrarily numbered and screened via PCR with MTT1 and Rad23 specific primers, found to optimally anneal at 58°C (Appendix B1). Presence of a ~650 base pair band was considered confirmation of the exogenous sequence.

Two different methods of PCR screening were conducted. The early screens were done using colony PCR, in which raw cells were used as a template for amplification. This method confirmed three positive colonies in the CU428 cell line (#1-3), one in the CU522 cell line (#19),

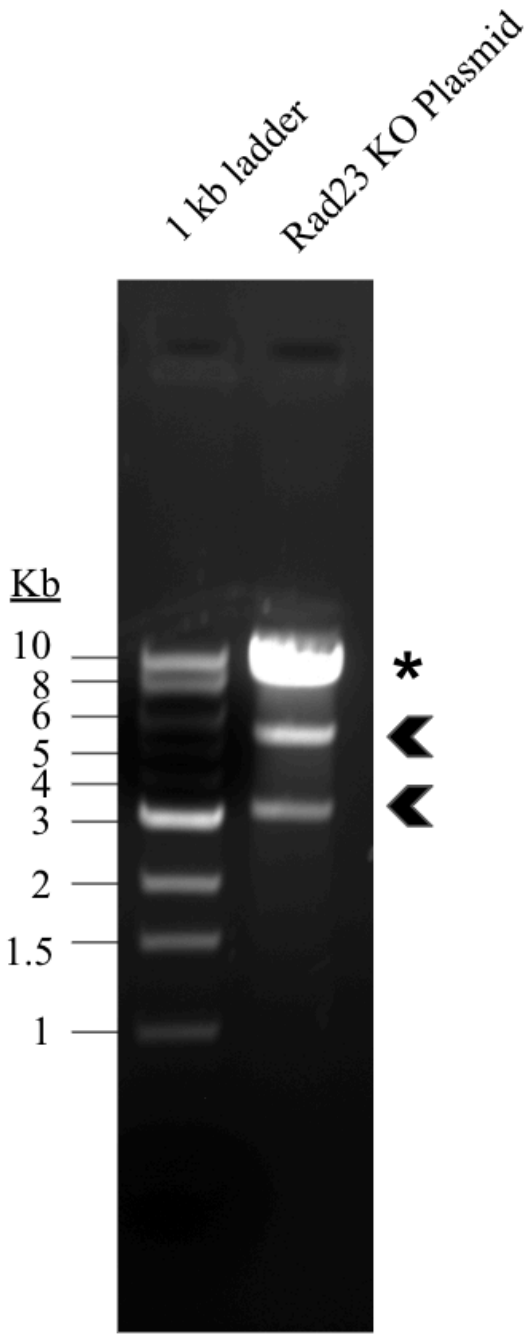


Figure 18: Rad23 KO Plasmid Preparation for Biolistic Transformation. Prior to transformation into *T. thermophila*, plasmids were restriction enzyme digested to linearize genes and remove the bacterial sequences of the plasmid backbone. The Rad23 knockout plasmid was digested with high fidelity SacI and SphI. Predicted band sizes for this digest were 5048 and 3123 base pairs, as indicated by arrowheads. The asterisk represents undigested plasmid.

and eighteen in the CU725 cell line (#2-3, #5-13, #15, #21-22, #26-29) (Figure 19C, Appendix C1). Additional transformant colonies were identified in the CU522 cell line using genomic DNA isolate (#1-2, #4-5), and the three CU428 transformants were re-confirmed using the more sensitive procedure (Figure 19A, 19B).

Three mutant cell lines were selected for subsequent data analysis, based largely on the band intensities observed in each screening gel. For the CU428 cell line, mutant #2 was selected. CU522 was represented by mutant #19, which appeared in the initial colony PCR screens. CU725 mutant #13 was used.

Isolation of RNA from Wildtype and *rad23*Δ Cell Lines

Because *Tetrahymena thermophila* possess 45 copies of each somatic chromosome, analysis of gene expression was required to confirm that the *Rad23* gene locus was adequately silenced by the knockout construct. Additionally, total RNA isolation would allow for analysis of several related gene products, to help further characterize the effects of Rad23 loss. For each of the three cell backgrounds (CU428, CU522, CU725), four cultures were prepared: untreated wild-type; wild-type, 100 J/m^2 UV; untreated *rad23*Δ; *rad23*Δ, 100 J/m^2 UV. Using an RNeasy kit obtained from Qiagen, total RNA was isolated from each cell line; concentration and purity were quantified (Appendix D1), and samples were diluted to equal concentrations of 1 μg/μL. Due to the instability of RNA, reverse transcriptase PCR was conducted in order to obtain a workable volume of cDNA representative of each condition. These samples were used to obtain all subsequent quantitative (q) PCR work.

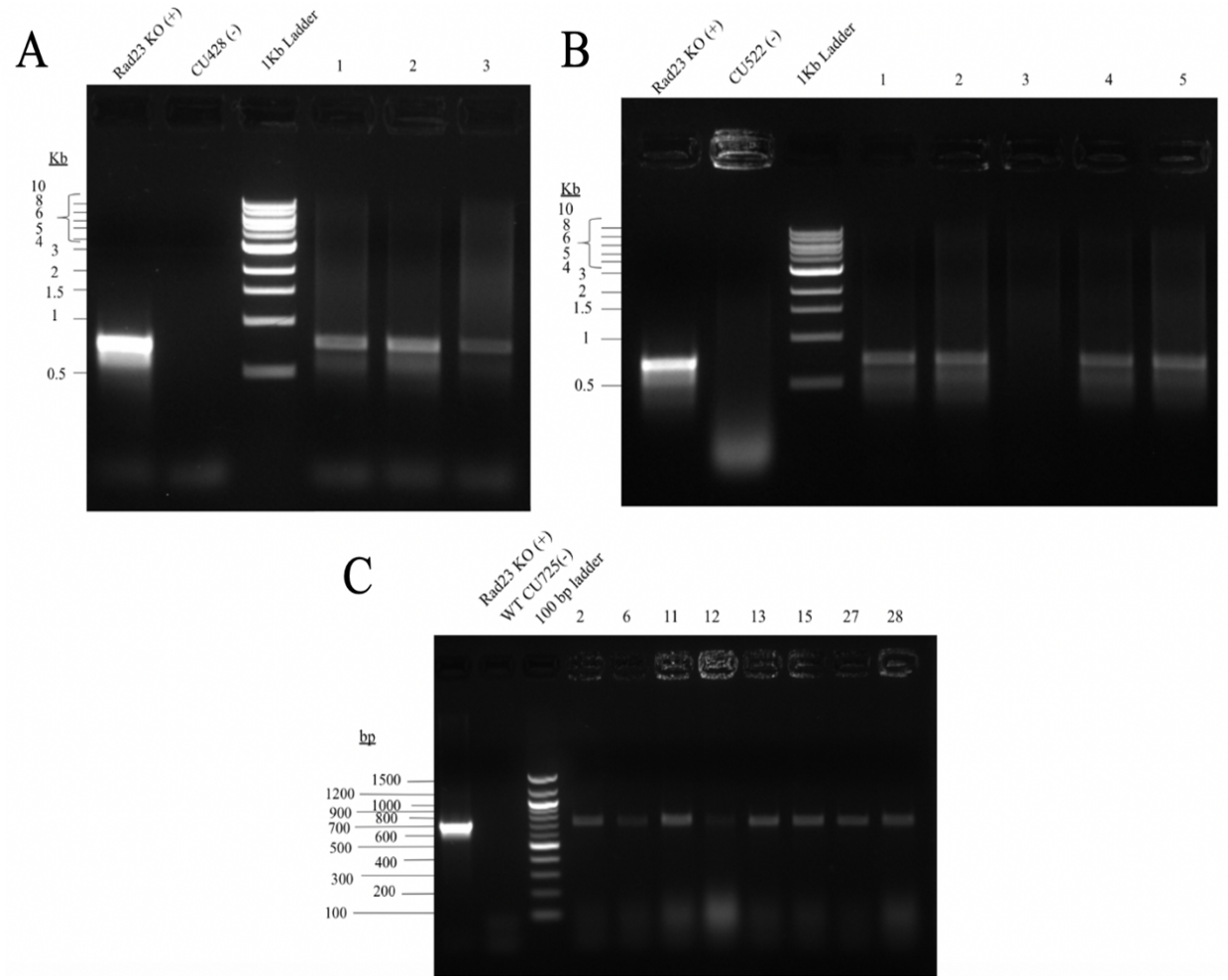


Figure 19: Colony and gDNA PCR Screening of *rad23*Δ *Tetrahymena thermophila.* *Tetrahymena thermophila* strains biolistically transformed with a Rad23 knockout construct were screened for recombination through either colony PCR or through PCR of isolated genomic DNA. In both cases, Rad23 -132F and MTT -815R primers were used for amplification. For all three backgrounds, the first lane contains the uncut plasmid as a positive control. The second lane contains untransformed *Tetrahymena* of the same background as the transformants, as a negative control. A 100 base pair or 1 kilobase ladder was used in lane 3 as indicated. The remaining lanes are individual clonal colonies of cells, arbitrarily numbered according to their plate positions. **A:** CU428 transformant gDNA PCR screening. **B**: CU522 transformant gDNA PCR screening of. **C:** CU725 transformant colony PCR screening.

RAD23 Expression Analysis

RAD23 expression levels were characterized in both wild-type and *rad23*Δ cell lines, before and after treatment with damaging UV-C irradiation. Following each cycle of PCR, a standard curve was generated using five standards of genomic DNA in varying concentrations (Appendix D2), which were in turn used to calculate the starting concentrations of *Rad23* and *Hhp1*. The first amplification products were also run on an agarose gel to confirm that only a single product of the predicted band size was produced from each sample (Appendix D3).

Single melt peaks (±0.5°C) were observed across all cell lines and in all replicates, with the exception of the CU725 *rad23*Δ mutants. These samples displayed an additional below-threshold peak at a slightly lower temperature, that is not thought to contribute to the calculated datapoints (Appendix D4, D5). In conjunction with the data obtained from the agarose gels, these melt peaks were used to confirm that all amplifications were specific to the genes of interest.

Amplification was measured in relation to mRNA transcripts: *Act1*, *Hhp1*, and *Rad23*. The actin-encoding gene, *ACT1*, was selected to serve as the standard (Appendix D4A). It is a conserved, single-copy gene, which makes it ideal for calculating starting concentrations of other single-copy genes within the model. Datapoints for the other two genes of interest were only collected if the resulting standard curve produced an r^2 value of at least 0.990. *HHP1* was selected as a normalization gene, as it is known to have fairly consistent expression when challenged with UV irradiation (Appendix D4B), and *Rad23* was analyzed to compare expression among wild-type and *rad23*Δ cells, both before and after treatment with UV (Appendix D4C).

In all samples, the *Hhp1* product amplified consistently across the cell line background,

regardless of *rad23*Δ mutation or UV treatment (Appendix D4B). The average starting quantity calculated for each of these samples were therefore used to normalize the expression levels of *RAD23*, so that any minor differences in expression due to total RNA or reverse transcriptase efficiency could be ignored.

Normalized values for each of the *Rad23* samples were subsequently set relative to the average *RAD23* expression observed in the wild-type, untreated cell lines. As such, the expression levels of untreated wild-types in each background (CU428, CU522, and CU725) were always set equal to 1.0 (Figure 20). This method of internal control ensured that minor experimental variations between replicates did not impact the overall landscape of the data.

The direct comparison of datasets revealed that *rad23*Δ mutants did show significantly decreased expression of *RAD23* compared to their wild-type counterparts, both before and after exposure to UV irradiation (Figure 20). This trend was most dramatically observed among the CU725 cell line family, in which *Rad23* was reduced over 5,000-fold in the *rad23*Δ mutant relative to the wild-type counterpart in the absence of UV damage. A decrease just shy of 14,000-fold occurred following UV exposure (Figure 20).

These results provided confidence that a phenotype associated with the loss of Rad23 could be adequately seen in the established cell lines. Specifically, the loss of Rad23 has been reported to depress the efficiency of nucleotide excision repair and enhance sensitivity to UV damage, resulting in decreased survival following stress. While all three mutants were analyzed for this purpose, the considerable efficiency of the CU725 knock-down line made the *rad23*Δ #13 mutant a major focus for subsequent analysis.

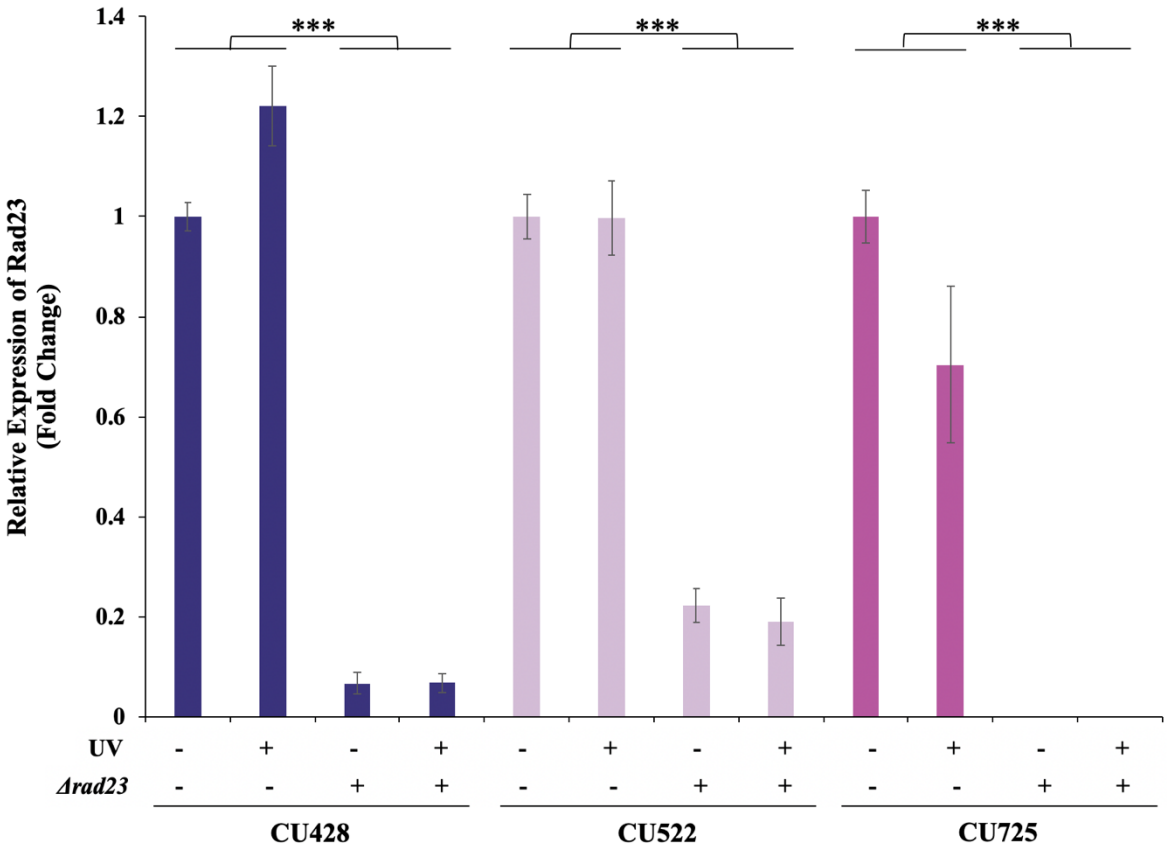


Figure 20: Relative Rad23 Expression in *T. thermophila* Cell Lines. Quantitative PCR analysis was performed on three CU cell lines, each with and without a *rad23*Δ mutation. The wild-type and mutant cell lines were each analyzed before and after treatment with 100 J/m^2 UV. Starting quantities were computationally determined using five genomic DNA standards of known concentration amplified using actin primers. Expression was normalized to a housekeeping gene, HHP1, and set relative to the average untreated wild-type expression for each cell background. Each dataset represents six to nine replicates. Cell backgrounds CU428 and CU522 were analyzed using a parametric two-way ANOVA. Cell background CU725 was analyzed using a nonparametric two-way ANOVA.

Characterization of *rad23*Δ Cell Lines: UV Survivability

As discussed in previous sections, loss of Rad23 has been shown to reduce survivability following exposure to UV damage in numerous model organisms. It was expected that decreased *RAD23* expression in the established *rad23*Δ cell lines would therefore present a similar phenotype when challenged with moderate intensity UV-C irradiation (75 J/m^2). Interestingly, the *rad23*Δ mutant cell lines showed consistently increased survival when challenged with UV when compared to a wild-type cell line of the same background (Figure 21). This effect was most substantially observed within the CU725 backgrounds, which was previously shown to express the most significantly depressed levels of *Rad23* (Figure 20). The mutant strain (CU725 *rad23*Δ #13) was so resistant to UV irradiation that the cells displayed over 150% survivability following exposure, actually producing more live wells in the treated cell populations than the untreated counterpart (Figure 21).

While the CU522 mutant strain exhibited a slightly less dramatic discrepancy, it was also significantly more resistant to UV irradiation when compared to a wildtype strain of the same cellular background. Conversely, the CU428 line, while appearing to be slightly more stable when challenged with UV, did not show significant differences from wild-type when analyzed. However, analysis of this strain also revealed an effect size of ~27.8%, indicating that additional trials and increased datapoints may show a similar, significant difference between wild-type and *rad23*Δ lines.

