Distribution Agreement

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

Signature:

_____ _____

Annie J. McPherson-Davie Date

Regulation of the base excision DNA repair pathway is critical for proper cell function

By
Annie J. McPherson-Davie
Doctor of Philosophy

Graduate Division of Biological and Biomedical Science
Genetics and Molecular Biology

_____
Anita H. Corbett, Ph.D.
Advisor

_____
Guy M. Benian, M.D.
Committee Member

_____
Paul W. Doetsch, Ph.D.
Committee Member

_____
David S. Yu, M.D., Ph.D.
Committee Member

_____
Wei Zhou, Ph.D.
Committee Member

Accepted:

_____
Lisa A. Tedesco, Ph.D.
Dean of the James T. Laney School of Graduate Studies

_____
Date

Regulation of the base excision DNA repair pathway is critical for proper cell function

By

Annie J. McPherson-Davie
B.S., North Carolina Agriculture & Technical State University, 2012

Advisor: Anita H. Corbett, Ph.D.

An abstract of
A dissertation submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
In partial fulfillment of the requirements for the degree of
Doctor of Philosophy
In Genetics and Molecular Biology
2020

Abstract

Regulation of the base excision DNA repair pathway is critical for proper cell function

By Annie J. McPherson-Davie

The accumulation of DNA damage within the genome can lead to numerous deleterious conditions including cancer. Thus, cells have evolved numerous, highly conserved DNA repair pathways to efficiently repair such damage and protect the genome. The base excision repair (BER) pathway repairs oxidative DNA damage. Although the biochemical steps in BER have been well defined, little is understood about how the pathway is regulated. The work described here exploits the Saccharomyces cerevisiae model to provide insight into how the BER pathway is regulated through analysis of a key BER protein termed NTHL1 in humans and Ntg1/2 in budding yeast. Previous work demonstrated that Ntg1 is sumoylated in response to oxidative damage. Here, we map the specific lysine residues that are sites of Ntg1 SUMO modification and then generate an Ntg1 variant where these five lysines are changed to arginine to create a variant of Ntg1 that cannot be modified by SUMO, ntg1ΔSUMO. When this Ntg1 variant is expressed in cells as the sole copy of Ntg1, cells show altered ability to arrest the cell cycle in response to DNA damage. This work begins to define how SUMO modification could regulate this key BER protein. To extend this work to mammalian cells, we demonstrated that human NTHL1 can also be modified by SUMO in response to oxidative insult, but the consequences of this modification have not yet been explored and the sites of modification have not been defined. This work was extended to create an S. cerevisiae system to explore the consequences of dysregulation of NTHL1, which has been linked to cancer. Overexpression of Ntg1 causes double-strand breaks and chromosome loss in budding yeast cells comparable to what occurs when NTHL1 is overexpressed in cultured cells. The budding yeast system facilitated genetic studies to define the pathways by which cells respond to the damage induced by overexpression of Ntg1 and provide insight into how different DNA repair pathways may intersect with one another. Taken together, this work provides important initial insights into potential molecular mechanisms that can regulate the BER pathways and coordinate cellular response to DNA damage.

Regulation of the base excision DNA repair pathway is critical for proper cell function

By

Annie J. McPherson-Davie
B.S., North Carolina Agriculture & Technical State University, 2012

Advisor: Anita H. Corbett, Ph.D.
A dissertation submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy
in Genetics and Molecular Biology
2020

Acknowledgments

This work is dedicated to my family, especially my mom (Umma) and dad (Appa). My parents have always encouraged me and my siblings to do well in school and achieve our dreams. They always prioritized the five of us and never asked anything for themselves. Through their selflessness, sacrifices, and encouragement, I was able to attend graduate school and successfully obtain a Ph.D. To my older brothers, Ho Young and Ho Lim Lee, thank you for always setting great examples in school and in life. My life would not have been the same without you guys, and I would not change our childhood for the world. To my younger siblings, Jonathan and Nikki McPherson, I hope I have in turn set a good example for you. I want you both to remember to follow your passions and live a happy life.

To my dear husband, Ronnie, thank you for loving me at both my high and low points and always caring for me. I cannot express how much I appreciate your sweet thoughtfulness throughout this process. You have also shared your lovely family with me. The Davies have taken me in and made me feel like their daughter. Thanks Davies for the love and laughs, and I look forward to our family's future together!

I would like to thank the past and present Corbett and Doetsch lab mates. Being co-mentored is not always easy, but you all have helped guide and mold me during my graduate experience. Thanks for always being willing to accept my questions and being there to bounce ideas off of. Through failed experiments, you guys were there to vent and commiserate with me. You guys were also there to celebrate the successes! The day to day of research was much more bearable with you all by my side. Additionally, I am thankful to have had the opportunity to mentor so many great students while at Emory. I wish you all success in your endeavors, academically and otherwise. Remember, once my mentee, always my mentee!

I thank my committee members, Anita, Paul, David, Guy, and Wei. It has been a long tough road, but you guys have always been there to push and steer me. Thank you for all the scientific, career, and life advice you all have given me over the years. Paul, you in particular are always very encouraging of me no matter what I do. Thank you for all of your support. Anita, you are basically superwoman, and I do not know how you do it. You have always prioritized your many students in the midst of wearing all of your other hats. You value mentoring and it shows. I wanted to come to Emory because of you, and you (and the training I knew you could provide) are what drove me to join your lab. I cannot thank you enough for the patience and understanding you have shown me. I really appreciate you and our friendship!

I would also like to thank the GMB program and the Emory University affiliates (students, professors, and staff). I have made many lifelong friends here, and I am grateful to have been a part of such a great University. I would especially like to thank my classmates, Liz, Marko, Kelly, and Kristie. I am so lucky to have joined the same year as these fabulous people. The first two years in particular were very difficult for me, but together we all made it through. I could not have asked for a better group to start graduate school with. Many members of the Emory staff have become like a second family to me, and I will always appreciate them for looking out for me. To the friends before and friends I made while in graduate school, thank you for being there for me.