Due to the surprising nature of these initial results, the CU725 cell line was reexamined for UV survivability in the context of more severe damage. Both wild-type and *rad23*Δ CU725 strains were divided and treated with either 0 J/m^2 UV or 100 J/m^2 UV, as done in the Figure 21, and allowed to recover for one week prior to survivability quantification (Figure 22). Again,

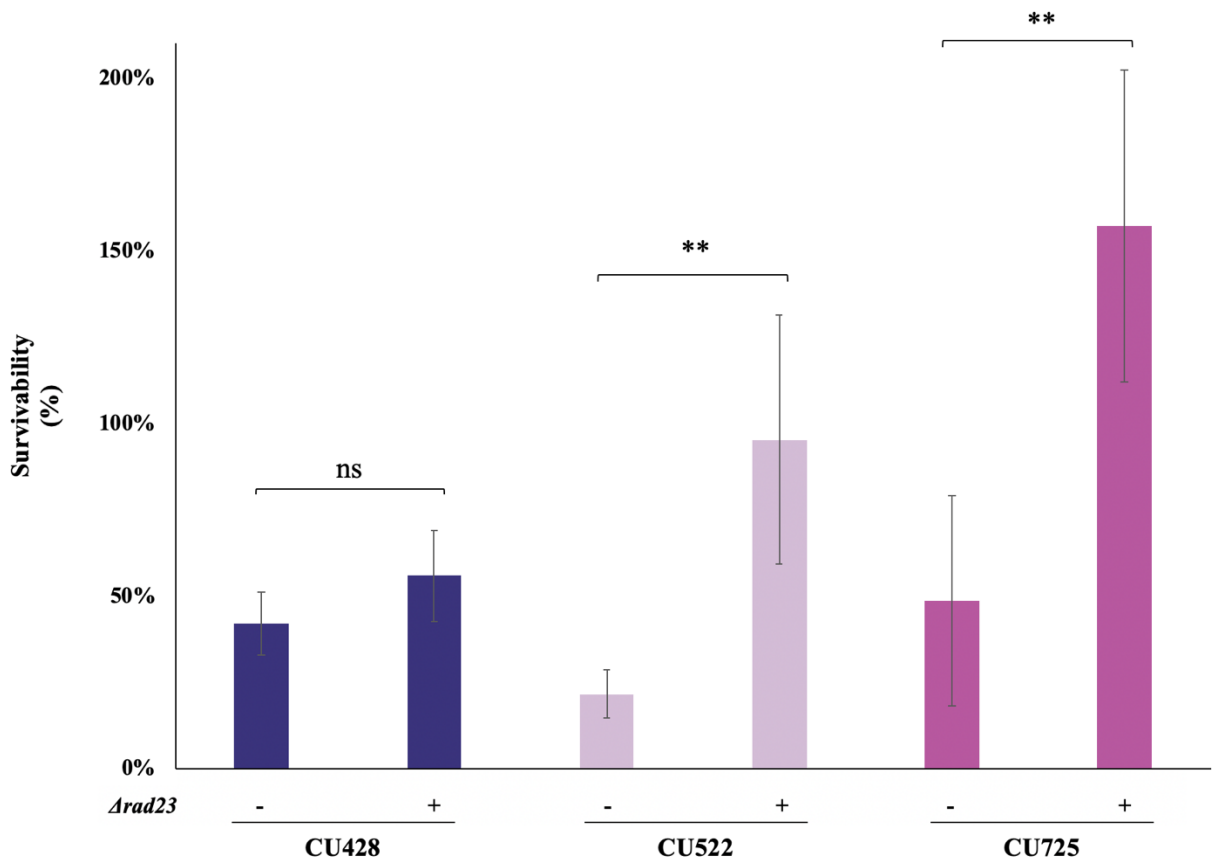


Figure 21: Rad23Δ Cell Survivability Following 75 J/m^2 UV-C Treatment. Wild-type and Rad23-depleted cell lines from the backgrounds CU428, CU522, and CU725 were treated with 0 J/m^2 UV-C and 75 J/m^2 UV-C irradiation. Following one week of recovery, living wells were counted and percent survival of each treated sample was calculated relative to the untreated counterpart (100% for each background). The graph depicts the average percent survival for each cell line and condition, relative to the corresponding untreated cell line. Error bars represent ± 1 standard deviation from the mean of all replicates represented. Data represents six replicates per sample. All cell backgrounds were analyzed using non-parametric two-tailed T-tests (Mann-Whitney).

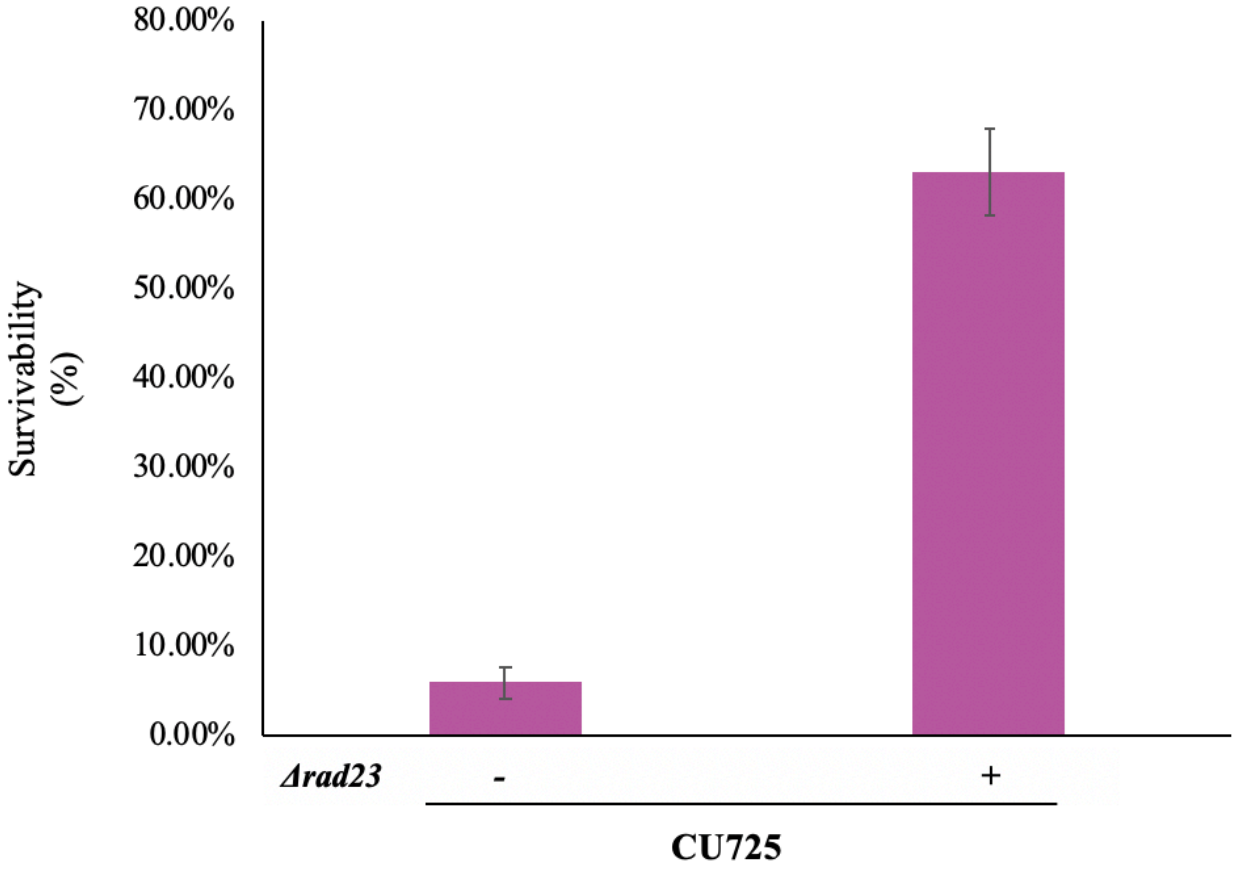


Figure 22: Rad23Δ Cell Survivability Following 100 J/m^2 UV-C Treatment. Wild-type (CU725) and Rad23-depleted (CU725 *rad23*Δ #13) cells were treated with 0 J/m^2 UV-C and 100 J/m^2 UV-C irradiation. Following one week of recovery, living wells were counted and percent survival of each treated sample was calculated relative to the untreated counterpart. The graph depicts the average percent survival relative to the corresponding untreated CU725 cell line. Error bars represent ±1 standard deviation from the mean of all replicates represented. Data is representative of six replicates per sample. Statistical analysis was not performed due to the limitations of the available datapoints.

these data supported increased stability of the strains exhibiting a *rad23*Δ mutation; the mutants lacking Rad23 survived 10 times better than the wild-type cells following treatment in initial trials. Initially, this assay was intended to simply confirm that the loss of Rad23 in *Tetrahymena thermophila* produced sensitivity to UV comparable to that seen in previous studies using *Saccharomyces cerevisiae* and other model organisms. These results indicate that loss of Rad23 has an unexpected, contradictory effect on UV resistance in *T. thermophila*.

Quantitative PCR: Additional DNA Repair Genes

The increased survival of Rad23-depleted cells following UV irradiation is largely unprecedented. With that said, it was hypothesized that the unique characteristics of *Tetrahymena thermophila* could be allowing other DNA repair pathways to compensate for the loss of efficient NER. To explore this possibility, four DNA repair proteins representing four distinct repair mechanisms, were analyzed via qPCR to determine whether the loss of Rad23 corresponded to altered transcriptional regulation.

APN2, *RAD6*, *RAD51*, and *TKU80* were selected for this analysis, due to their involvement in base excision repair, post-replication repair, homologous recombination repair, and non-homologous end joining, respectively. Relative expression levels of these genes in the context of both wild-type and Rad23-depleted CU725 cell lines, both before and after treatment with 100 J/m^2 UV, were compared using specific qPCR primers (Figure 23). It was determined that no significant changes to *APN2*, *RAD51*, or *TKU80* expression levels could be attributed to loss of Rad23; significance was only observed in treated vs. untreated populations regardless of the presence or absence Rad23. The RAD6 expression levels showed increased expression in Rad23-depleted cells only following treatment with UV. Based on these results, the resilience of

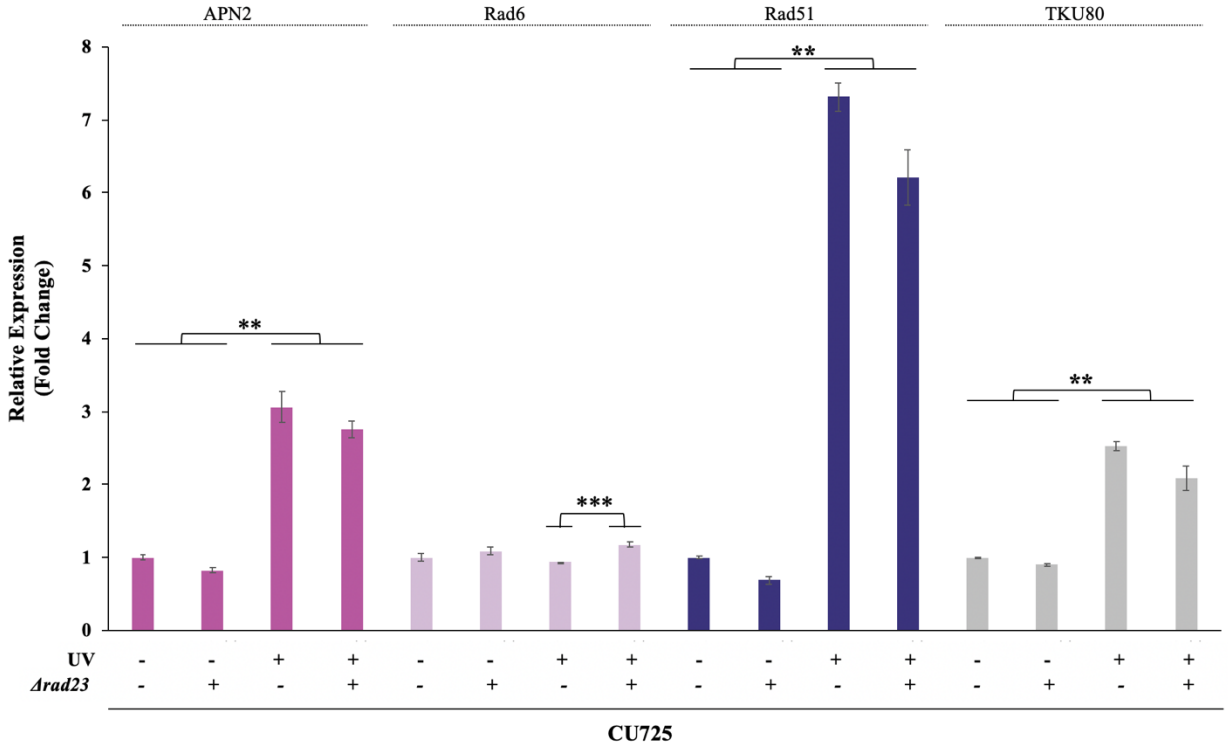


Figure 23: qPCR analysis of additional DNA repair gene expression levels. DNA repair genes APN2 (BER), RAD6 (PRR), RAD51 (HRR), and TKU80 (NHEJ) were analyzed for expression differences in between wild-type and Rad23-depleted CU725 cell lines, both before and after treatment with 100 J/m^2 UV. APN2, RAD51, and TKU80 were analyzed using a nonparametric two-way ANOVA test; RAD6 was analyzed using a parametric two-way ANOVA test. Error bars represent $\pm$1 standard deviation from the mean of all replicates represented. Data represents six replicates per sample. Datasets corresponding to the genes APN2, RAD51, and TKU80 were analyzed using an non-parametric two-way ANOVA. The Rad6 dataset was analyzed using a parametric two-way ANOVA.

the *rad23*Δ cell lines could not be easily attributed to compensation by base excision repair or double-stranded break repair pathways, and other methods of analysis had to be pursued.

Generating a Rad23 Overexpression Cell Line

With an unusual phenotype of UV resistance established for the Rad23-depleted *T. thermophila* cell lines, and no apparent smoking gun provided by qPCR analysis of other DNA repair genes, it was decided to establish additional cell lines in order to further elucidate potential mechanisms of UV stability. A Rad23 rescue cell line was attempted through the addition of a RAD23 gene back into the *rad23*Δ background. A Rad23 overexpression cell line was attempted through the addition of Rad23 into a wild-type cell line already expressing normal, endogenous levels of the protein.

A 2HA-tagged Rad23 plasmid was selected for this transformation process (Appendix A2). Once it had been isolated and sequenced, the plasmid was ready to be prepared for biolistic transformation into the *T. thermophila* cell lines. As described previously, *T. thermophila* require linear DNA fragments separate from any bacterial plasmid backbone to facilitate successful transformation. Plasmids were therefore digested with the restriction enzyme Spe1 following their isolation from DH10B cells. Of the two colonies screened by agarose gel following digestion, both produced the predicted band sizes of a successful enzyme digest, and were therefore both utilized for biolistic transformation into wild-type CU725 cells and the CU725 *rad23*Δ #13 mutant strain (Figure 24).

Following selection of transformant *Tetrahymena* colonies in Paclitaxel, GoTaq PCR was utilized to screen for positive transformants and to eliminate any false positive colonies. Initially, these screenings were done using colony PCR on whole cell templates. Some cell lines failed to

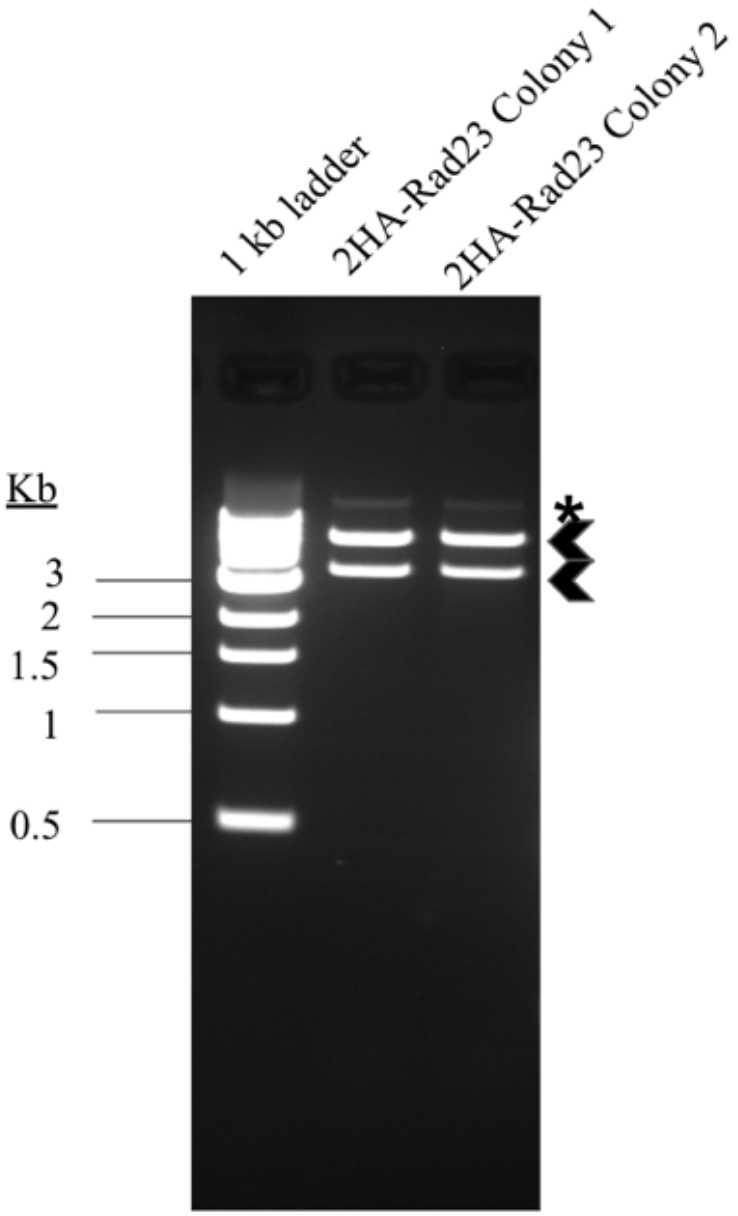


Figure 24: 2HA-Rad23 Plasmid Preparation for Biolistic Transformation. Prior to transformation into *T. thermophila*, plasmids were restriction enzyme digested to linearize genes and remove the bacterial sequences of the plasmid backbone. The 2HA-Rad23 plasmid, isolated from two distinct transformant DH10B colonies, was digested with SpeI. Predicted band sizes for this digest were 5318 and 3211 base pairs, indicated by arrowheads. The asterisk (*) represents undigested plasmid.

produce conclusive results through the use of this method, however, and were therefore screened using isolated genomic DNA as the template source.

The success of the rescue attempt to add exogenous Rad23 back into the *rad23*Δ cell line was unfortunately inconclusive based on these two methods. Where a clean band at the size corresponding to the 2HA-Rad23 positive control lane would indicate incorporation of the recombinant DNA into the genome, the PCR products of the *rad23*Δ background produced only nondescript smears (Appendix C2). The overexpression cell line, however, produced distinct banding patterns more capable of supporting or contradicting successful transformation (Figure 25). Six clonal colonies were identified as likely transformants based on the presence of a band at the size indicated by the positive control product: #2, #3, #5, #6, #8, #9 (Figure 25). These strains were maintained in stock tubes for additional confirmation and analysis.

Confirmation of 2HA-Rad23 Protein Expression

Because of the inconclusive results of the CU725 *rad23*Δ #13 mutant screen, and in order to further verify the over-expression of Rad23 in the wild-type CU725 background, proteins were isolated from both 2HA-Rad23 transformed strains for use in a western blot analysis. Additionally, as the 2HA-Rad23 gene was recombined under the control of a cadmium chloride inducible MTT1 promoter, cells were either left uninduced or were treated with $CdCl_2$ prior to protein isolation, in order to assess the levels of protein expression under either cellular condition. Anti-HA antibodies were used in order to visualize only the expression of the exogenous Rad23 protein, without any background contamination from endogenous levels.

Ponceau staining of each nitrocellulose membrane prior to visualization confirmed that protein transfer had occurred in both cell backgrounds (Appendix E1 and E2) Unfortunately, the

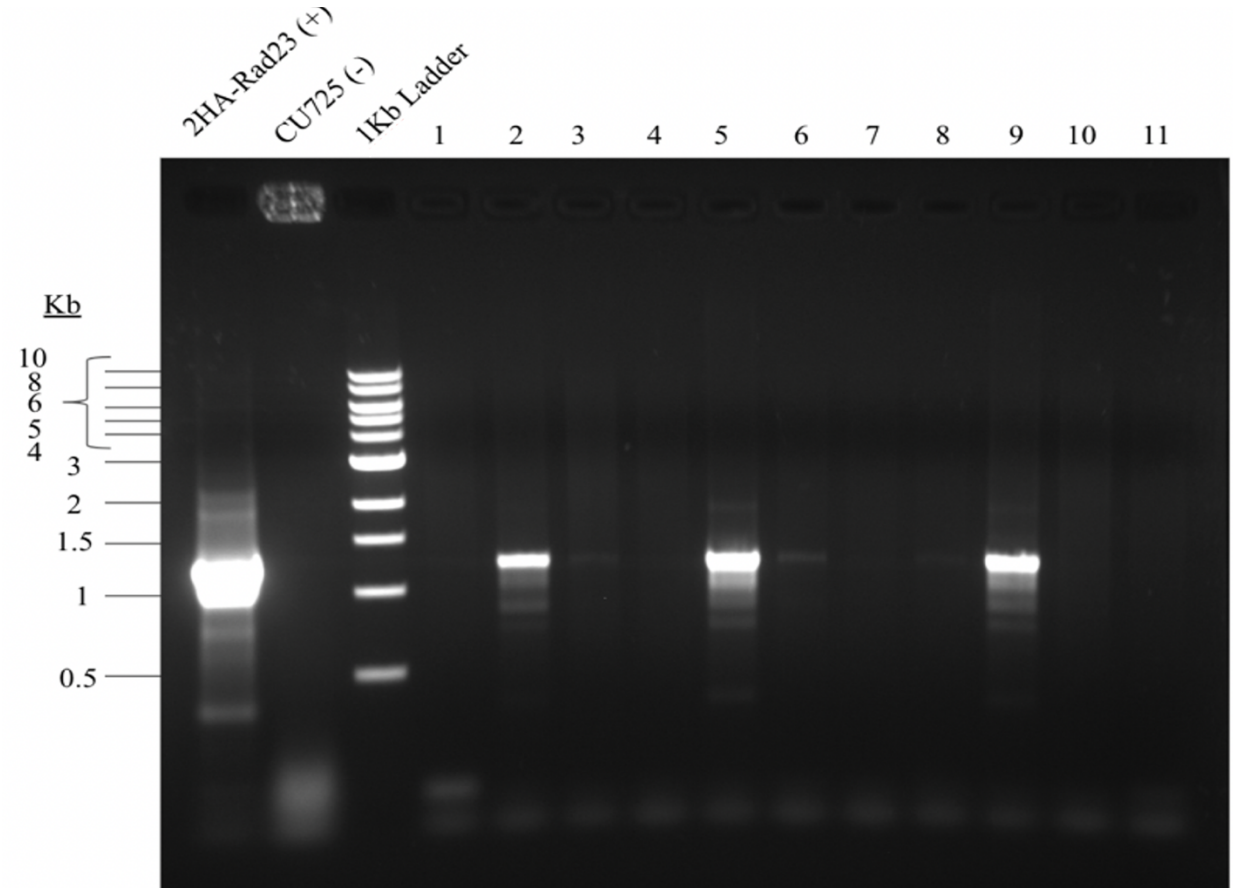


Figure 25: Colony PCR Screening of 2HA-Rad23 *Tetrahymena thermophila.* CU725 *Tetrahymena thermophila* was biolistically transformed with a 2HA-Rad23 construct and screened for recombination. Using the MTT-151F and Rad23 E3R primer pair. The first lane contains the uncut plasmid template as a positive control. The second lane contains untransformed wild-type CU725 cells as a negative control, followed by a 1 Kb ladder in lane 3. The remaining lanes represent individual clonal colonies of 2HA-Rad23 transformants of CU725, arbitrarily numbered according to their plate positions.

Western Blot analysis revealed that the CU725 *rad23*Δ #13 cell lines did not express any perceivable 2HA-Rad23 protein, which was likely the cause of the inconclusive streaking pattern observed on the GoTaq PCR gel screenings (Appendix C2, Appendix E3). This made analysis of a rescue phenotype unachievable, but provided an important potential future direction.

Fortunately, the CU725 cell line transformed with 2HA-Rad23 was confirmed to express the tagged recombinant protein, both prior to and following induction with $CdCl_2$ (Figure 26). As the MTT1 promoter is known to be "leaky" in some contexts, allowing expression to occur in the absence of heavy metal promoter induction, the presence of expression in the uninduced fractions was not surprising. Still, five confirmed overexpression cell lines provided a unique opportunity to analyze both ends of the Rad23 expression spectrum: significant downregulation (*rad23*Δ), and observable overabundance (2HA-Rad23).

UV Survivability of Overexpression Cell Lines

As the depletion of Rad23 had been shown to significantly decrease cellular sensitivity to UV damage, it was possible that an overexpression cell line could show decreased resistance when challenged with the same conditions. To test the possibility, wild-type CU725 cell lines were matched with the 2HA-Rad23 transformants of the CU725 background, and each dosed with the standard 0 J/m^2 and 75 J/m^2 UV-C. Following the incubation and recovery period, the 2HA-Rad23 transformants not only showed no increased sensitivity, but showed a significantly increased resistance (Figure 27). The phenotype, while not as strong as that observed in the Rad23-depleted CU725 line, does provide some insight into what may be occurring at a

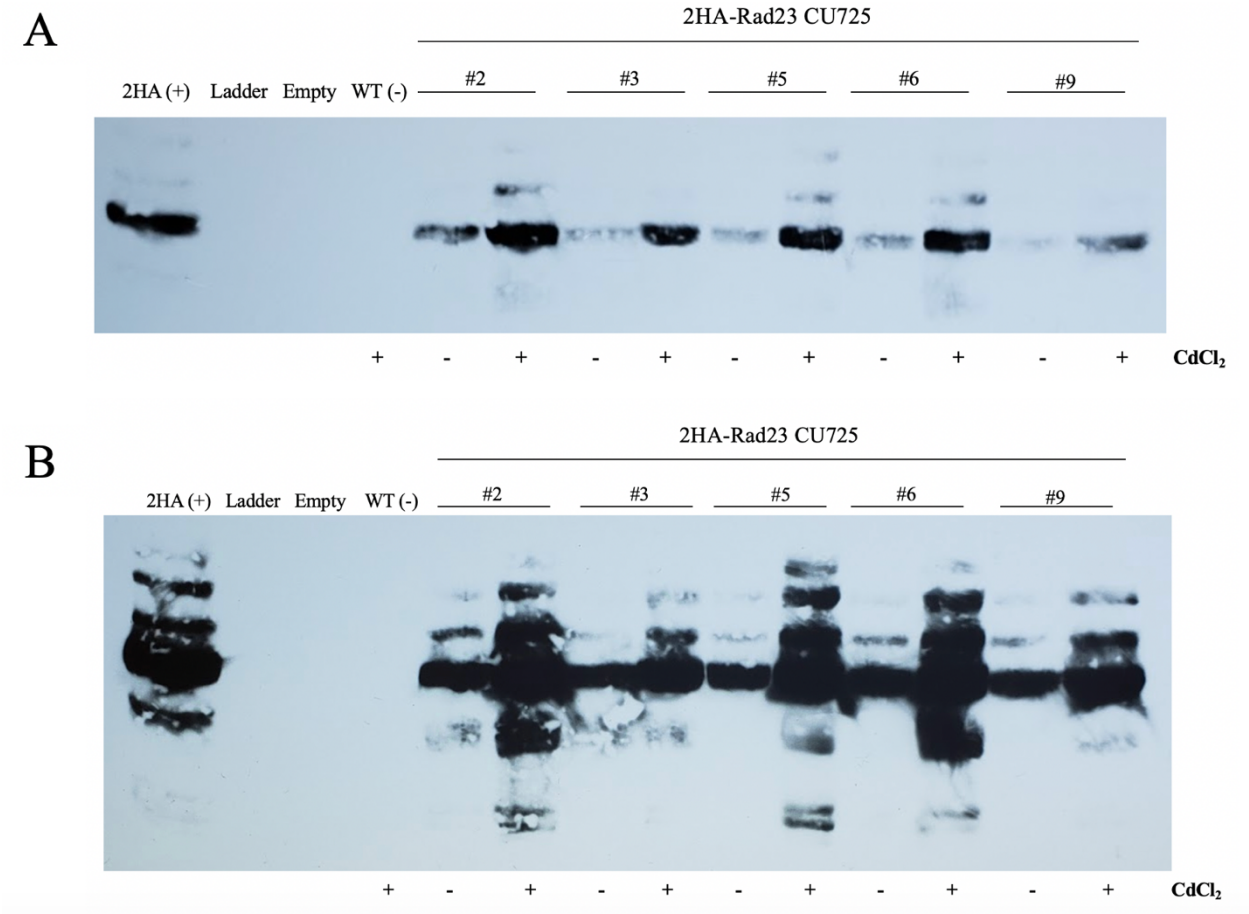


Figure 26: Western Blot Confirmation of 2HA-Rad23 Expression. Protein isolates from CU725 cells transformed with a 2HA-Rad23 plasmid were analyzed via Western Blot. CU522 transformed 2HA-Rad23 (Carpenter) was used as a positive control in the first lane, followed by a protein ladder and protein isolates from wild-type CU725 cells as a negative control. The five cell lines analyzed were treated with or without $CdCl_2$ to induce activity of the MTT1 promotor prior to isolation. Exposure times of 3 seconds (**A**) and 1 minute (**B**) are shown.

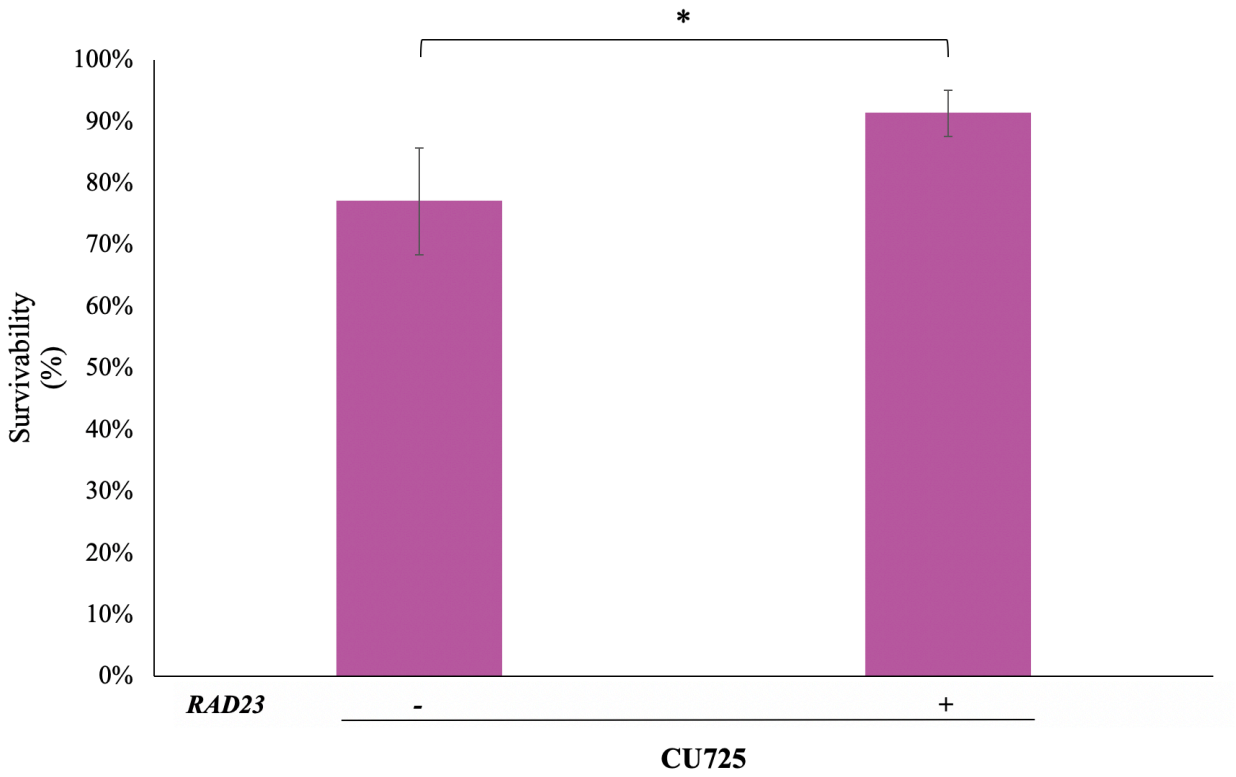


Figure 27: Rad23 Overexpression Cell Survivability Following 75 J/m^2 UV-C Treatment. Wild-type and 2HA-Rad23 (#5) overexpression cell lines from the background CU725 were treated with 0 J/m^2 UV-C and 75 J/m^2 UV-C irradiation. Following one week of recovery, living wells were counted and percent survival of each treated sample was calculated relative to the untreated counterpart. The graph depicts the average percent survival for each cell line and condition, relative to the corresponding untreated wild-type CU725. Error bars represent ± 1 standard deviation from the mean of all replicates represented. Data represents six replicates per sample. Significance was analyzed using a non-parametric two-tailed T-test (Mann-Whitney).

molecular level following UV. While the loss of Rad23 may disrupt any number of cellular process, resulting in a disregulated system and the reduced cell death following damage, an increased pool of the protein is not enough, on its own, to produce an inverse effect. There must be additional factors at play, and this will be further explored throughout the discussion and future direction sections.

Mutagenesis

The original goal of this research project was to determine which ubiquitin residues, if any, are required for the efficient function of Rad23 in nucleotide excision repair. This was to be accomplished through site-directed mutagenesis of a handful of lysine residues which reside in the ubiquitin-like (UbL) domain of the Rad23 protein. The 2HA-Rad23 plasmid discussed previously in the context of Rad23 overexpression was to serve as the template for mutagenesis, and the *rad23*Δ cell lines were to be used as the background for biolistic transformation to ensure that the exogenous mutant could function without endogenous Rad23 competition. Science, however, rarely conforms to these ideal plans and timelines.

The Kunkel method of site-directed mutagenesis requires a single-stranded DNA template, which is mixed with a variety of mutagenic primers in a single reaction step to generate multiple unique mutation combinations. Unfortunately, multiple attempts at isolation failed to produce single-stranded DNA (Appendix F), so the Phusion method of mutagenesis was selected as an alternative. Establishment of the "lysine-less" mutants, lacking all conserved lysine residues within the UbL, was not feasible due to the time constraints of the project, paired with the more time-consuming nature of the Phusion mutagenesis method. The complete protocol to do so was created, however, along with a detailed plan regarding the order and combinations of

mutations to be induced (Figure 28). Because several of the mutagenic primers (Table 5) were designed to cover multiple point mutation sites, these three progressions would provide an efficient pathway to a variety of partially and fully mutagenized 2HA-Rad23 constructs.