Table of Contents

List of Figures

Chapter 1: General Introduction

1.1 Types of DNA damage, frequencies, and consequences

The DNA in cells is constantly in danger of damage from a variety of DNA damaging agents[1–3] (Figure 1.1). These DNA damaging agents come from both outside (exogenous) and within (endogenous) the cell and can impact cellular function[1,4] (Figure 1.2). Both nuclear and mitochondrial DNA are subject to these potential deleterious chemical processes[4–6] (Figure 1.2). DNA damaging agents can lead to a number of different DNA damages called lesions[1,7] (Figure 1.1). If not addressed quickly and efficiently, these lesions can lead to genetic or genomic instability[1,7]. Exogenous sources of DNA damage include ultraviolet (UV) light, ionizing radiation (IR), and chemical combustion[7,8] (Figure 1.1). Simply stepping outside can expose you to low levels of harmful UV light from the sun[9–11]. The photochemical reaction induced by UV light can cause two neighboring pyrimidines in DNA (cytosine and thymine) or RNA (cytosine, thymine, and uracil) to dimerize[9–11] (Figure 1.1). This dimerization lesion is estimated to occur up to 50-100 times per second during exposure to sunlight[7]. The most common dimers are cyclobutene pyrimide and pyrimidine-pyrimidone (6-4) photoproducts, and their signature mutations are prevalent in genes mutated in skin cancers[7,9–11]. Ultimately, if left unrepaired, these dimers can inhibit proper replication of DNA by disrupting polymerases and therefore preventing proper cellular function[7,10] (Figure 1.2).

IR is composed of a number of rays (i.e. cosmic and gamma rays) that can vary in energy and have wavelengths less than 100nm[12] (Figure 1.1). IR is abundant in the environment and arise from radioisotopes found in places like rocks, soil, building material in old basements, and from medical equipment[12]. IR can either directly or indirectly damage the DNA[8]. Direct IR can cause fragmented sugar derivatives, single-strand breaks (SSB), and double-strand breaks (DSB)[7,13,14]

(Figure 1.1). IR damage can ultimately lead to chromosome loss, genomic rearrangements, cell death, and tumorigenesis[7,14].

Indirect IR can occur when the surrounding water molecules are hit by IR and are turned into a cluster of highly reactive free radicals (i.e. hydroxyl radical (•OH))[7,14] (Figure 1.1). These free radicals then damage the DNA forming lesions and DSBs[7,14]. 65% of the IR DNA damage is caused by •OH activity and resembles damages induced by reactive oxygen species (ROS) including thymine glycol to be discussed later[14,15] (Figure 1.1).

Endogenous sources of DNA damaging agents include spontaneous base hydrolysis, ROS, and enzymatic activity[7,16–18] (Figure 1.1). Chemical modifications are the most common damage that occur in the DNA [3,7,19]. There are an estimated 20,000-120,000 base lesions that occur per human cell per day[7,20]. This range is likely an underestimate as we do not currently have tools sensitive enough to accurately detect all the base lesions that occur in the cell[21]. These DNA modifications can occur in a variety of ways, but there are four major base lesion classes (hydrolysis, oxidation, deamination, and alkylation)[7,18,22,23] (Figure 1.1). These lesions, if left alone, can ultimately lead to mutations and genomic instability[24] (Figure 1.2). Spontaneous base hydrolysis is estimated to occur 2,000-10,000 times per human cell per day and occurs primarily at guanine bases[3,7,25]. These damages, known as abasic sites, or apurinic/apyrimidinic (AP) sites, can be particularly mutagenic as they can inhibit transcription[3,26]. Oxidation is caused by ROS (hydrogen peroxide, hydroxyl radical, and super oxide anion)[15,27,28]. Of the four bases, guanines are particularly sensitive to oxidation, leading to 8-oxo-7,8-dihydroguanine (8-oxoG), and these lesions are estimated to occur 100-500 times per human cell per day[3,7].

Due to the proximity of the mitochondria DNA (mtDNA) to the electron transport chain (ETC), which produces energy for the cell and ROS as a byproduct, and the method of replicating

the mitochondria genome, mtDNA is mutated more frequently than nuclear DNA[28–30]. Per year, the point mutation rate in mtDNA is about $6 \bullet 10^{-8}$ per base pair (bp) and in general the mutation rate increases about 5-fold from the age of one to 80[31]. Deletions and point mutations that occur in mtDNA can lead to serious consequences like impaired heart, muscle, and nervous system tissues, Mitochondrial Encephalomyopathy, Lactic Acidosis, Stroke-like Episodes (MELAS) and have been associated with Parkinson's and Alzheimer's diseases[32,33].

Enzymatic induced damages include polymerase base incorporation errors and the intermediates formed during DNA repair[7,34] (Figure 1.1). Although the replication polymerases are highly accurate (making mistakes at a frequency of 10^{-3} to $>10^{-6}$) and have the ability to proofread, they still make mistakes especially in highly redundant regions like the centromeres[34]. Common polymerase mistakes include base-base mispairing, insertions and deletions[34] (Figure 1.1).

In the process of repairing the DNA, 15,000-30,000 enzymatically induced damages occur per human cell per day[7]. Repair intermediates include AP sites, SSBs, DSBs and sometimes loss of genetic information (Figure 1.1). SSBs occur at around 10,000 times per human cell per day[7]. DSBs, although less common, are the most lethal form of DNA damage to cells and are estimated to occur around 10-50 times per human cell, per day, however this estimate can vary based on the cell type and cell cycle stage[7] (Figure 1.1). While SSBs do not normally compromise the integrity of dsDNA, if left unrepaired, a SSB could be converted to a DSB during replication[19]. Additionally, a DSB can occur when a nuclease attacks damaged DNA at a replication fork[35]. Sometimes, replication fork reversal can also result in a DSB[11,36,37]. Unrepaired, or incorrectly repaired DSBs, can result in mutations, chromosomal rearrangements, or even cell death[1,7,22] (Figure 1.2).

DNA damage can ultimately have pathological consequences, such as aging, degenerative disorders, and cancer development[11,27,28,38] (Figure 1.2). Paradoxically, cancer treatment often involves inducing DNA damage with chemotherapy and radiotherapy[39–42]. In the context of cancer therapy, treatments cause a large number of DNA damages that force the cell into apoptosis[39–42]. High levels of DNA damage are particularly harmful to quickly dividing cells, such as cancer cells[39–42]. Unfortunately, normal fast dividing cells such as gut epithelium, bone marrow hematopoietic cells and hair follicle cells can also be affected.

Altogether, there are many types of DNA damages that occur every day caused by a variety of sources[7]. Although some changes in DNA are important for genetic diversity and natural evolution, the majority of DNA damages can have lasting negative effects on cell function[1,7]. These damages, once incorporated into mutations, can lead to cell death, aberrant cell growth, or cancer development[1,11,26]. The importance of maintaining genetic and genomic stability in spite of the large number and variety of DNA damages highlights the need for DNA repair pathways (Figure 1.2).

1.2 DNA repair pathways

DNA damage can threaten the normal functions of the cell and lead to pathological consequences[1,7] (Figure 1.2). To combat the large variety of DNA damages, cells employ a number of DNA repair pathways. Six major DNA repair pathways exist to repair, mediate, or tolerate DNA damage[4,12] (Figure 1.3). These repair pathways are highly evolutionarily conserved, and each pathway is specifically suited to repair certain types of DNA damage. Much is known about the steps involved in each of these repair pathways. All six pathways have been identified as present and active in the nucleus[4] (Figure 1.3). In the mitochondria, some components from every pathway

are present, however, in some cases the repair pathway does not seem functional in this organelle[4] (Figure 1.3). Repair pathways are regulated both with and independent of the cell cycle and range in accuracy and efficiency. Each repair pathway is unique and defective or missing repair components often result in tumorigenesis.

i. The base excision repair pathway

The base excision repair (BER) pathway is a versatile repair pathway that targets a large number of small non-helix distorting base damages[43–46] (Figure 1.3). Major targets of BER are oxidized, alkylated, and deaminated bases[3,5,47,48]. BER can take place in two forms, long or short patch (to be discussed later). Some components of BER are coordinated with the cell cycle, but generally BER is active in G_1 phase[49]. In general, the steps of BER are: 1) Excision; 2) incision; 3) end processing; 4) repair synthesis; and 5) ligation (to be discussed later). BER is fully functional in the mitochondria and the most well understood mtDNA repair pathway. Missing or mutant components of BER lead to colorectal cancer, general cancer predisposition, and immunological defects[11,50,51].

ii. The mismatch repair pathway

The mismatch repair (MMR) pathway targets replication errors that were missed by the proof-reading machinery of the polymerase[1,4,52–54] (Figure 1.3). MMR excises a region of DNA including the mismatch and effectively gives a polymerase a second chance to synthesize the DNA[52–54]. This type of repair is dependent on the ability to distinguish between the newly formed DNA and the template strand as the template is assumed to be accurate[52–54]. The substrates of MMR are insertions, deletions and misincorporated bases[52–54]. MMR is active during replication and recombination and therefore primarily active during S phase[52–54]. This repair pathway is less active, but still functional during other phases of the cell cycle[52–55]. The steps of MMR are: 1)

damage recognition; 2) template identification; 3) base excision; 4) error free gap filling; and 5) ligation.

Only a few enzymes involved in MMR have been identified in the mitochondria so far, and the ones that have seem to function differently in this compartment[4,55]. As a protein involved in both nuclear BER and NER, the protein, YB-1, is confirmed to be involved in mitochondria MMR, but no connection to mitochondrial BER or NER has yet been found[55] (Figure 1.3). Technically, mismatches in mtDNA can be repaired by BER. Further research is needed to determine if MMR is fully functional in mitochondria[4,55].

In humans, how the template is distinguished from nascent DNA is not known, however, the leading theory is that nicks in the backbone of newly synthesized DNA are used to distinguish the newly synthesized strand from the template strand[52]. Interestingly, MMR can remove from a few to thousands of nucleotides[52–54]. Mutations in this repair pathway lead to microsatellite instability, hereditary nonpolyposis colorectal cancer (HNPCC) or Lynch syndrome, and constitutional mismatch repair deficiency syndrome (CMMR-D) which can lead to sporadic cancers[11,52,56].

iii. The nucleotide excision repair pathway

Nucleotide excision repair (NER) is the primary pathway for repairing helix distorting DNA damage, with the most common damage repaired being pyrimidine dimers that result from UV irradiation[16,57,58] (Figure 1.3). NER is active throughout the cell cycle[59]. The general steps of NER are: 1) damage recognition; 2) DNA unwinding; 3) excision; 4) gap filling; and 5) ligation. NER can be divided into two subcategories, transcription-coupled repair (TCR) and global genome repair (GGR)[57] (Figure 1.3). TCR occurs on transcriptionally active genes and detects when an RNA polymerase stalls at a lesion on the DNA[60]. The RNA polymerase is removed, and the

transcript is terminated so NER can repair the lesion[60]. GGR occurs throughout the genome, including non-transcribed strands of DNA[60].

Components of the NER pathway were only recently discovered in mitochondria[55]. This could be because these enzymes are endogenously maintained at low levels in the mitochondria and the optimal conditions to increase the abundance of these proteins have not yet been defined. Further research is needed to determine if NER is fully functional in mitochondria.

Xeroderma pigmentosum (XP) is a skin disorder that develops in people missing components of NER[11]. People who suffer from XP cannot be exposed to the sun because they can develop skin lesions that often become cancerous[11]. XP is characterized by about a 10,000-fold increased risk of skin cancer associated with sunlight exposure[11]. Cockayne syndrome (CS) is another disorder that can be traced back to mutations in the NER pathway[11]. CS is a rare and fatal disorder characterized by impaired development, sensitivity to sunlight, and premature aging[11].

iv. The direct reversal repair pathway

The direct reversal repair pathway removes the chemical addition rather than excising the base[53]. (Figure 1.3). This process targets alkylated groups and is error-free[53]. Proteins involved in direct damage reversal locate to replication foci during S phase[53,59]. Once damage is detected, direct reversal repair proteins catalyze the removal of the alkyl group[53]. A protein variant of a nuclear direct reversal repair protein has been identified in the mitochondria[4]. Mutants in direct reversal repair are associated with non-small-cell lung, colorectal, and prostate cancers[11].

v. The double-strand break repair pathway

The double-strand break repair (DSBR) pathway repairs or mediates the most lethal type of DNA damage, DSBs[4,61–63] (Figure 1.3). DSBR is further broken down into two major subcategories, homologous recombination (HR), and end joining (EJ)[61–63]. The key difference

between HR and EJ is the initial processing of the DSB[61–63]. Loss of components of DSBR can lead to loss of heterozygosity and tumorigenesis[11].