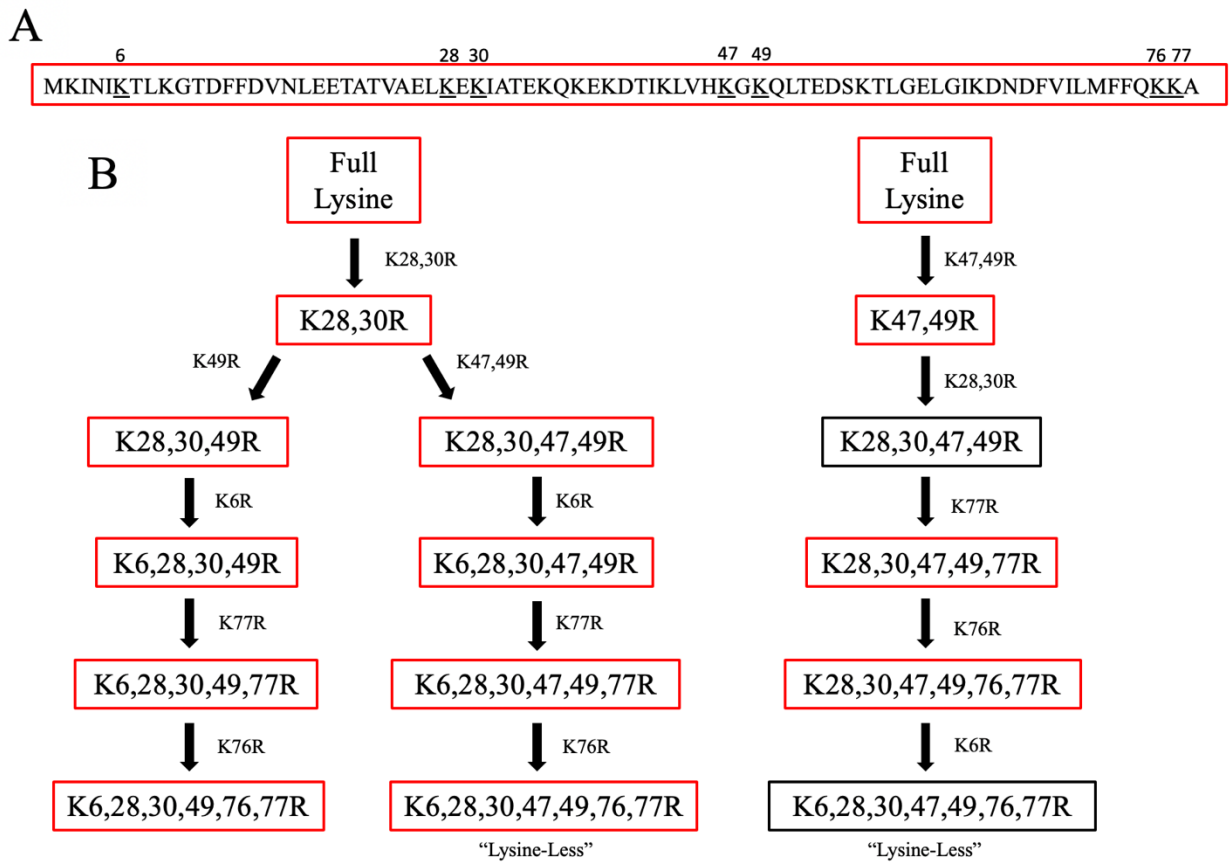


Figure 28: Proposed Mechanism of Rad23 Mutagenesis. A: The Rad23 UbL domain contains 15 lysine residues, 7 of which were targeted for mutagenesis. Lysine residues 6, 28, 30, 49, 76, and 77 were chosen due to their conservation, as well as lysine 47 due to its proximity to conserved residue 49 and easy inclusion into the required K49R mutagenic primer. **B:** A possible order of mutagenesis to be performed on the 2HA-Rad23 plasmid using the Phusion method would ensure that a variety of mutants, especially those with disruption of the most highly conserved UbL lysines, would be obtained in the fewest number of steps. The primer used (Table 5) at each step appears besides the corresponding arrow, and the resulting plasmid is boxed below. Twelve unique mutation combinations are boxed in red, while two repeated products appear in black. The varieties established could be used to identify which residues, if any, are necessary and/or sufficient to support efficient NER *in vivo*.

DISCUSSION

Rad23 is highly conserved across organisms in sequence, structure, and function (Tables 8, 9; Figures 16, 17). It is largely accepted that Rad23 plays a role in efficient GG-NER, and that that loss of the protein results in increased sensitivity to UV irradiation. The unprecedented UV resistance of *rad23*Δ *Tetrahymena thermophila* cell lines was therefore a striking new development for the field of DNA repair research.

Studies similar to this one have been done in labs all across the world, using common protocols and model organisms; the loss of Rad23 has consistently been shown to increase sensitivity to UV damage. The data presented here are not nearly sufficient to contradict the findings of these previous studies, nor would these findings follow the logic of the established system; loss of Rad23 *should* decrease NER efficiency, and therefore reduce survivability in response to UV damage. This level of divergence would not be expected according to its phylogenic similarities to the commonly used *S. cerevisiae* model, and therefore *T. thermophila* may provide a fresh, new window into studying the complexities of genomic stability. Despite the conservation of the protein, and presumably its activities between models, there is clearly something out of place that allows these cell lines to evade death.

The effect of Rad23 loss on UV survivability was assessed by comparing UV treated cell lines to otherwise-identical untreated counterparts to establish percent survival. Percent survival was then directly compared between the three "wild-type" cell backgrounds, CU428, CU522, and CU725, and their Rad23-depleted counterparts, CU428 *rad23*Δ #2, CU522 *rad23*Δ #19, and CU725 *rad23*Δ #13, respectively. This method provided a great deal of internal control in assessing the significance of survivability changes, but also ensured that a single cell background

was not used in order to evaluate all of the data. With the exception of the CU428 pairing, which still produced a reasonably promising effect size of 28%, the other two backgrounds showed significantly increased survival when Rad23 was depleted below wild-type expression levels (Figure 21, 22).

One important trend to note in unraveling these unusual results is the variation between the "wild-type" CU428 cell background and the two β-tubulin mutant backgrounds, CU522 and CU725. The β-tubulin mutation serves an important purpose in recombinant DNA technique, in that the altered BTU2 locus provides an efficient target for inducing knock-in mutations. While this mutation is not known to have major a major impact on overall operations within the cell, it has been shown to slow division rate (4-5 hours versus a normal 2.5-3 hours). It is therefore possible that UV-induced genomic instability is accumulating in all three strains at a comparable rate, but that the slowed division of the CU522 and CU725 lines may enhance, inhibit, or otherwise alter the natural damage response in cells that lack Rad23. In other words, the varied growth rate of these cells may provide a way to look at these pathways in a new context, to better understand the chain of events that occurs when an organism incurs stress.

Initially, it was thought that a prolonged window for repair could allow other DNA repair pathways to compensate, or even over-compensate, for the loss of fully efficient NER. While this still could be the case, qPCR analysis of several members of these pathways revealed that loss of Rad23 did not account for significant up-regulation in gene expression in any pathway except post-replication repair (Figure 23). As base excision repair is the pathway most similar to NER, and is the only mechanism shown to have some compensatory capacity for repairing NER lesions (Fleck and Nielsen 2004), it is unlikely that a small upregulation of Rad6 at the transcriptional level could provide the degree of UV resistance observed.

The next step was to analyze the effects of Rad23 overexpression *in vivo*. If increased expression produced reciprocal effects compared to the *rad23*Δ mutants, then expression level itself could be tied to survivability. Interestingly, the 2HA-Rad23 overexpression cell lines *also* showed increased survival following UV damage (Figure 27), though not as significant as that observed in the *rad23*Δ mutants. The increase in UV resistance is consistent with slightly improved efficiency of NER; it is not consistent with the possibility that Rad23 levels are inversely related to survivability in *T. thermophila*. With no low-hanging fruit left to explain the resistance, future research must explore more complex mechanisms to explain the phenomenon (See FUTURE DIRECTIONS).

Reflecting over the original aims of this research project, most were met only partially. The establishment and phenotypic analysis of a "lysine-less" Rad23 *Tetrahymena thermophila* cell line remains an unwritten chapter in the narrative (Figure 28). However, the *rad23*Δ cell line provided a unique new look into the understanding of this versatile protein. As this phenotype has not been observed in the context of other organisms, this *Tetrahymena thermophila* model could provide a rare perspective in examining apoptosis, DNA damage checkpoint response, gene expression, and post-translational modification (i.e. ubiquitin code). The potential remains also to identify whether or not the ubiquitylation patterns of Rad23 affect the function of the protein *in vivo*, and to analyze the effects of these modifications on transcriptionally silent micronuclear global genome repair. This project remains a large step in the proper direction, and has provided an unexpected number of new avenues for future researchers to pursue. The fulfillment of the remaining original goals, and the exploration of the related research targets not previously identified, could significantly contribute to this complex tapestry surrounding Rad23.

FUTURE DIRECTIONS

This project is still very much in its infancy; these data have raised more questions than they have successfully put to rest. In order to explore these issues, additional cell lines must be established, many of which require the completion of the Rad23 "lysine-less" UbL constructs. The protocol for inducing the desired mutations has been established as part of this project, as well as a proposed order for cumulating the point-changes (Figure 28). The potential creation of 12 partially- and completely-mutagenized 2HA-Rad23 constructs would serve as a way to explore the importance of the targeted residues in efficient NER function, an original aim proposed for this project. A truncated 2HA-Rad23 construct lacking the UbL domain entirely could also be added to the mix, to more thoroughly reflect the previous work done in yeast (see Preliminary Data). These sequences, alongside the unmutagenized 2HA-Rad23 construct, would be transformed into the established Rad23-depleted cell lines (particularly CU725 *rad23*Δ #13) to observe the "rescue" potential within the context of UV survivability. These strains can be further used in a number of additional assays, in order to assess the efficiency of other Rad23-dependent pathways discussed in more detail below.

Challenging the cell lines with alternative damaging agents, such as MMS or H_2O_2 would provide an idea as to whether or not the observed phenotype is NER specific, or may be impacting other DNA damage response pathways, such as double-stranded break repair, post-replication repair, and base excision repair. This could be done currently using the *rad23*Δ cell lines, to determine whether Rad23-loss plays any roles in DNA damage response beyond its activity in NER, but inclusion of the mutant Rad23 constructs could provide additional insight into the ubiquitylation-based regulation of these mechanisms.

Also listed among the original aims, and pushed aside due to the unusual nature of the data, *in vivo* NER assays would more thoroughly measure the effects of Rad23 loss and mutation. In most previous studies, loss of Rad23 or Rad23 function has correlated with decreased efficiency in NER, which results in accumulation of DNA lesions and eventual cell death. As increased cell death is not observed in these mutant lines, it is possible that assessment of NER *in vivo*, especially targeting the transcriptionally silent, GG-NER dependent micronucleus, may provide a better understanding of whether or not this new phenotype is actually governed by the targeted DNA repair pathway. If micronuclear DNA is, in fact, being damaged and left unrepaired, it would follow that impairment of Rad23's role is masked by the lack of transcriptional requirement in the micronucleus. While UV survivability assays look only at the overall fate of the cell, an NER assay would reveal any damage that accumulates in the micronucleus.

These assays could be further complemented by the analysis of strains expressing the "lysine-less" or truncated UbL-less Rad23 mutants, to determine which residues, if any, are required for the performance of this function. Because both wild-type and mutant Rad23 proteins will have been transformed into *Tetrahymena thermophila* lines with a conjugated 2HA-tag, ubiquitin shift assays could be used to further determine how post-translational modification correlates to the efficiency of the repair. Utilization of a Rad4 antibody (available commercially or from collaborating labs) could be used to assess the stability levels of Rad4 in the Rad23-depleted or Rad23-modified cell backgrounds. Through these simple procedures, the existing and planned cell lines could be used to dig much deeper into this unique phenotype, while still focusing entirely on the players already outlined in the original hypothesis.

Moving beyond the scope of the original goals, and utilizing the cell lines that have been and will be established, there are several new directions in which this project could proceed. As Rad23 is closely linked to the ubiquitin proteasome system and degradation, research could extend into the effects of the mutations on both NER and the UPS hand-in-hand, rather than as two separate sides of the same coin. It would be interesting to see whether eliminating post-translational modification potential would push Rad23 to be more active in one pathway as opposed to another, and common protocols, such as cycloheximide chase assays using Rad23-dependent proteolytic substrates and the aforementioned *in vitro* NER assays, could be used to see how the wild-type and "lysine-less" mutants act in either path prior to or following UV damage.

Interaction between wild-type and "lysine-less" mutant Rad23 could also be examined through the use of co-immunoprecipitation, again thanks to the 2HA-tag that would be conjugated to both wild-type and mutant protein. This process could be extremely helpful in identifying ubiquitylation patterns required for Rad23 interaction with its numerous partners, and potentially in identifying the conditions in which one partner is preferred over another. Specifically, Rad4 (NER), Rpn10 (UPS), Png1 (ERAD), E2F1 (expression), Snf2 (expression), and AIF1 (Apoptosis) and could be targeted to dramatically narrow the role of Rad23 under different post-translational conditions and cellular contexts.

To give some context to the genes mentioned above, it has already been established that the post-translational modifications to Rad23's UbL domain may be the key to determining its participation in NER over the UPS or vice versa (see INTRODUCTION). It is therefore possible that examining the pool of Rad23 interacting with Rad4, versus the pool interacting with Rpn10, could illustrate how the mutations to the UbL tips the scale towards one pathway over the other.

Beyond these two main routes, Rad23 has been shown to interact with Endoplasmic Reticulum Associated Degradation (ERAD) protein Png1 via its Rad4 binding domain, in order to facilitate efficient degradation of misfolded proteins (Dantuma *et al.* 2009). Loss of this function could increase pools of aberrant proteins, which may alter the ability of the cell to respond appropriately to DNA damage. Transcription factor E2F1 also associates with the Rad4 binding domain of Rad23. Rad23 protects E2F1 from proteasomal degradation in a fashion similar to its protection of Rad4, and E2F1 modulates the subcellular localization of Rad23, allowing the protein to localize to CPDs in the nucleus in response to UV induced damage (Singh and Dagnino 2016). Loss of Rad23 would reduce the stability of E2F1, which could have serious effects in transcriptional regulation.

While these pathways are slightly less explored, Rad23 also interacts with the protein kinase Snf2 to modulate the expression of DNA damage response genes in response to UV exposure, and with Rad4 in a transcriptional regulatory capacity to inhibit production of ribonucleotide reductase transcripts in the absence of damage (Wade *et al.* 2009; Zhou *et al.* 2015). Along with the increased mutation rate associated with inefficient NER, the loss of Rad23 could therefore affect the cell at a transcriptional level through dysregulation of genes required for normal damage response. Without further analysis, there is no way to know exactly what the implications could be. It is possible that, through a collaboration with the Institute of Molecular Biology in Mainz, Germany, a complete proteomic analysis of wildtype or mutant Rad23-expressing strains could reveal some of these potential consequences. Perhaps most compelling, Rad23 acts as a shuttle for apoptosis inducing factor AIF1 from the mitochondria to the nucleus in order to induce programmed cell death (Sudhakar and Chow 2014). This shuttling is a crucial requirement in the caspase-independent pathway of apoptosis, and therefore the loss of Rad23

could inhibit normal cell death. TUNEL assays could confirm this phenomenon, which alone could explain the increased UV survivability observed in the *rad23*Δ cell lines.

In short, this project poses more questions than it answers. It is known that this cell line displays a unique phenotype not seen in previous studies using any other known model organism. It is also understood that this phenotype defies the common consequence of Rad23 and NER loss. What remains unclear is the reason that this UV resistance occurs, and whether the mechanisms at play are unique to *Tetrahymena thermophila*, or if nuclear dimorphism may simply allow a more prominent display of what occurs universally in damage-incurring cells prior to or in concert with apoptosis. If the latter is proven to be true, this model could serve as an entirely unique opportunity to study details of damage response never before revealed, and to better understand exactly how Rad23 serves as a mediator of these inseparable pathways. These data merely represents chapter one of an ongoing saga surrounding Rad23.

REFERENCES

Bergink S., W. Toussaint, M. S. Luijsterburg, C. Dinant, S. Alekseev, *et al.*, 2012 Recognition of DNA damage by XPC coincides with disruption of the XPC–RAD23 complex. J. Cell Biol. 196: 681–688.

Benaroudj N., P. Zwickl, E. Seemüller, W. Baumeister, and A. L. Goldberg, 2003 ATP hydrolysis by the proteasome regulatory complex PAN serves multiple functions in protein degradation. Mol. Cell 11: 69–78.

Berneburg M., and A. R. Lehmann, 2001 3 Xeroderma pigmentosum and related disorders: Defects in DNA repair and transcription, pp. 71–102 in *Advances in Genetics*, Academic Press.

Bertolaet B. L., D. J. Clarke, M. Wolff, M. H. Watson, M. Henze, *et al.*, 2001 UBA domains of DNA damage-inducible proteins interact with ubiquitin. Nat. Struct. Biol. 8: 417–422.

Bootsma D., 2001 Nucleotide excision repair syndromes: xeroderma pigmentosum Cockayne syndrome, and trichothiodystrophy. The metabolic and molecular bases of inherited disease 677–703.

Brown T. A., 2002 *Mutation, Repair and Recombination*. Wiley-Liss.

Carpenter, A., 2012 Characterization of Rad23 in *Tetrahymena thermophila*. Unpublished.

Chen L., U. Shinde, T. G. Ortolan, and K. Madura, 2001 Ubiquitin-associated (UBA) domains in Rad23 bind ubiquitin and promote inhibition of multi-ubiquitin chain assembly. EMBO reports 2: 933–938.

Chen L., and K. Madura, 2002 Rad23 promotes the targeting of proteolytic substrates to the proteasome. Mol. Cell. Biol. 22: 4902–4913.

Coin F., V. Oksenych, and J.-M. Egly, 2007 Distinct roles for the XPB/p52 and XPD/p44 subcomplexes of TFIIH in damaged DNA opening during nucleotide excision repair. Mol. Cell 26: 245–256.

Collins K., and M. A. Gorovsky, 2005 Tetrahymena thermophila. Curr. Biol. 15: R317–8.

Compe E., and J.-M. Egly, 2012 TFIIH: when transcription met DNA repair. Nat. Rev. Mol. Cell Biol. 13: 343–354.