HR is the more accurate of the two DSBR subcategories[61,62,64] (Figure 1.3). This process is slow, and only occurs when a sister chromatid is available and therefore, only during S phase and G2[61,62,65]. Initially, long stretches DNA are resected, and a homology search ensues to identify complementary DNA sequences and limit the loss of genetic information[61,62,64]. In general, the steps of HR are: 1) resection; 2) strand invasion; 3) DNA synthesis; and 4) resolution. While HR is often accurate, with enough homology, two non-homologous loci could be joined, leading to genome rearrangements[24].

Outside of repair, HR occurs during meiosis and T and B cell maturation to increase genetic and immune diversity[61,62,64]. Although DSBR is known to occur in the mitochondria, most components of HR have not yet been identified in this compartment[4]. HR deficiency leads to breast, ovarian, skin, and bone cancers, and Fanconi anemia, a cancer predisposition syndrome[11].

During the cell cycle when a sister chromatid is not accessible, a faster, less accurate repair pathway is used for DSBs[63,66]. This repair pathway is called end joining (EJ)[63,66] (Figure 1.3). EJ ligates two DSB ends together whether they originally resided next to one another or not[63,66]. Initially, very little end processing is done before end joining occurs. The steps of EJ are: 1) recognition; 2) end processing; and 3) ligation. EJ generally leads to limited loss of genetic information but can result in large deletions and chromosomal rearrangements and therefore could be potentially mutagenic[63,66]. In mitochondria, a sub pathway of EJ, microhomology mediated end joining (MMEJ), has been detected[4]. Dysfunction in EJ proteins include consequences like Fanconi anemia and breast and ovarian cancers[11].

vi. The DNA damage tolerance pathway

When replicative polymerases encounter a DNA adduct or an abnormal base, such as a pyrimidine dimer the polymerase is unable to continue and stalls[19,53,67,68]. If not resolved quickly, this can result in a replication fork collapse, which in turn can lead to SSBs, DSBs, and ultimately genomic instability[19,53,67,68]. While DNA damage tolerance (DDT) is not technically a repair pathway, DDT does prevent further DNA damage, and allows the replication process to continue without further incident[19,53,67,68]. The damaged DNA can then be repaired later by a traditional repair pathway[19,53,67,68]. DDT is active when the DNA is being replicated, and therefore only during S phase[59].

There are two sub pathways of DDT, translesion synthesis (TLS) and template switching (TS)[19,53,67,68]. TLS occurs when a high-fidelity replicative polymerase with low-processivity switches with the replicative polymerase to bypass the lesion[19,53,67,68]. Then the replicative polymerase returns and continues replication[19,53,67,68]. TS is not well understood but is thought to involve the sister chromatid[19,53,67,68]. One of the TLS polymerases is located in the mitochondria, suggesting mtDNA also can undergo TLS[4]. People who are deficient in a TLS polymerase can develop a variant of XP [11].

1.3 The base excision repair pathway

The base excision repair (BER) pathway is the major pathway for repair of oxidized DNA damage[18]. BER is evolutionarily conserved from *E. coli* to humans[18,45,69]. The BER pathway processes a large number of non-helix distorting base lesions caused by oxidation, deamination, and alkylation[20,44,46]. Both the nucleus and the mitochondria contain components of the BER pathway that allow these lesions to be repaired[4].

BER takes place in a five-step process illustrated in Figure 1.4: 1) excision; 2) incision; 3) end processing; 4) repair synthesis; and 5) ligation.

Step 1) excision, consists of eleven *N*-glycosylases that can detect specific lesions with some overlapping specificity (e.g. OGG1) (Figure 1.4)[18,70]. These *N*-glycosylases identify base lesions, flip the base out of the helix and sever the glycosidic bond[70]. This step results in an abasic site (Figure 1.4).

Step 2) incision refers to cleaving the sugar phosphate backbone next to the abasic site (Figure 1.4). The abasic site is recognized by an AP endonuclease (e.g. APE1) [18,70]. The AP endonuclease will cleave the DNA backbone on the 5' side leaving an available OH for base replacement (Figure 1.4). Some *N*-glycosylases are bi-functional, meaning they have both glycosylase and AP lyase ability[71] (e.g. NTHL1). An AP lyase will cleave the 3' side of an abasic site[71] (Figure 1.4).

Step 3) end processing, can differ depending on which enzyme (an AP endonuclease or an AP lyase) cleaved the DNA backbone[18,70]. Ultimately, both ends need to be processed to have the appropriate terminal groups to complete repair (Figure 1.4). The 5' end must have an available OH group, and the 3' end must have an available phosphate group. In the event that the abasic site is cleaved on the 3' side by an AP lyase, this processing can be completed by an AP endonuclease (e.g. APE1) (Figure 1.4).

Step 4) repair synthesis, occurs when a polymerase (e.g. Pol β) fills in the single missing base[18,70] (Figure 1.4). This path is called the short patch repair[18,70] (Figure 1.4). Alternatively, a different polymerase (e.g. Pol δ/ε) can polymerize 2-10 bases and displace the preceding DNA bases, known as long patch repair[18,45,70]. Following this extension, the displaced DNA bases, known as a flap, must be removed by Flap endonuclease (e.g. FEN1) before the final step of BER

can be completed[18,70]. While it is not well understood how cells decide to utilize short patch verses long patch repair, the decision seems to depend on which *N*-glycosylase catalyzes the initial reaction[18,70]. Short patch repair is used most often. This path selection also appears to be cell stage dependent.

Step 5) ligation, refers to sealing the nick within the DNA backbone following polymerization (Figure 1.4). A ligase joins the two ends (OH and phosphate group) in the nucleus (e.g. LIG1) and the mitochondria (e.g. LIGIII) [18,70].

1.4 Bi-functional *N*-glycosylase, NTHL1, and orthologs

One BER pathway initiator is the *N*-glycosylase, NTHL1[72]. NTHL1 is categorized as an AP endonuclease[73]. Within this category, NTHL1 is a member of the evolutionarily conserved Endonuclease III (EndoIII) or Nth family proteins (Figure 1.5)[72,73]. AP endonucleases are divided into two classes, class I, which cleave 3' of an abasic site, or class II, which cleave 5' of an abasic site[73]. Human NTHL1, and orthologs in other species (*S. cerevisiae*: Ntg1 and Ntg2, *E. coli*: EndoIII), are classified as class I [73]. Nth family proteins are known for two highly conserved domains, the ENDO3c domain (smart00478) and the iron-sulfur cluster loop (Fe-S domain) (smart 00525)[70]. Within this ENDO3 domain resides a Helix-hairpin-Helix (H-h-H) segment responsible for non-specific DNA binding[70] (Figure 1.5). The consensus for this sequence is $L_{111}X_2LP_{115}GVG_{118}XK_{120}TA_{122}$[74]. This sequence contains the catalytic residue responsible for AP lyase activity, K_{120} in *E. coli*[70] (Figure 1.5). These residues are required to maintain the shape of the H-h-H and position K_{120} into the active site cavity[70].

The second domain present in Nth family proteins, an iron-sulfur cluster loop (Fe-S domain), is located at the C-terminus of the proteins and is present in most of the Nth family

proteins[72,75] (Figure 1.5). This iron-sulfur cluster loop, which is comprised of 21 residues, is important for DNA binding[75]. The consensus sequence is $G_{183}X_3C_{187}X_6C_{194}X_2C_{197}X_5C_{203}$[75]. In addition to AP endonuclease activity, Nth family proteins also can hydrolyze the *N*-glycosidic bond between the base and the sugar moiety, thus making them bi-functional glycosylases[70].

The *S. cerevisiae* genome encodes two homologs of EndoIII that arose from a whole genome duplication, Ntg1 and Ntg2[76,77] (Figure 1.5). They both are bi-functional glycosylases[78]. Ntg1 and Ntg2 share 41% identity and 63% similarity[78] (Figure 1.5). Both Ntg1/2 have the H-h-H motif, however, only Ntg2 retains the Fe-S cluster loop[78] (Figure 1.5). In proteins that lack the Fe-S cluster loop (e.g. Ntg1), it is hypothesized that activity and specificity are reduced[74]. While Ntg1 and Ntg2 both contain a nuclear localization sequence (NLS), only Ntg1 has a mitochondrial targeting sequence (MTS)[78,79]. The catalytic site of Ntg1 is lysine 243[70] (Figure 1.5). Both proteins can efficiently remove the following substrates: 5-hydroxy-6-hydrothymine, 5-hydroxy-6-hydrouracil, 5-hydroxy-5-methylhydantoin, 5-hy-droxyuracil, 5-hydroxycytosine, thymine glycol (Tg), 4,6-diamino-5-formamidopyrimidine (fapyAdenine and fapyGuanine), however the two proteins process each lesion at different rates[80].

1.5 *Saccharomyces cerevisiae* as a model system

Saccharomyces cerevisiae is a eukaryotic model organism that is used extensively to study the highly conserved DNA repair pathways. This organism can exist in either the diploid or haploid state. The haploid state allows for easy genetic manipulation as only a single copy needs to be removed to study gene function. In fact, a budding yeast gene deletion collection of 4,200 non-essential gene knockouts has been available since 2000. DNA repair systems are highly evolutionarily conserved, including BER, allowing researchers to exploit this model system. In

fact, much of the accumulated knowledge of BER comes from studies performed in budding yeast. Any information gathered in yeast is highly translatable to humans, making *S. cerevisiae* an excellent model organism to study the regulation of BER.

1.6 Regulation of NTHL1 and orthologues

While the biochemical mechanism of the BER pathway is well understood, not much is known about how this pathway is regulated. Such regulation is critically important because this pathway must be available for rapid response to any insults that cause DNA damage, but also regulated to ensure that the pathway is not improperly initiated. Some recent studies have begun to provide insight into the regulation of this critical DNA repair pathway. There are number of ways proteins can be regulated, including regulation of steady-state levels at either the transcript or protein levels, post-translational modifications, localization, and protein-protein interactions.

At the gene level, *NTHL1* is comprised of six introns and five exons (8052 bases) and is located on chromosome 16[71]. *NTHL1* is arranged 5' to 5' with the Tuberous sclerosis 2 gene (*TSC2*) with only 357bp separating the two genes and presumably providing the promoter activity for both genes[72,81]. This promoter is either overlapping or bidirectional and contains a CpG island[72,81]. This promoter region contains consensus binding sites for transcription factors Sp1, Ets1, LBP-1, and more but does not contain a typical TATA box[72,81]. *NTHL1* has multiple transcription start sites and resides tail to tail with the 3' neighboring gene, *OCTS2*[71]. The gene loci of *S. cerevisiae NTG1* and *NTG2* are located on chromosome I and XV, respectively[78]. The promoter of *NTG1* has a cis-acting element, a 19bp sequence, that is located at nucleotide -360[82].

At the RNA level, *NTHL1* mRNA is expressed ubiquitously throughout the human body[71]. The mRNA expression level can vary greatly in different cell types with the highest observed

expression in the heart and the lowest in lung and kidney tissues[71]. High levels of *NTHL1* mRNA expression in the heart are suspected to be caused by the high levels of ROS, a byproduct of ATP production, and the resulting demand for oxidative phosphorylation that can trigger oxidative stress[72]. Interestingly, alternative splice variants of *NTHL1* mRNA are detected in the liver, hippocampus (800 bases), and blood (1030 bases), but no studies have been performed to assess the functional consequences of this alternative splicing[83].