Dantuma N. P., C. Heinen, and D. Hoogstraten, 2009 The ubiquitin receptor Rad23: at the crossroads of nucleotide excision repair and proteasomal degradation. DNA Repair 8: 449–460.

De Boer J., and J. H. Hoeijmakers, 2000 Nucleotide excision repair and human syndromes. Carcinogenesis 21: 453–460.

De Laat W. L., E. Appeldoorn, and K. Sugasawa, 1998 DNA-binding polarity of human replication protein A positions nucleases in nucleotide excision repair. Genes.

Elsasser S., D. Chandler-Militello, B. Müller, J. Hanna, and D. Finley, 2004 Rad23 and Rpn10 serve as alternative ubiquitin receptors for the proteasome. J. Biol. Chem. 279: 26817–26822.

Fagbemi A. F., B. Orelli, and O. D. Schärer, 2011 Regulation of endonuclease activity in human nucleotide excision repair. DNA Repair 10: 722–729.

Fischer E. S., A. Scrima, K. Böhm, S. Matsumoto, G. M. Lingaraju, *et al.*, 2011 The molecular basis of CRL4DDB2/CSA ubiquitin ligase architecture, targeting, and activation. Cell 147: 1024–1039.

Fitch M. E., S. Nakajima, A. Yasui, and J. M. Ford, 2003 *In vivo* recruitment of XPC to UV-induced cyclobutane pyrimidine dimers by the DDB2 gene product. J. Biol. Chem. 278: 46906–46910.

Fleck O., and O. Nielsen, 2004 DNA repair. J. Cell Sci. 117: 515–517.

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

García-de Teresa B., M. Hernández-Gómez, and S. Frías, 2017 DNA Damage as a Driver for Growth Delay: Chromosome Instability Syndromes with Intrauterine Growth Retardation. Biomed Res. Int. 2017: 8193892.

Gartenberg M. R., and J. S. Smith, 2016 The Nuts and Bolts of Transcriptionally Silent Chromatin in Saccharomyces cerevisiae. Genetics 203: 1563–1599.

Ghaemmaghami S., W.-K. Huh, K. Bower, R. W. Howson, A. Belle, *et al.*, 2003 Global analysis of protein expression in yeast. Nature 425: 737–741.

Gillet L. C. J., and O. D. Schärer, 2006 Molecular mechanisms of mammalian global genome nucleotide excision repair. Chem. Rev. 106: 253–276.

Glickman M. H., and A. Ciechanover, 2002 The ubiquitin-proteasome proteolytic pathway: destruction for the sake of construction. Physiol. Rev. 82: 373–428.

Gödderz D., T. A. Giovannucci, J. Laláková, V. Menéndez-Benito, and N. P. Dantuma, 2017 The deubiquitylating enzyme Ubp12 regulates Rad23-dependent proteasomal degradation. J. Cell Sci. 130: 3336–3346.

Heessen S., M. G. Masucci, and N. P. Dantuma, 2005 The UBA2 domain functions as an intrinsic stabilization signal that protects Rad23 from proteasomal degradation. Mol. Cell 18: 225–235.

Her J., and S. F. Bunting, 2018 How cells ensure correct repair of DNA double-strand breaks. J. Biol. Chem. 293: 10502–10511.

Hershko A., and A. Ciechanover, 1998 The ubiquitin system. Annu. Rev. Biochem. 67: 425–479.

Hicke L., and R. Dunn, 2003 Regulation of membrane protein transport by ubiquitin and ubiquitin-binding proteins. Annu. Rev. Cell Dev. Biol. 19: 141–172.

Janićijević A., K. Sugasawa, Y. Shimizu, F. Hanaoka, N. Wijgers, *et al.*, 2003 DNA bending by the human damage recognition complex XPC–HR23B. DNA Repair 2: 325–336.

Jaspers N. G. J., 2001 *Encyclopedia of Genetics*. Elsevier Science Inc.

Kemp M. G., J. T. Reardon, L. A. Lindsey-Boltz, and A. Sancar, 2012 Mechanism of release and fate of excised oligonucleotides during nucleotide excision repair. J. Biol. Chem. 287: 22889–22899.

Kisselev A. F., T. N. Akopian, K. M. Woo, and A. L. Goldberg, 1999 The sizes of peptides generated from protein by mammalian 26 and 20 S proteasomes Implications for understanding the degradative mechanism and antigen presentation. J. Biol. Chem. 274: 3363–3371.

Komander D., and M. Rape, 2012 The ubiquitin code. Annu. Rev. Biochem. 81: 203–229.

Kraemer K. H., 1997 Sunlight and skin cancer: another link revealed. Proc. Natl. Acad. Sci. U.S.A. 94: 11–14.

Krebs J. E., E. S. Goldstein, and S. T. Kilpatrick, 2017 *Lewin's GENES XII*. Jones & Bartlett Learning.

Laczmanska I., J. Gil, P. Karpinski, A. Stembalska, A. Trusewicz, *et al.*, 2007 Polymorphism in nucleotide excision repair gene XPC correlates with bleomycin-induced chromosomal aberrations. Environ. Mol. Mutagen. 48: 666–671.

Lecker S. H., A. L. Goldberg, and W. E. Mitch, 2006 Protein Degradation by the Ubiquitin–Proteasome Pathway in Normal and Disease States. J. Am. Soc. Nephrol. 17: 1807–1819.

Li J., A. Bhat, and W. Xiao, 2011 Regulation of nucleotide excision repair through ubiquitination. Acta Biochim. Biophys. Sin. 43: 919–929.

Li L., S. J. Elledge, C. A. Peterson, E. S. Bales, and R. J. Legerski, 1994 Specific association between the human DNA repair proteins XPA and ERCC1. Proc. Natl. Acad. Sci. U.S.A. 91: 5012–5016.

Li L., C. A. Peterson, X. Lu, and R. J. Legerski, 1995 Mutations in XPA that prevent association with ERCC1 are defective in nucleotide excision repair. Mol. Cell. Biol. 15: 1993–1998.

Madura K., 2004 Rad23 and Rpn10: perennial wallflowers join the melee. Trends Biochem. Sci. 29: 637–640.

Maillard O., S. Solyom, and H. Naegeli, 2007 An aromatic sensor with aversion to damaged strands confers versatility to DNA repair. PLoS Biol. 5: e79.

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

Nance M. A., and S. A. Berry, 1992 Cockayne syndrome: review of 140 cases. Am. J. Med. Genet. 42: 68–84.

Ng J. M. Y., 2003 A novel regulation mechanism of DNA repair by damage-induced and RAD23-dependent stabilization of xeroderma pigmentosum group C protein. Genes & Development 17: 1630–1645.

Nishi R., Y. Okuda, E. Watanabe, T. Mori, S. Iwai, *et al.*, 2005 Centrin 2 stimulates nucleotide excision repair by interacting with xeroderma pigmentosum group C protein. Mol. Cell. Biol. 25: 5664–5674.

Nowicka U., D. Zhang, O. Walker, D. Krutauz, C. A. Castañeda, *et al.*, 2015 DNA-Damage-Inducible 1 Protein (Ddi1) Contains an Uncharacteristic Ubiquitin-like Domain that Binds Ubiquitin. Structure 23: 542–557.

Ortolan T. G., P. Tongaonkar, D. Lambertson, L. Chen, C. Schauber, *et al.*, 2000 The DNA repair protein Rad23 is a negative regulator of multi-ubiquitin chain assembly. Nature Cell Biology 2: 601–608.

Ortolan T. G., 2004 Rad23 stabilizes Rad4 from degradation by the Ub/proteasome pathway. Nucleic Acids Research 32: 6490–6500.

Preston B. D., T. M. Albertson, and A. J. Herr, 2010 DNA replication fidelity and cancer. Seminars in Cancer Biology 20: 281–293.

Raasi S., and C. M. Pickart, 2003 Rad23 ubiquitin-associated domains (UBA) inhibit 26 S proteasome-catalyzed proteolysis by sequestering lysine 48-linked polyubiquitin chains. J. Biol. Chem. 278: 8951–8959.

Ramsey K. L., J. J. Smith, A. Dasgupta, N. Maqani, P. Grant, *et al.*, 2004 The NEF4 complex regulates Rad4 levels and utilizes Snf2/Swi2-related ATPase activity for nucleotide excision repair. Mol. Cell. Biol. 24: 6362–6378.

Rock K. L., C. Gramm, L. Rothstein, K. Clark, R. Stein, *et al.*, 1994 Inhibitors of the proteasome block the degradation of most cell proteins and the generation of peptides presented on MHC class I molecules. Cell 78: 761–771.

Ruehle M. D., E. Orias, and C. G. Pearson, 2016 Tetrahymena as a Unicellular Model Eukaryote: Genetic and Genomic Tools. Genetics 203: 649–665.

Schärer O. D., 2008 Hot topics in DNA repair: the molecular basis for different disease states caused by mutations in TFIIH and XPG. DNA Repair 7: 339–344.

Schärer O. D., 2013 Nucleotide Excision Repair in Eukaryotes. Cold Spring Harbor Perspectives in Biology 5: a012609–a012609.

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

Sherman M. Y., and A. L. Goldberg, 2001 Cellular Defenses against Unfolded Proteins. Neuron 29: 15–32.

Shrestha A., 2011 The Role of DNA Damage Dependent Rad23 Ubiquitination in Stabilization of Rad23. MSU Graduate Theses. 2837.

Sigismund S., S. Polo, and P. P. Di Fiore, 2004 Signaling through monoubiquitination. Curr. Top. Microbiol. Immunol. 286: 149–185.

Singh R. K., and L. Dagnino, 2016 E2F1 interactions with hHR23A inhibit its degradation and promote DNA repair. Oncotarget 7.

Smerdon M. J., and M. W. Lieberman, 1978 Nucleosome rearrangement in human chromatin during UV-induced DNA- reapir synthesis. Proceedings of the National Academy of Sciences 75: 4238–4241.

Smith D. M., G. Kafri, Y. Cheng, D. Ng, T. Walz, *et al.*, 2005 ATP binding to PAN or the 26S ATPases causes association with the 20S proteasome, gate opening, and translocation of unfolded proteins. Mol. Cell 20: 687–698.

Smith J.J., Ubiquitin and NER: Making a Complex More Complex. Unpublished.

Spivak G., 2015 Nucleotide excision repair in humans. DNA Repair 36: 13–18.

Sudhakar J. N., and K.-C. Chow, 2014 Human RAD23 homolog A is required for the nuclear translocation of apoptosis-inducing factor during induction of cell death. Biology of the Cell 106: 359–376.

Sugasawa K., C. Masutani, A. Uchida, T. Maekawa, P. J. van der Spek, *et al.*, 1996 HHR23B, a human Rad23 homolog, stimulates XPC protein in nucleotide excision repair *in vitro*. Molecular and Cellular Biology 16: 4852–4861.

Sugasawa K., T. Okamoto, Y. Shimizu, C. Masutani, S. Iwai, *et al.*, 2001 A multistep damage recognition mechanism for global genomic nucleotide excision repair. Genes Dev. 15: 507–521.

Sugasawa K., Y. Okuda, M. Saijo, R. Nishi, N. Matsuda, *et al.*, 2005 UV-Induced Ubiquitylation of XPC Protein Mediated by UV-DDB-Ubiquitin Ligase Complex. Cell 121: 387–400.

Theil A. F., J. Nonnekens, N. Wijgers, W. Vermeulen, and G. Giglia-Mari, 2011 Slowly progressing nucleotide excision repair in trichothiodystrophy group A patient fibroblasts. Mol. Cell. Biol. 31: 3630–3638.

Theil A. F., J. Nonnekens, B. Steurer, P.-O. Mari, J. de Wit, *et al.*, 2013 Disruption of TTDA results in complete nucleotide excision repair deficiency and embryonic lethality. PLoS Genet. 9: e1003431.

Vélez-Cruz R., A. S. Zadorin, F. Coin, and J.-M. Egly, 2013 Sirt1 suppresses RNA synthesis after UV irradiation in combined xeroderma pigmentosum group D/Cockayne syndrome (XP-D/CS) cells. Proc. Natl. Acad. Sci. U. S. A. 110: E212–20.

Verma R., R. Oania, J. Graumann, and R. J. Deshaies, 2004 Multiubiquitin Chain Receptors Define a Layer of Substrate Selectivity in the Ubiquitin-Proteasome System. Cell 118: 99–110.

Wade S. L., K. Poorey, S. Bekiranov, and D. T. Auble, 2009 The Snf1 kinase and proteasome-associated Rad23 regulate UV-responsive gene expression. The EMBO Journal 28: 2919–2931.

Wade S. L., and D. T. Auble, 2010 The Rad23 ubiquitin receptor, the proteasome and functional specificity in transcriptional control. Transcription 1: 22–26.

Walters K. J., M. F. Kleijnen, A. M. Goh, G. Wagner, and P. M. Howley, 2002 Structural Studies of the Interaction between Ubiquitin Family Proteins and Proteasome Subunit S5a†. Biochemistry 41: 1767–1777.

Wilson, A., 2018 Development of Endogenous Tagging Plasmids for Characterization of Protein Interactions, Localization, and Post-Translational Modifications of Tetrahymena Thermophila Rad23. MSU Graduate Theses. 3268.

Xie Z., 2004 Roles of Rad23 protein in yeast nucleotide excision repair. Nucleic Acids Research 32: 5981–5990.

Ye Y., and M. Rape, 2009 Building ubiquitin chains: E2 enzymes at work. Nat. Rev. Mol. Cell Biol. 10: 755–764.

Yeh J. I., A. S. Levine, S. Du, U. Chinte, H. Ghodke, *et al.*, 2012 Damaged DNA induced UV-damaged DNA-binding protein (UV-DDB) dimerization and its roles in chromatinized DNA repair. Proc. Natl. Acad. Sci. U. S. A. 109: E2737–46.

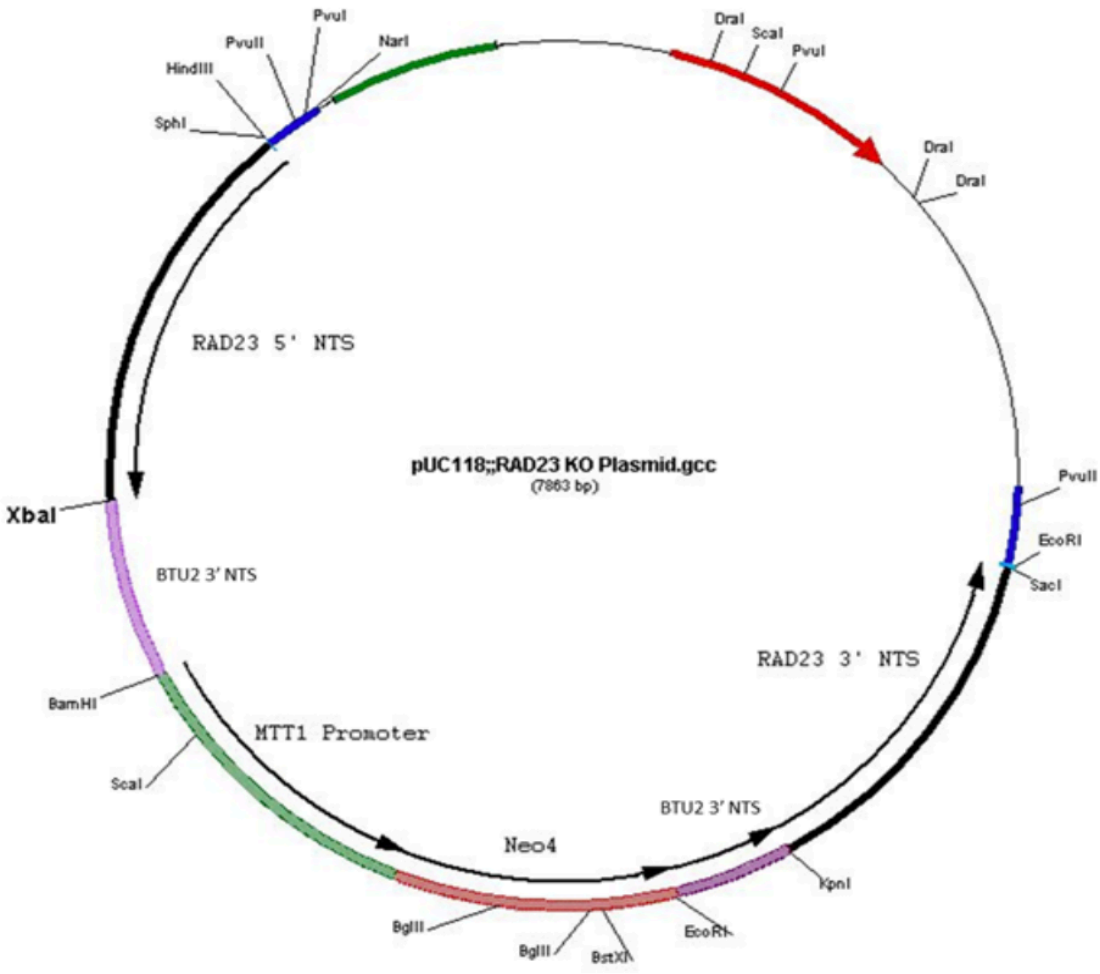
Zhou Z., N. Humphryes, P. van Eijk, R. Waters, S. Yu, *et al.*, 2015 UV induced ubiquitination of the yeast Rad4-Rad23 complex promotes survival by regulating cellular dNTP pools. Nucleic Acids Res. 43: 7360–7370.