One study estimates that the expression of *NTG1* mRNA is one transcript per cell[69]. Consistent with this prediction, we were unable to detect *NTG1* or *NTG2* mRNA by northern blot under endogenous conditions[78]. We were, however, able to detect both transcripts via qRT-PCR[78]. Interestingly, *NTG2* mRNA is present at a relatively lower level than *NTG1* mRNA[78]. This finding could suggest that Ntg1 is the primary Nth family member present in budding yeast under resting conditions.

Not much is known about the transcriptional regulation of endonuclease III/Nth family glycosylases. *NTHL1* mRNA expression is regulated by the cell cycle[83,84]. In human keratinocytes, expression is low at the G_0 and G_1 phases of the cell cycle[83,84]. Expression begins to increase at the start of S phase and remains high throughout S phase[83,84]. After mitosis, expression decreases[83]. *NTHL1* mRNA expression could be differentially expressed in different cell types, or throughout development and factors that control this cell cycle dependent expression have not yet been identified.

NTG1 mRNA expression may be inducible in response to DNA damaging agents, although there are some conflicting results on this point. One study found that *NTG1* mRNA expression was induced after exposure to hydrogen peroxide (H_2O_2), methylmethane sulfonate (MMS), and 4-NQO[69]. They found particularly high expression in the presence of menadione, suggesting some

specificity of the DNA damage to the inducibility of Ntg1[69]. Our group only detected increased expression in response to menadione[21]. This discrepancy between studies could be attributed to differences in experimental or growth conditions. The 19bp sequence in the promoter region, mentioned above, is required to induce expression in response to hydroxyurea (HU) in another gene, *DIN7,* but has not yet been investigated in connection with Ntg1[82]. Ntg2 expression has never been readily detectable under any of the conditions tested[78]. Currently, we cannot address whether an increase in mRNA levels corresponds to an increase in protein levels as no antibodies are available to analyze expression of endogenous Ntg1 or Ntg2. Based on this evidence in budding yeast, there is a probability that expression of NTHL1 can be induced in response to a DNA damaging agent, but further studies will need to be conducted.

At the protein level, the N-terminus of NTHL1 is a region important for proper subcellular protein localization. The MTS and bipartite NLS and are located in this region[83] (Figure 1.5). In cells, NTHl1 can be detected in the nucleus, the mitochondria and the cytoplasm depending on the cell type or species examined[83]. This localization also appears to be tissue dependent[83]. Whether NTHL1 can be re-localized to a specific compartment in response to DNA damage is not known. In budding yeast, both Ntg1 and Ntg2 also have localization sequences located at the N-terminus[78]. Ntg1, which contains both a bipartite NLS and an MTS, is localized to both the nucleus and the mitochondria[69]. Ntg2 only contains an NLS and therefore is only located in the nucleus[69]. Ntg1 can be recruited to the nucleus in response to DNA damage induced by H_2O_2, or the mitochondria in response to DNA damage induced by a combination of H_2O_2 and antimycin A[79]. These findings highlight how cellular localization is one mechanism that can regulate Ntg1 activity.

The N-terminal domain of NTHL1 and Ntg1 is also critical for proper protein regulation. The N-terminal domain of NTHL1 inhibits the rate of lesion release[85]. Additionally, this domain

allows NTHL1 to homodimerize leading to an auto-inhibitory effect on lesion processing in a concentration dependent manner both *in vitro* and *in vivo*[85]. Loss of the first 55-80aa in NTHL1 stimulates the activity of NTHL1, as does the presence of the downstream BER player, APE1[85,86]. The N-terminus of NTHL1 and Ntg1 has a long positively charged region that may be necessary for protein-protein interactions[78,87]. For example, the Y box-binding protein 1 (YB-1) physically interacts with NTHL1 and stimulates both glycosylase and AP lyase activity of NTHL1[55]. Also, *in vitro* excision experiments show that the excision of the DNA lesion, Tg, by NTHL1 is greatly increased when binding partner and NER protein, XPG, is added[88].

Post-translational modifications (PTMs) also play a major role in regulating protein function and interactions[89]. Both Ntg1 and Ntg2 are post-translationally modified by the Small Ubiquitin-like MOdifier (SUMO), but whether this modification is also present on the human NTHL1 protein had not yet been explored prior to this thesis work[90]. The sites, the relevant proteins involved, and the functional consequence of Ntg1 and Ntg2 sumoylation are currently unknown. However, sumoylation of Ntg1 is only detected in the nucleus, and not seen in the mitochondria[90].

1.7 Dysregulation of human base excision repair protein, NTHL1

As a DNA repair protein, loss of NTHL1 could have detrimental effects on cellular function. A recessive homozygous loss of function germline mutation in *NTHL1* has been identified as a contributor to a novel colon cancer predisposition syndrome[91]. This nonsense mutation results in loss of NTHL1 protein[91]. NTHL1 is localized to both the nucleus and the cytoplasm, but localization can differ markedly based on cell type[92]. However, a subset of gastric and colon cancer patient samples had NTHL1 excluded from the nucleus and restricted to the

cytoplasm[92,93]. Effectively, these two examples demonstrate the need for properly controlled regulation of NTHL1 to maintain genome integrity.

Although loss of NTHL1 function can contribute to tumorigenesis, *NTHL1* is amplified or the mRNA is upregulated in a variety of tumor types far more often than *NTHL1* is deleted[94]. A recent study demonstrated that NTHL1 overexpression causes genomic instability, replication stress signaling, and an increased reliance on EJ in normal human bronchial epithelial cells[94]. Cells engineered to overexpress NTHL1 showed a number of early cancer hallmarks, including loss of contact inhibition[94]. Interestingly, these results were not dependent on NTHL1 catalytic activity[94].

1.8 NTHL1 and DNA repair pathway crosstalk

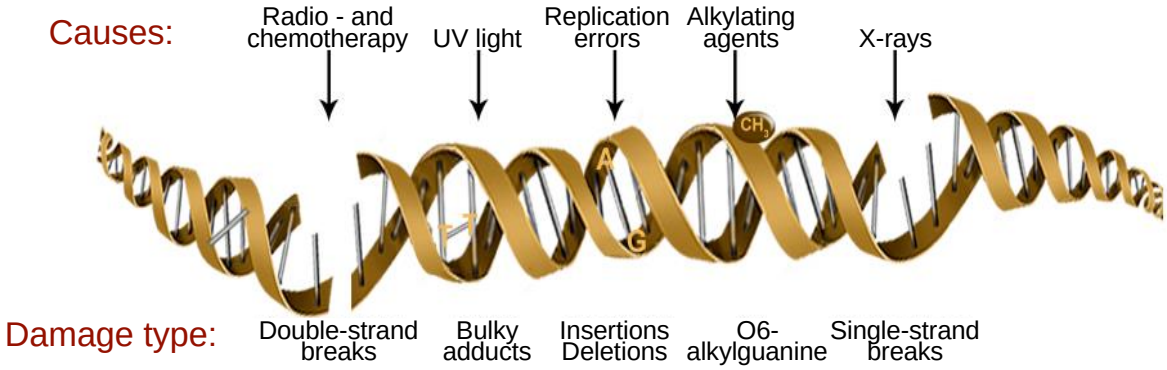
DNA repair pathways were originally characterized as stand-alone pathways. However, beyond the interactions within a given DNA repair pathway, there is evidence of interplay between DNA repair pathways[43,86,88,95,96]. This is not surprising since some repair intermediates, such as AP sites and SSBs, can be repaired by multiple pathways. While the idea of DNA repair pathway crosstalk is still in its infancy, there is evidence to suggest we have a lot to learn about this topic.

Most of the known BER crosstalk with other DNA repair pathways takes place at the first step. *In vitro* studies of NTHL1 show poor excision of NTHL1 substrate, Tg[88]. However, *in vitro* studies, indicate that the addition of either BRCA1 (a component of HR and involved in MMR) or XPG (a component of NER), but not other similar proteins, greatly increases the efficiency of Tg excision[88,97]. One study found that knocking down NTHL1 significantly decreased alt-NHEJ. Indeed, overexpression of NTHL1 resulted in an increase in NHEJ and a decrease in HR, consistent with crosstalk between BER and other repair pathways[94]. Taken together, these data suggest that NTHL1 plays a role in promoting EJ and suppressing HR.

1.9 Summary

In this thesis, the work addresses two aspects of regulation of the BER pathway taking advantage of the budding yeast model system. In Chapter 2, the sites of SUMO modification on the Ntg1 protein were mapped to identify five major sites of modification. A variant of Ntg1 that could not be modified by SUMO was generated and analyzed. In addition, this work demonstrated that the human NTHL1 protein can also be modified by SUMO. In Chapter 3, we developed a budding yeast system to model how overexpression of NTHL1 leads to early hallmarks of cancer. Studies were performed to demonstrate that overexpression of Ntg1 causes similar DNA damage as that induced by overexpression of NTHL1 in mammalian cells. This system could then be employed to screen a panel of budding yeast mutants in a variety of DNA repair pathways. These studies provide insight into how cells respond to damage that occurs upon overexpression of Ntg1 as well as potential pathways that interface with BER. Taken together, this work provides important initial insights into how the BER pathway is regulated at a critical first step that initiates this repair pathway. These studies exploiting the budding yeast system can be used to develop hypotheses for how dysregulation of NTHL1 leads to cancer.

1.10 Figures and tables



Modified from Stephen Jackson and Cheryl Bishop 2006

Figure 1.1: Schematic summarizing causes of DNA damage and the resulting types of damage.

The schematic depicts different types of DNA damages (listed below the image): double-strand breaks, bulky adducts, insertions and deletions, O_6-alkylguanine, and single-strand breaks. Additionally, this figure depicts examples of causes for each respective type of DNA damage (listed above the image); radio- and chemotherapy, ultraviolet (UV) light, replication errors, alkylating agents, and x-rays.