Zientara-Rytter K., and S. Subramani, 2019 The Roles of Ubiquitin-Binding Protein Shuttles in the Degradative Fate of Ubiquitinated Proteins in the Ubiquitin-Proteasome System and Autophagy. Cells 8. https://doi.org/10.3390/cells8010040

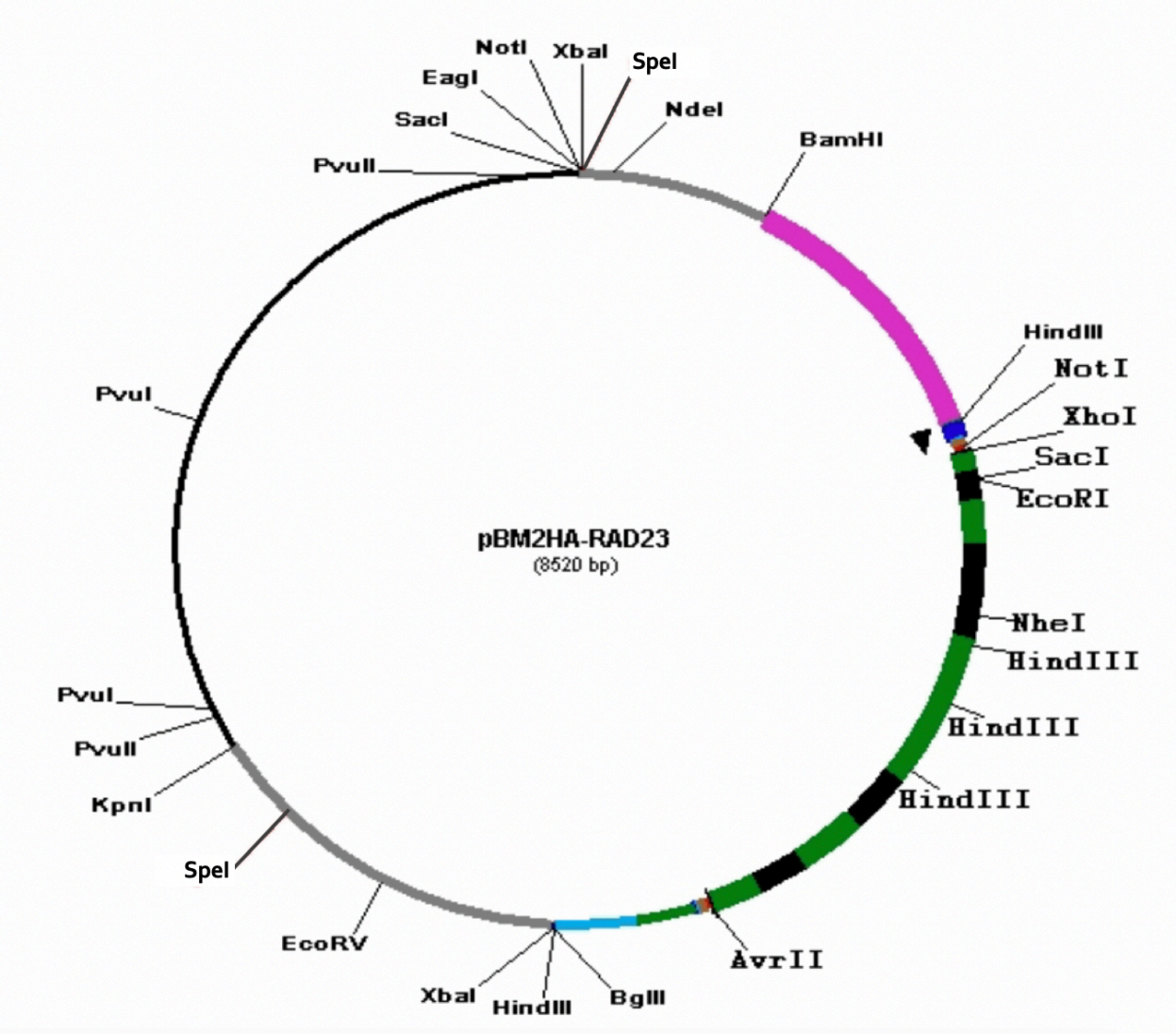
APPENDICES

Appendix A: Plasmid Maps

The two plasmids used for this study were designed and synthesized prior to the start of this project. Plasmid design was performed in silico using the Gene Construction Kit. Select restriction enzyme cut sites are included, which were used to generate predicted band sizes for RFLP analysis.



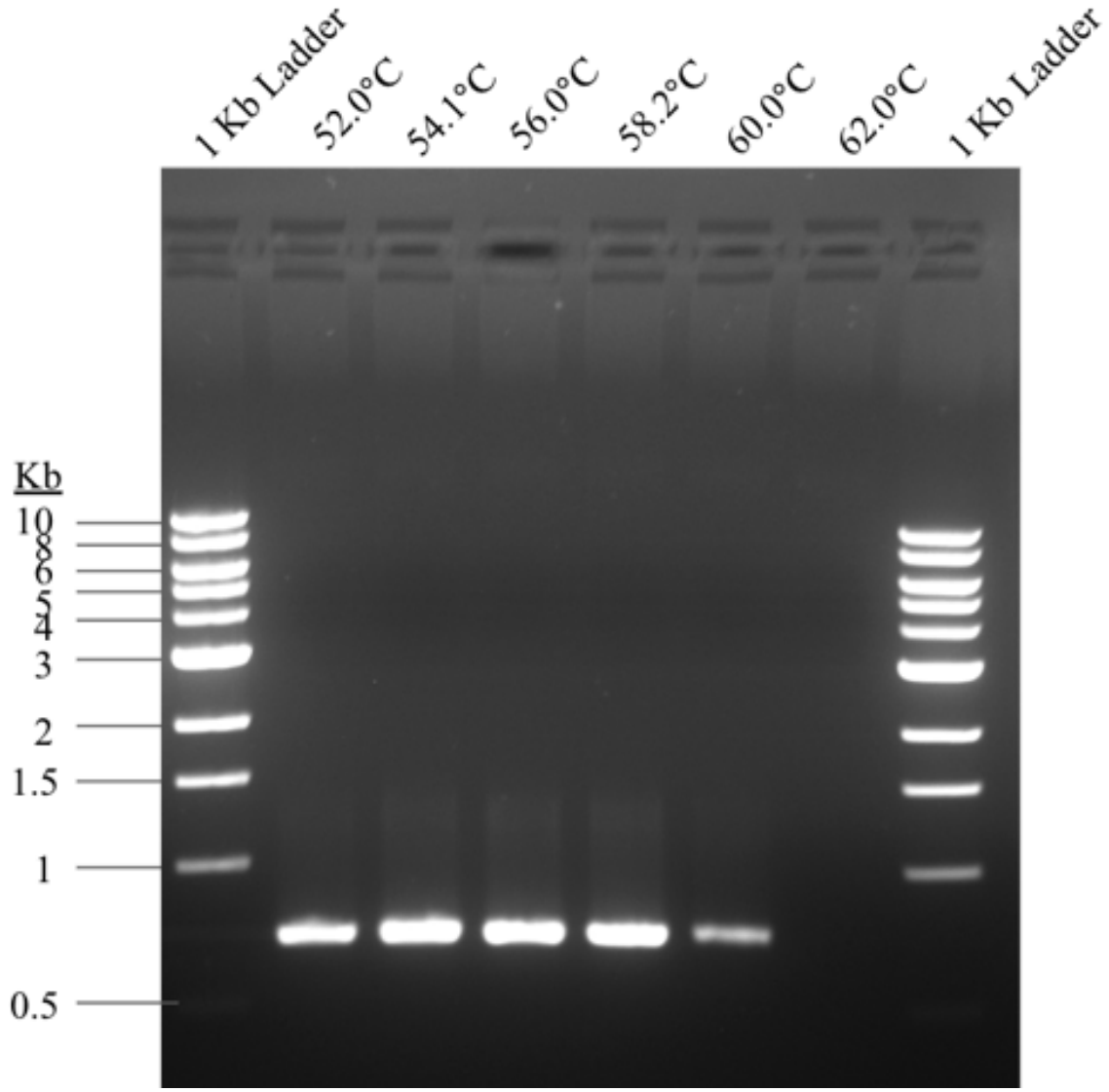
Appendix A1: Rad23 KO Plasmid Map. Plasmid map of pUC118b containing the MTTNEO4 cassette, generated using Gene Construction Kit. The pUC118b backbone consists of two multiple cloning sites (MCS), shown in light blue; the dark blue bars flanking the MTTNEO4 cassette represent the disrupted LacZ gene; Ampicillin resistance gene is represented by the red arrow; f1 replication origin is shown as the solid green bar; the MTT1 promoter (inducible by $CdCl_2$) is shown in light green; the *Tetrahymena thermophila* codon-optimized Paromomycin resistance gene is shown in light red; the BTU2 3' NTS (containing a poly-adenylation signal) is shown in purple; the *RAD23* 5' and 3' NTS are shown as black bars and will facilitate the insertion of the KO cassette into the endogenous *RAD23* loci of *Tetrahymena thermophila*. Black arrows indicate the direction of transcription. From Wilson, 2018.



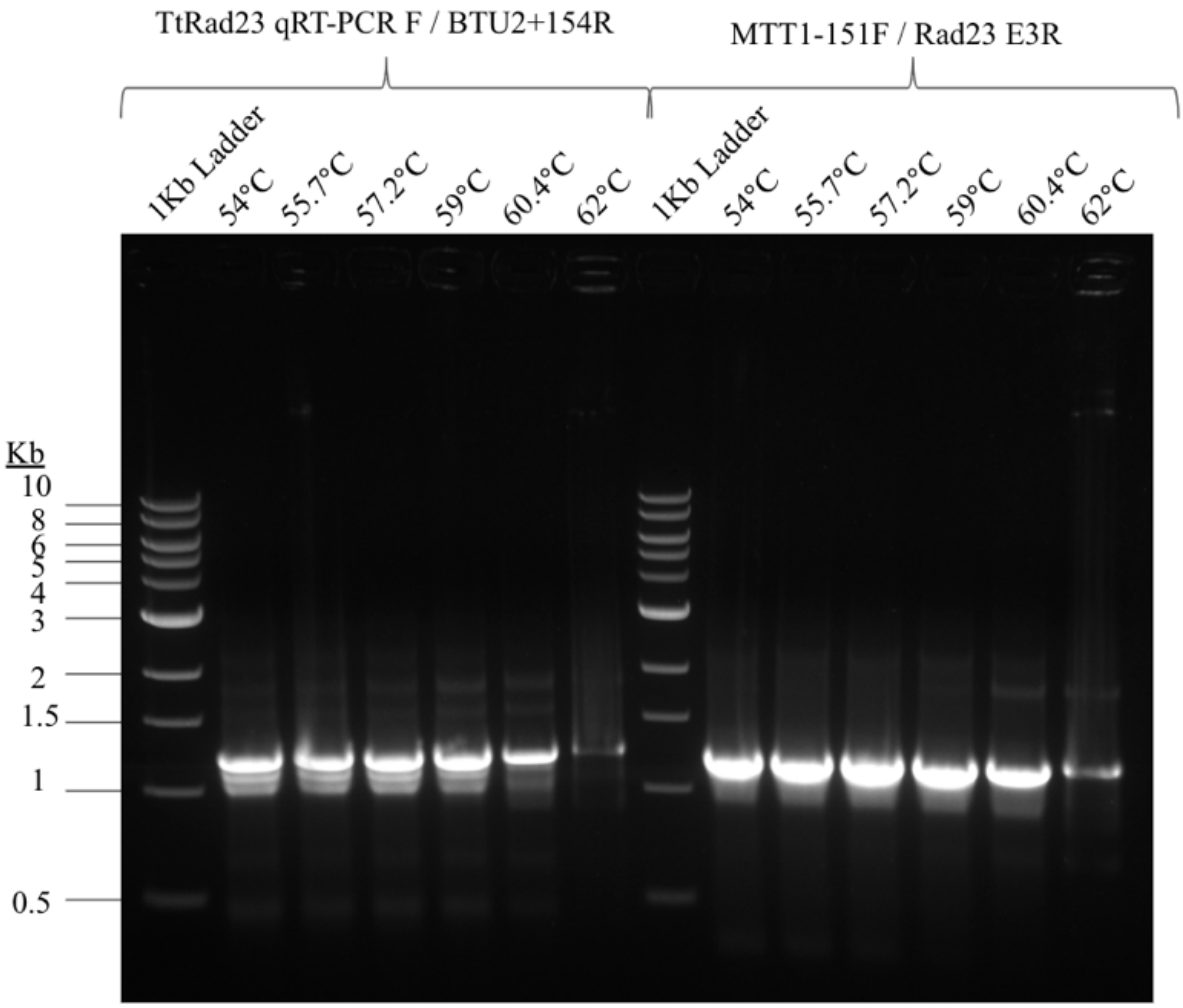
Appendix A2: 2HA-Rad23 Plasmid Map. Plasmid map of 2HA-Rad23 contained with a pUC118 gateway plasmid generated using Gene Construction Kit. The HA tag is shown in dark blue, between the magenta MTT1 promoter and Rad23 gene, shown in green (exons) and black (introns). Following the gene, a beta-tubulin poly-A addition sequence is shown in green and light blue; the junction of which represents the site of addition. Restriction enzyme cut sites appear where indicated. Adapted from Carpenter, 2012.

Appendix B: GoTaq Primer Optimization

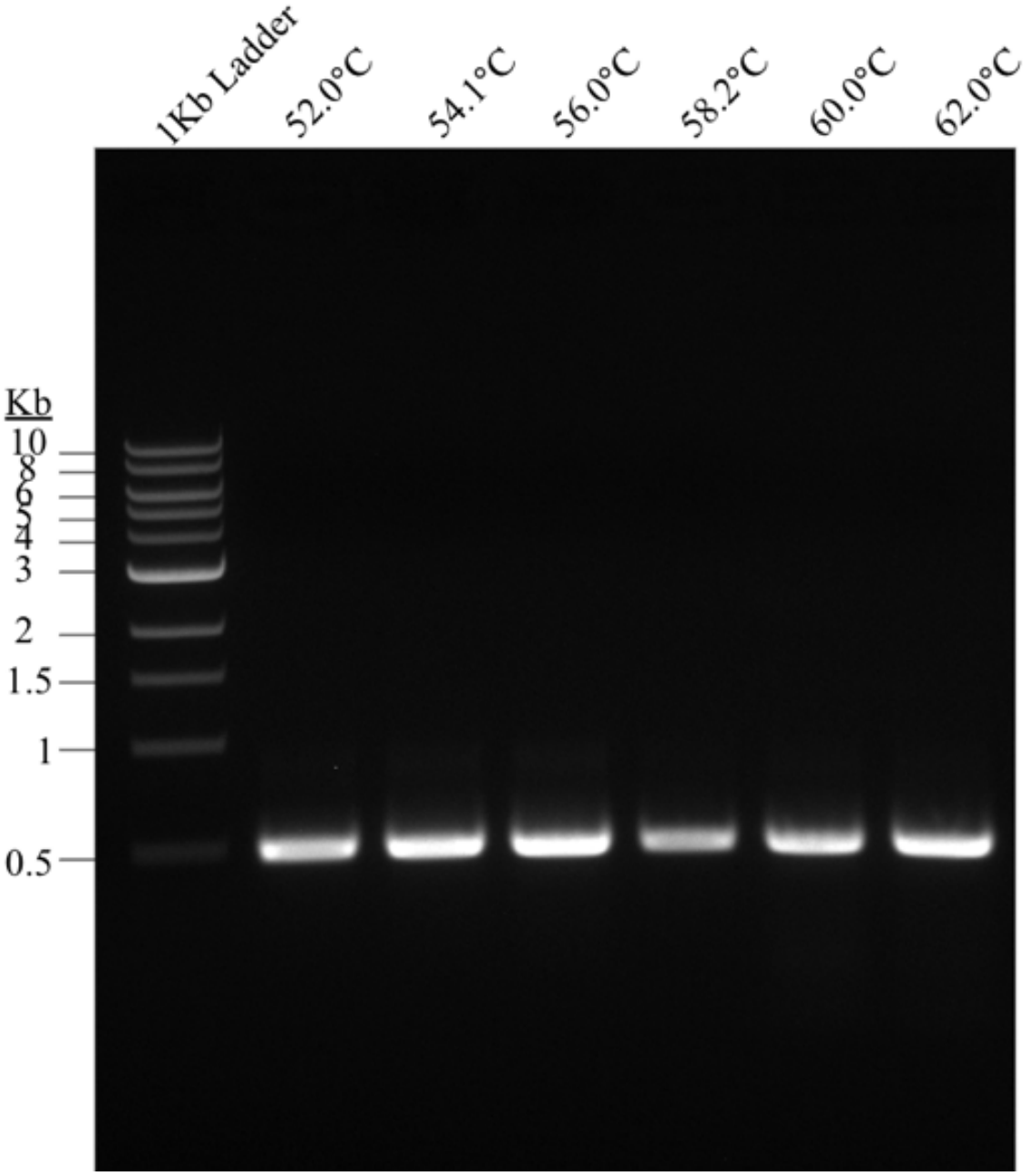
Efficient GoTaq PCR screening of transformant *Tetrahymena thermophila* colonies requires optimization of primer pairs. This process involves the amplification of pure target plasmid DNA across a temperature gradient of annealing temperatures. Products of the PCR reaction are run on a 1% agarose gel and visualized to determine the efficiency of the primer set and the optimal annealing temperature. Ideal pairs and temperatures will result in a clear, bright, single band of a specified product size.



Appendix B1: Optimization of Rad23 Knockout Plasmid Screening Primers. The primer set comprised of Rad23-132F and MTT-815R were used to amplify pure Rad23 knockout plasmid isolate at six different annealing temperatures. Products of the amplification were run on a 1.0% agarose gel for visualization. The annealing temperature 58.2°C was selected for subsequent use in colony and genomic DNA PCR screenings.



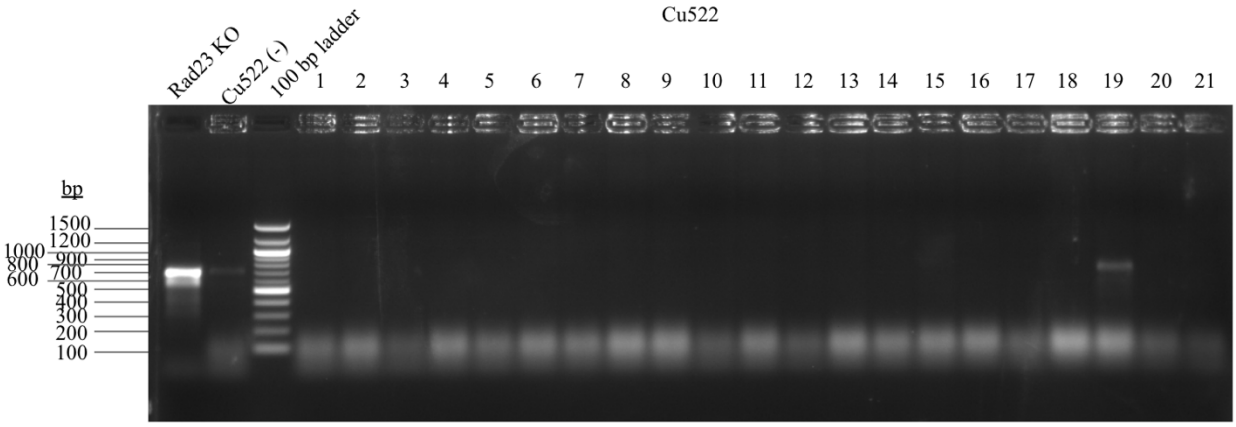
Appendix B2: Optimization of 2HA-Rad23 Plasmid Screening Primers. Two sets of screening primers were used to amplify pure 2HA-Rad23 plasmid isolate at six different annealing temperatures. Products of the amplification were run on a 1.0% agarose gel for visualization. The pair consisting of MTT1-151F and Rad23 E3R was chosen. The annealing temperature 57.2°C was selected for subsequent use in colony and genomic DNA PCR screenings.



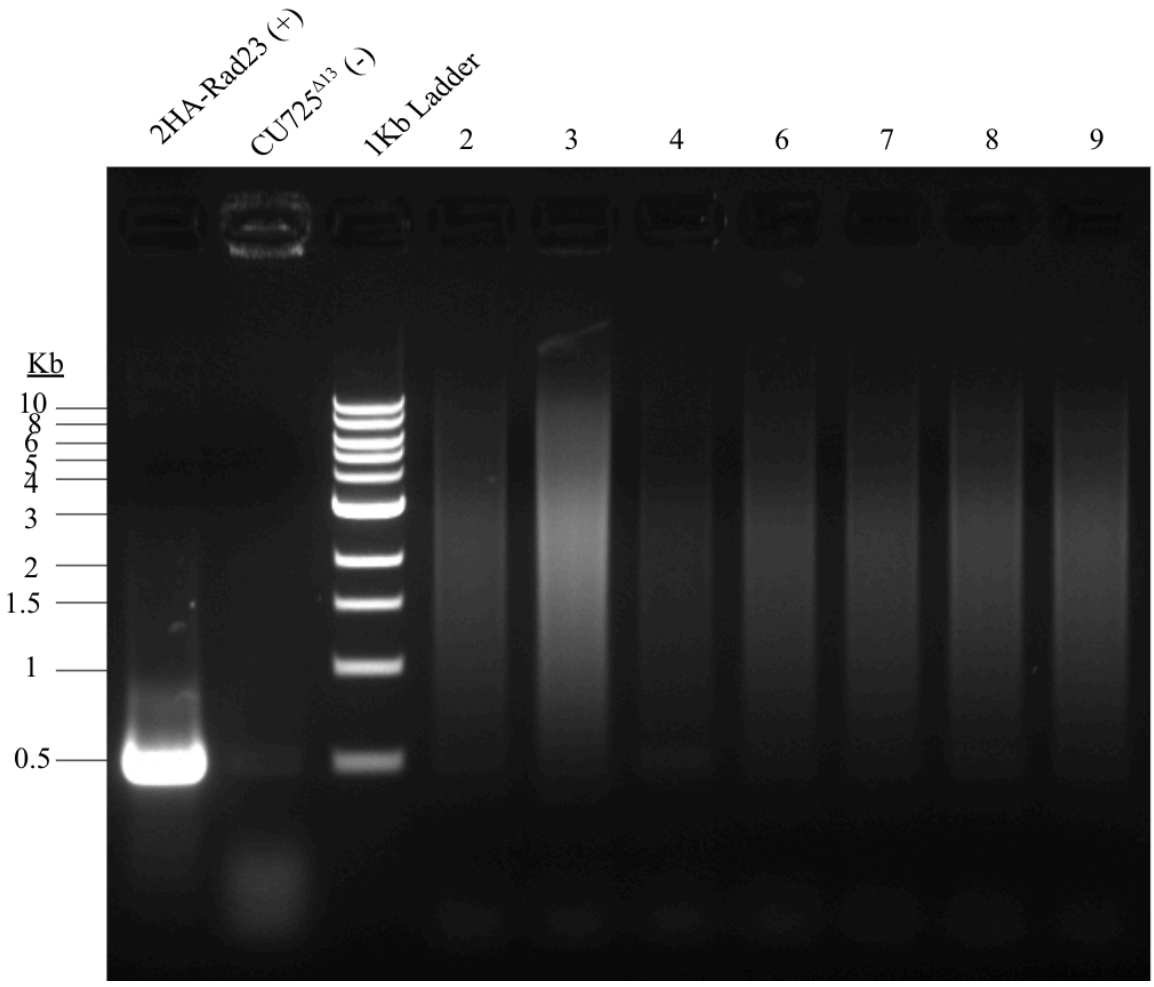
Appendix B3: Optimization of Additional 2HA-Rad23 Plasmid Screening Primers. Due to inconclusive screening results using the MTT1-151F and Rad23 E3R primer pair, a new reverse primer, Rad23 204R, was optimized for use opposite the MTT1-151F primer. Six annealing temperatures were used for amplification, and products were run on a 1.0% agarose gel for visualization. The annealing temperature 56.0°C was selected for subsequent use in colony and genomic DNA PCR screenings.

Appendix C: Additional Screening Data

Representative screening images for the *rad23*Δ and 2HA-Rad23 cell lines are included in the results section of this thesis. However, screenings were often done in multiple parts, in order to obtain conclusive confirmation of transformants. Typically, screenings continued until each desired cell line had between three and five clonal colonies confirmed by PCR. An additional screening (Appendix C1), as well as a representation of the inconclusive 2HA-Rad23 PCR in the CU725 *rad23*Δ #13 cell line (Appendix C2), are included here.



Appendix C1: Colony PCR Screening of CU522 *rad23*Δ Transformants. Figure 19B depicts the genomic DNA screen of five CU522 colonies transformed with the Rad23 knockout plasmid, four of which possess the plasmid. Prior to conducting the more sensitive gDNA screen, colony PCR was used in an attempt to identify colonies. This figure depicts a single confirmed colony, #19, which was identified through this screening method. This colony was used for subsequent gene expression analysis and survivability assays.



Appendix C2: Genomic DNA PCR Screening of 2HA-Rad23 in CU725 *rad23*Δ #13 Transformants. The CU725 *rad23*Δ #13 cell line, which was confirmed via qPCR to have significantly decreased Rad23 expression, was transformed using 2HA-Rad23 plasmid as an attempt to rescue Rad23 expression and wild-type cellular phenotypes. Use of the optimized primer pair MTT1-151F and Rad23 204R yielded non-specific "streaked" banding, concentrated at a molecular size larger than expected based on the positive control (~500 bp). This image is representative of multiple screening trials.

Appendix D: RNA Isolation and qPCR of *Rad23*

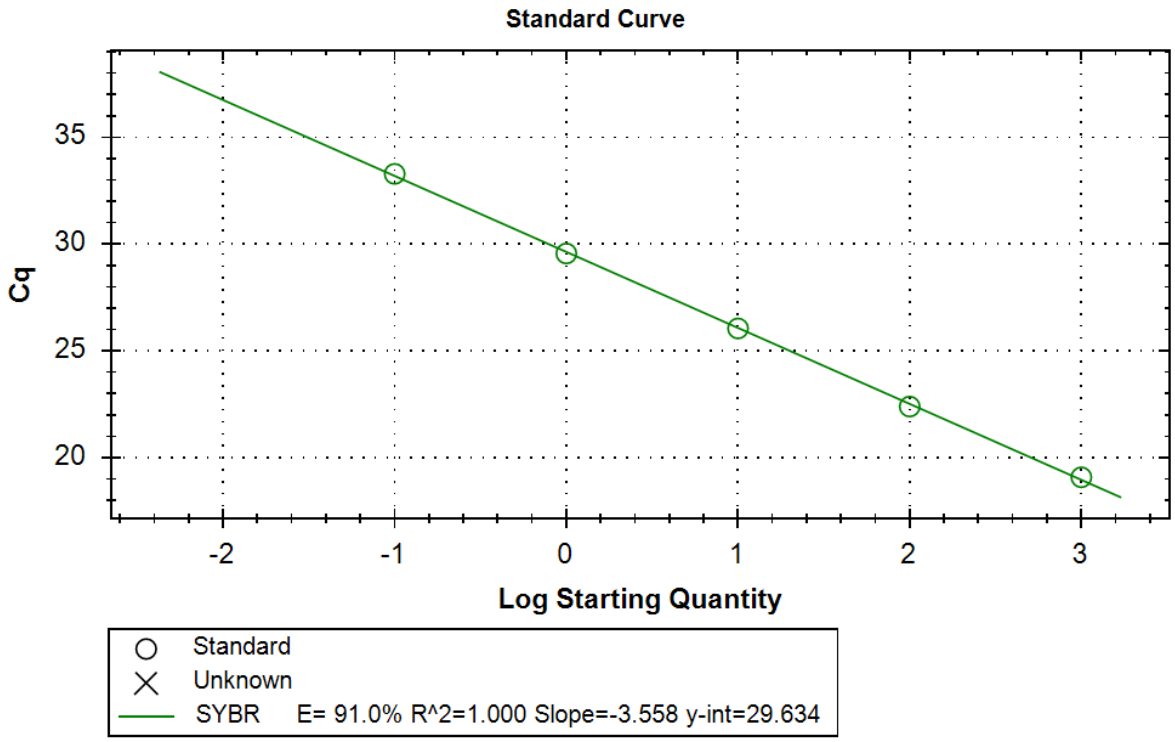
Total RNA was isolated from wild-type and Rad23-depleted cell lines prior to and following treatment with 100 J/m^2 UV. This RNA was used in an RT-PCR reaction to synthesize stable cDNA, which was in turn used for qPCR analysis of Rad23. Each qPCR experiment utilized a set of five genomic DNA standards and actin primers to generate a standard curve for starting quantity calculations. These actin standards were also analyzed via amplification curves and melt peaks during each run.

In addition to the standards, each cDNA sample was analyzed in the context of a housekeeping gene, HHP1, and Rad23. Amplification data and melt peaks were observed for both genes in each replicate. Following the first amplification attempt, PCR products were also run on an agarose gel to ensure only a single product was produced.

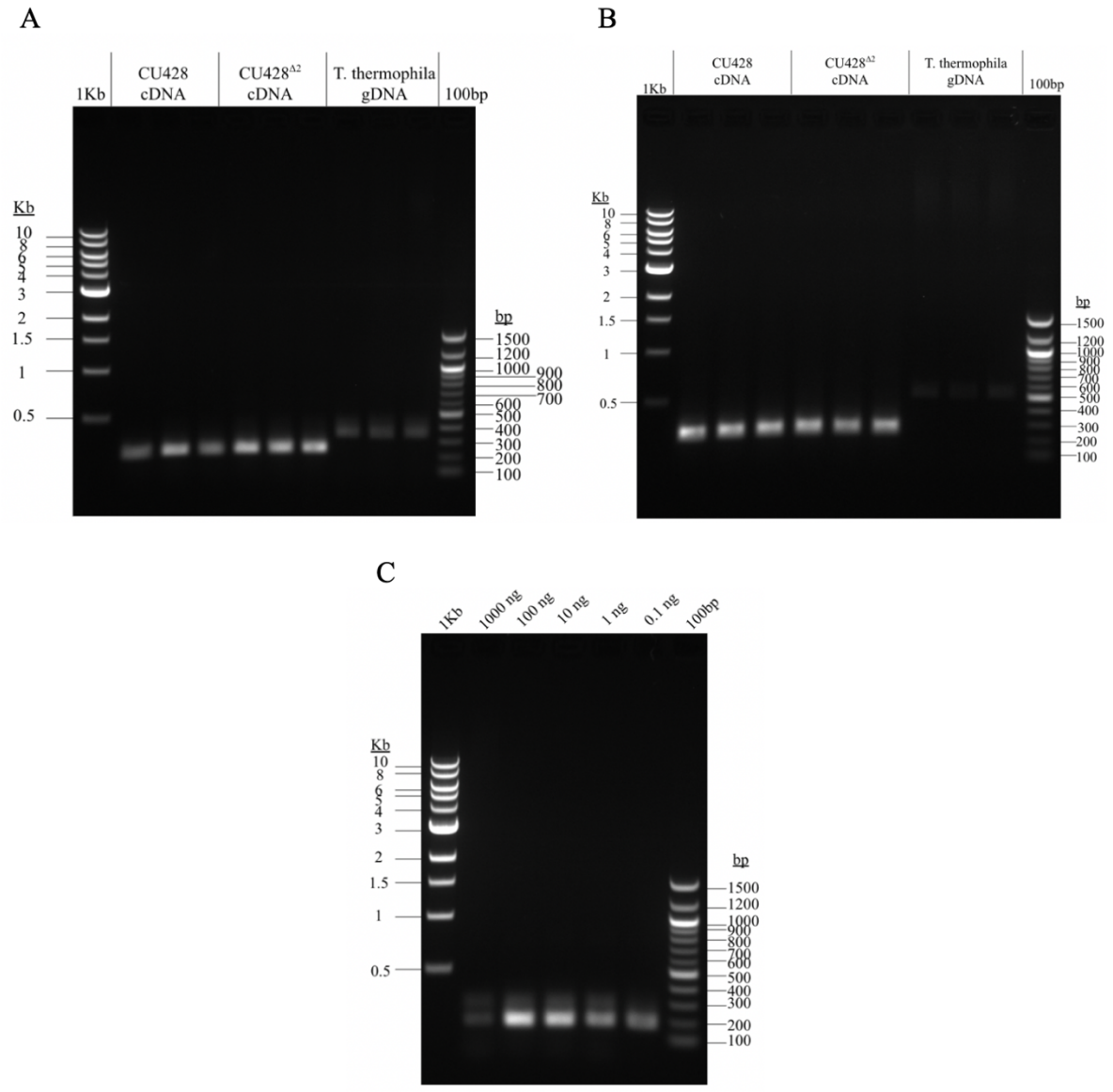
Appendix D1: RNA Isolate Concentration and Purity

Cell Line	Rad23[†]	UV (100J/m^2)	Concentration (µg/µL)	A_{260}/A_{280}
CU428	+	-	1.6618	2.427
	+	+	1.3803	2.501
	-	-	2.5499	2.471
	-	+	1.8455	2.486
CU522	+	-	1.4307	2.516
	+	+	0.9457	2.522
	-	-	1.3663	2.567
	-	+	0.9477	2.500
CU725	+	-	1.4397	2.601
	+	+	1.3298	2.541
	-	-	1.5663	2.464
	-	+	1.3530	2.474

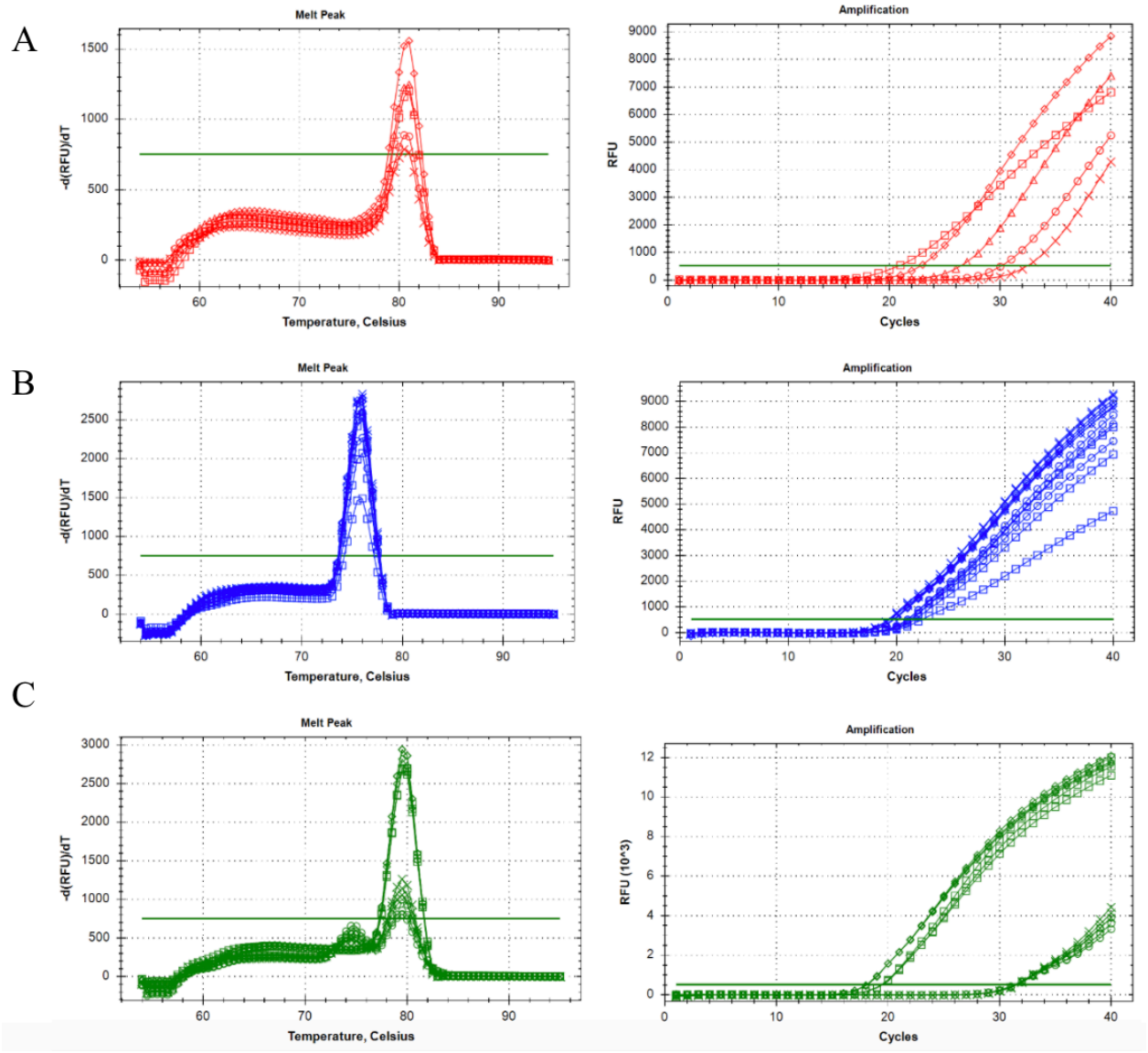
[†] Positive (+) indicates wild-type, negative (-) indicates *rad23*Δ mutant



Appendix D2: Representative Actin Standard Curve. Genomic DNA at concentrations of 1000 ng/µL, 100 ng/µL, 10 ng/µL, 1 ng/µL, and 0.1 ng/µL were amplified using actin (ACT1) primers. The resulting standard curve was used for starting quantity computation.



Appendix D3: qPCR Single-Product Confirmation. qPCR was run on *T. thermophila* genomic DNA and cDNA from wild-type and *rad23*Δ CU428 cells using HHP1 (**A**) and Rad23 (**B**) primer sets, to ensure the presence of only one PCR product. (**C**) Five dilutions of *T. thermophila* genomic DNA were amplified using actin primers. All products were run on a 1.0% agarose gel for visualization.



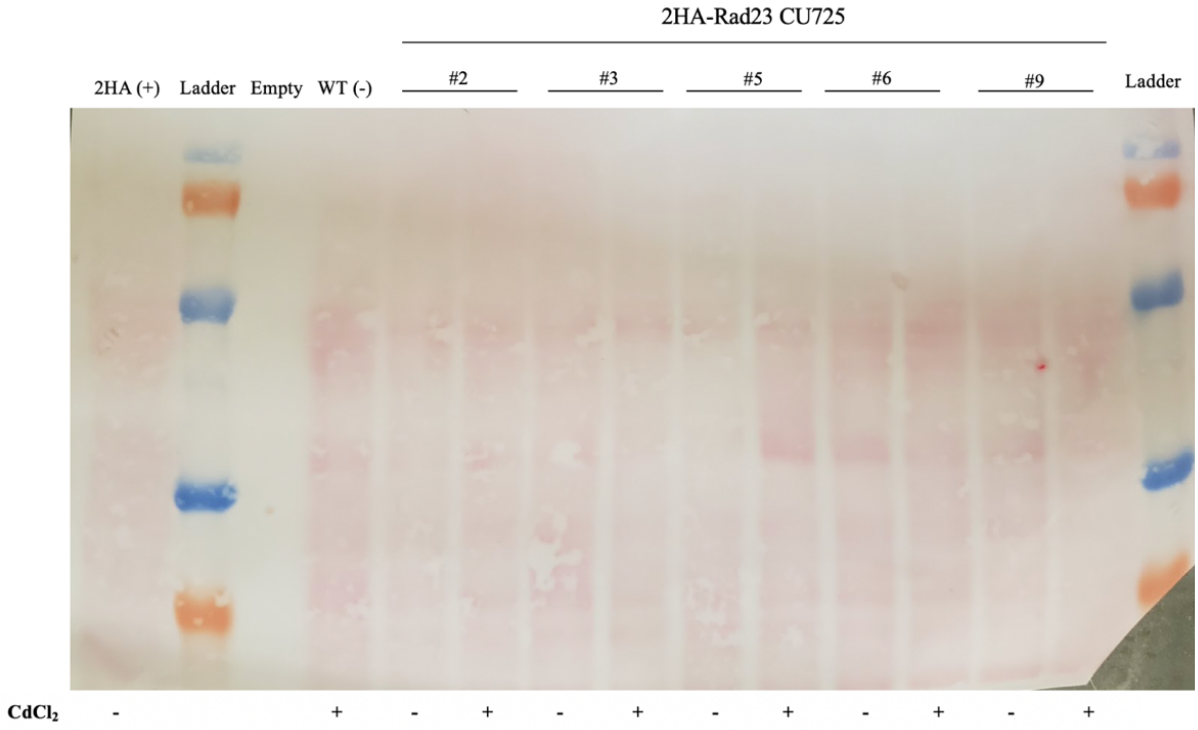
Appendix D4: Representative qPCR Traces of Background CU725. A: Melt peak (left) and amplification curve (right) of a *T. thermophila* genomic DNA standard using ACT1 primers (Squares: 1,000 ng/µL, Diamonds: 100 ng/µL, Triangles: 10 ng/µL, Circles: 1 ng/µL, Crosses: 0.1 ng/µL). **B-C**: Melt peak (left) and amplification curves (right) of *T. thermophila* cDNA isolated from wild-type and *rad23*Δ cell lines, before and after treatment with 100 J/m^2 UV. HHP1- (**B**) and Rad23-specific (**C**) primers were used. (Squares: untreated wild-type, Diamonds: treated wild-type; Circles: untreated *rad23*Δ, Crosses: treated *rad23*)

Appendix D5: Primer Melt Peaks

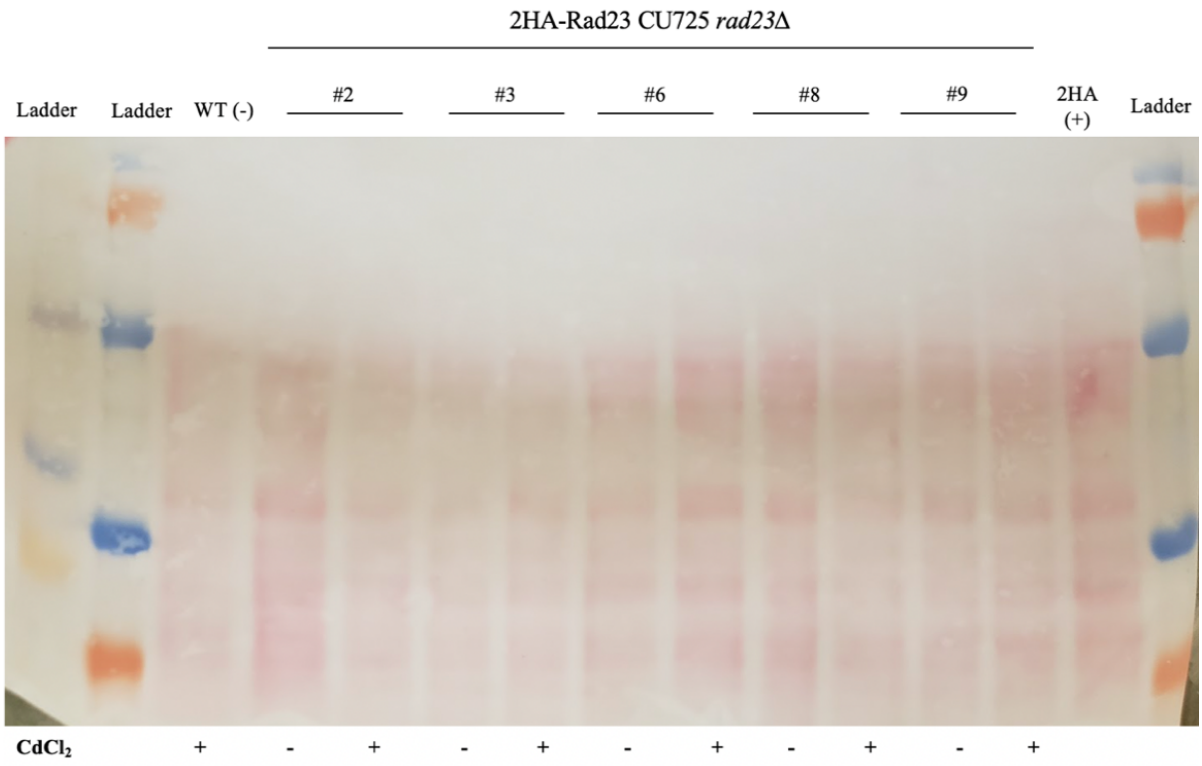
Target	Melt Peak (°C)
Act1	80.0-81.0
Hhp1	75.5-76.0
Rad23	79.5-80.0
Apn2	75.5
Rad6	77.0
Rad51	79.5-80.0
Tku80	75.0

Appendix E: Western Blot

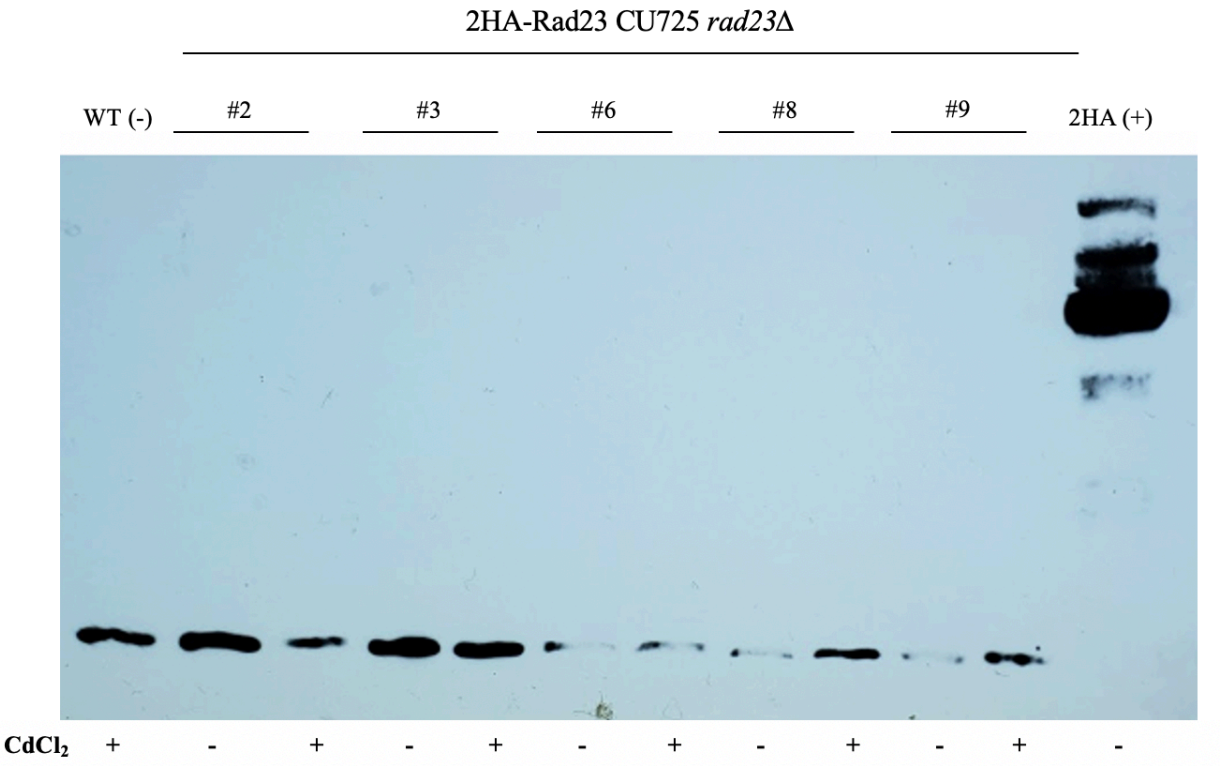
Ponceau staining was used as a loading control, to ensure that protein was adequately transferred from the gel to the membrane (Appendix E1 and E2). Western blotting for the 2HA tag in transformed *T. thermophila* cell lines was used to evaluate 2HA-Rad23 expression prior to and following induction with cadmium chloride. The CU725 *rad23*Δ #13 cell line (Appendix E3).



Appendix E1: Ponceau Stain of 2HA-Rad23 in CU725 Cell Lines. Protein isolated from 2HA-Rad23 transformed CU725 cell line was Ponceau stained following transfer from the gel to the nitrocellulose membrane. See Figure 26 for corresponding blot.



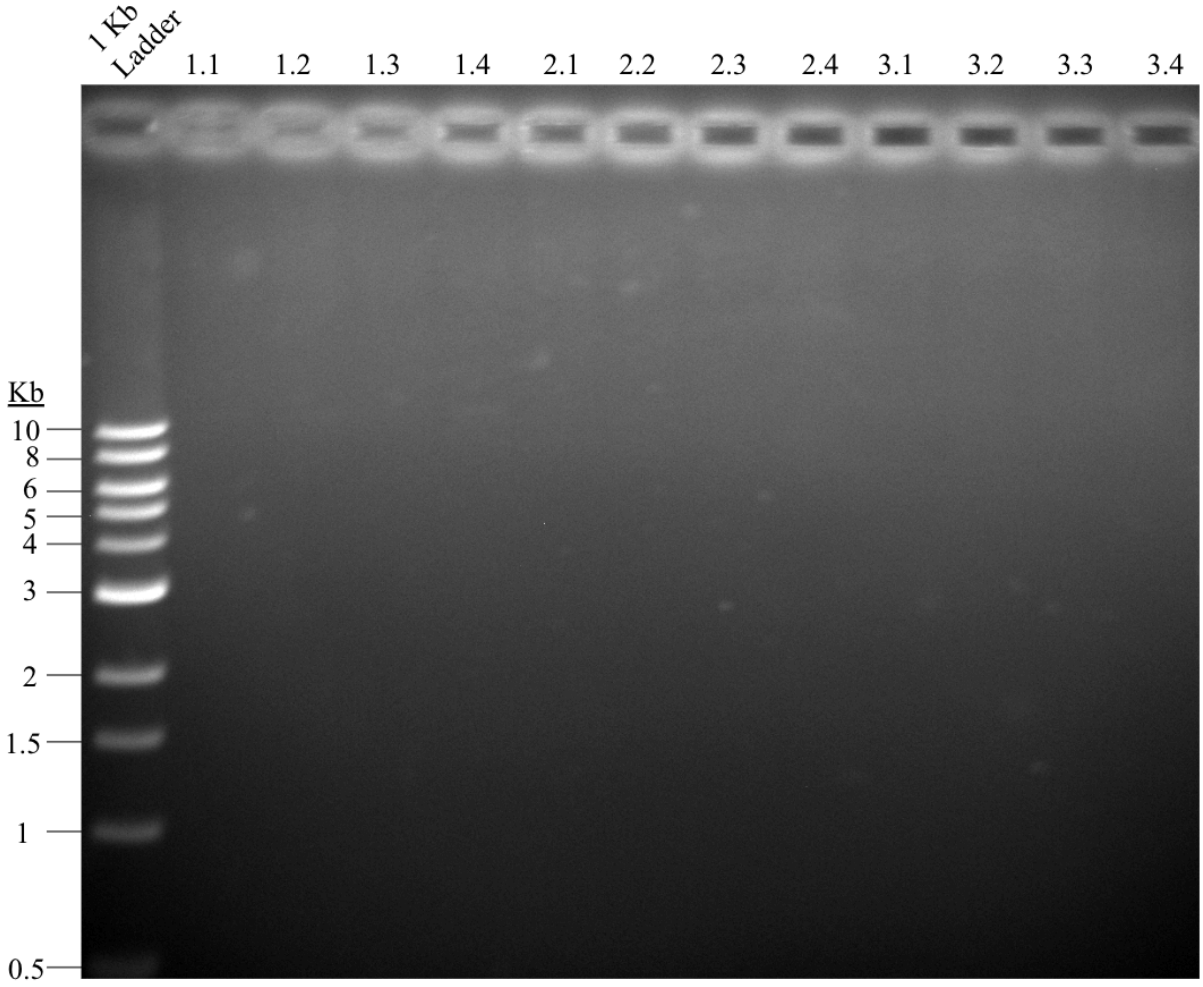
Appendix E2: Ponceau Stain of 2HA-Rad23 in CU725 *rad23*Δ #13 Cell Lines. Protein isolated from 2HA-Rad23 transformed CU725 *rad23*Δ #13 cell line was Ponceau stained following transfer from the gel to the nitrocellulose membrane.



Appendix E3: Western Blot of 2HA-Rad23 in CU725 *rad23*Δ #13 Cell Lines. A one-minute exposure was captured of the *rad23*Δ cell line transformed with 2HA-Rad23. The positive control (2HA) band reflects a positive banding pattern. While non-specific banding was observed in each of the protein isolates, including the untransformed CU725 *rad23*Δ #13 cell line (WT), no positive banding patterns appeared in the transformant lanes.

Appendix F: Site-Directed Mutagenesis

Site-directed mutagenesis was originally approached through Kunkel method, which requires isolation of single-stranded plasmid DNA for a template. Isolation was attempted multiple times on DH10B *E. coli* expressing the 2HA-Rad23 plasmid. When visualized on a 1.0% agarose gel, no bands were observed in any attempt (Appendix F1). Phusion mutagenesis was therefore pursued as an alternative method (Data not shown).



Appendix F1: Single-Stranded DNA Isolate. In an attempt to prepare for Kunkel site-directed mutagenesis, single stranded DNA isolation was performed on DH10B *E. coli* expressing 2HA-Rad23 plasmid. Three separate attempts were made (#1-3), with four colonies picked in each attempt (#.1-#.4). The isolate was run on a 1.0% agarose gel for visualization. No bands were observed.