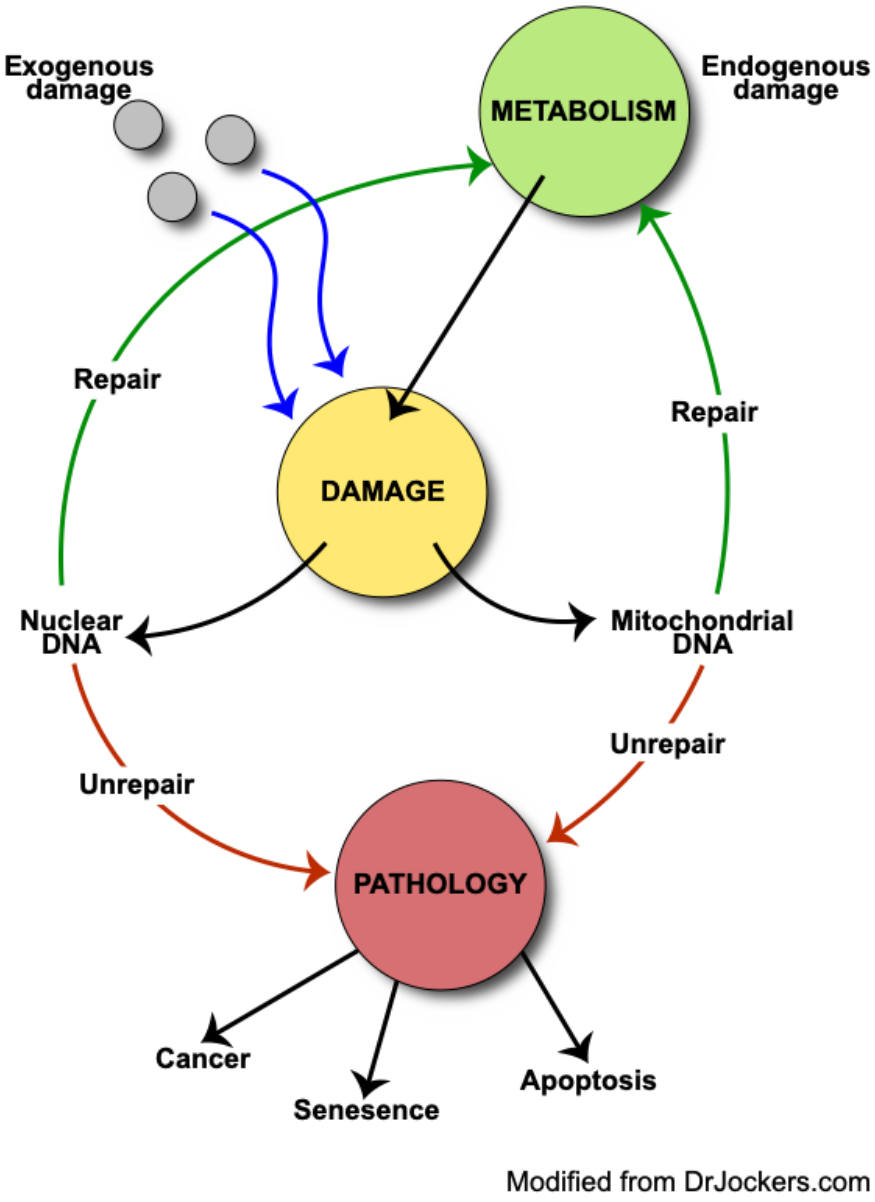


Figure 1.2: Pathway of the biological consequences of DNA damage.

The schematic depicts DNA damage (yellow circle) from either exogenous or endogenous sources (blue lines) to nuclear or mitochondrial DNA. When repaired (green lines), the cell continues normal function and metabolism (green circle). However, if the damage is left unrepaired (red lines), the cell can undergo abnormal function and develop different pathologies (red circle) including cancer, cell senescence, and apoptosis.

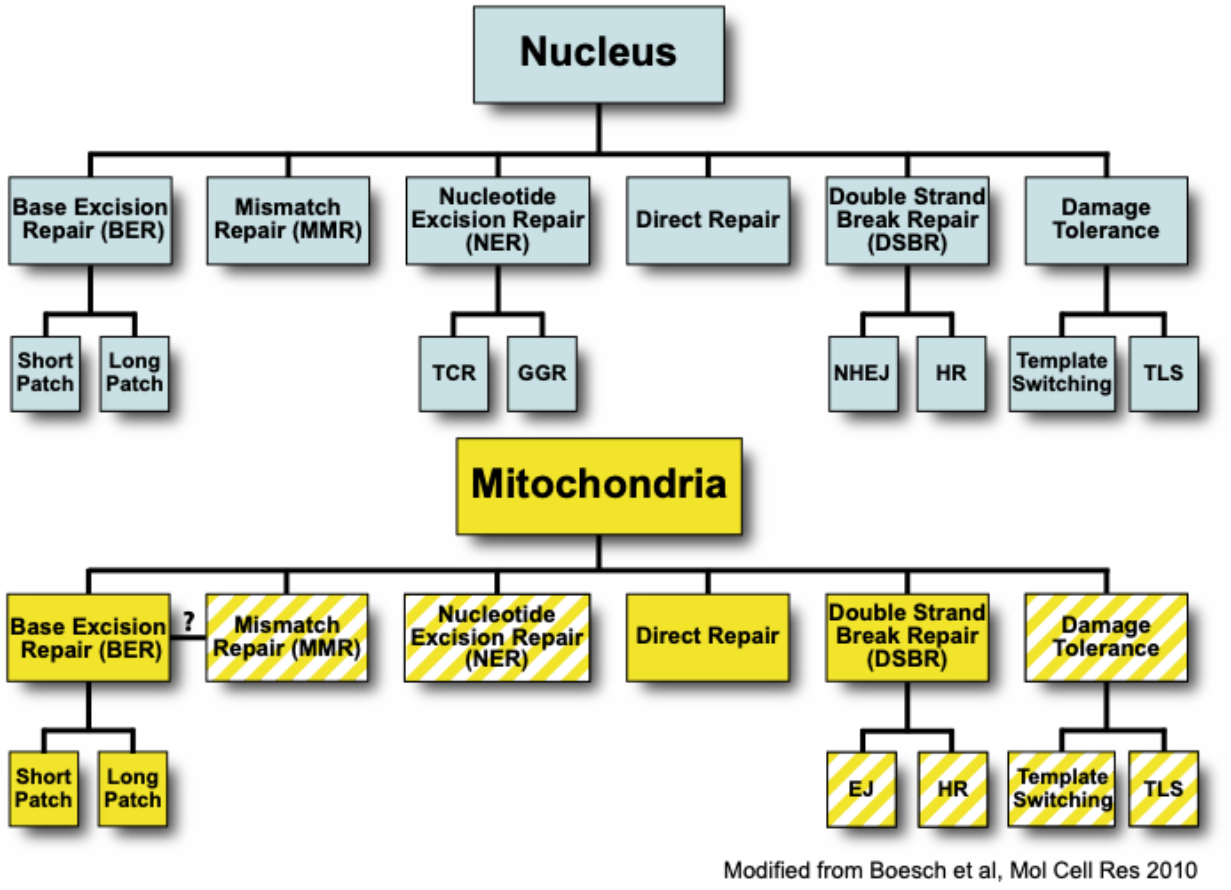


Figure 1.3: DNA repair pathways in the nucleus and the mitochondria.

Six major pathways of repair are present in both the nucleus and the mitochondria. Solid colored backgrounds (blue: nuclear, and yellow: mitochondrial) identify established pathways, while hatched backgrounds indicate pathways that need further confirmation or have some, but not all, established necessary nuclear components localized to the mitochondria. In humans, a component traditionally identified as a base excision repair protein, YB-1, appears to be critical for mismatch repair in mitochondria (identified by a ?)[55]. Abbreviations: TCR, transcription-coupled repair; GGR, global genome repair; EJ, end joining; HR, homologous recombination; TLS, translesion synthesis.

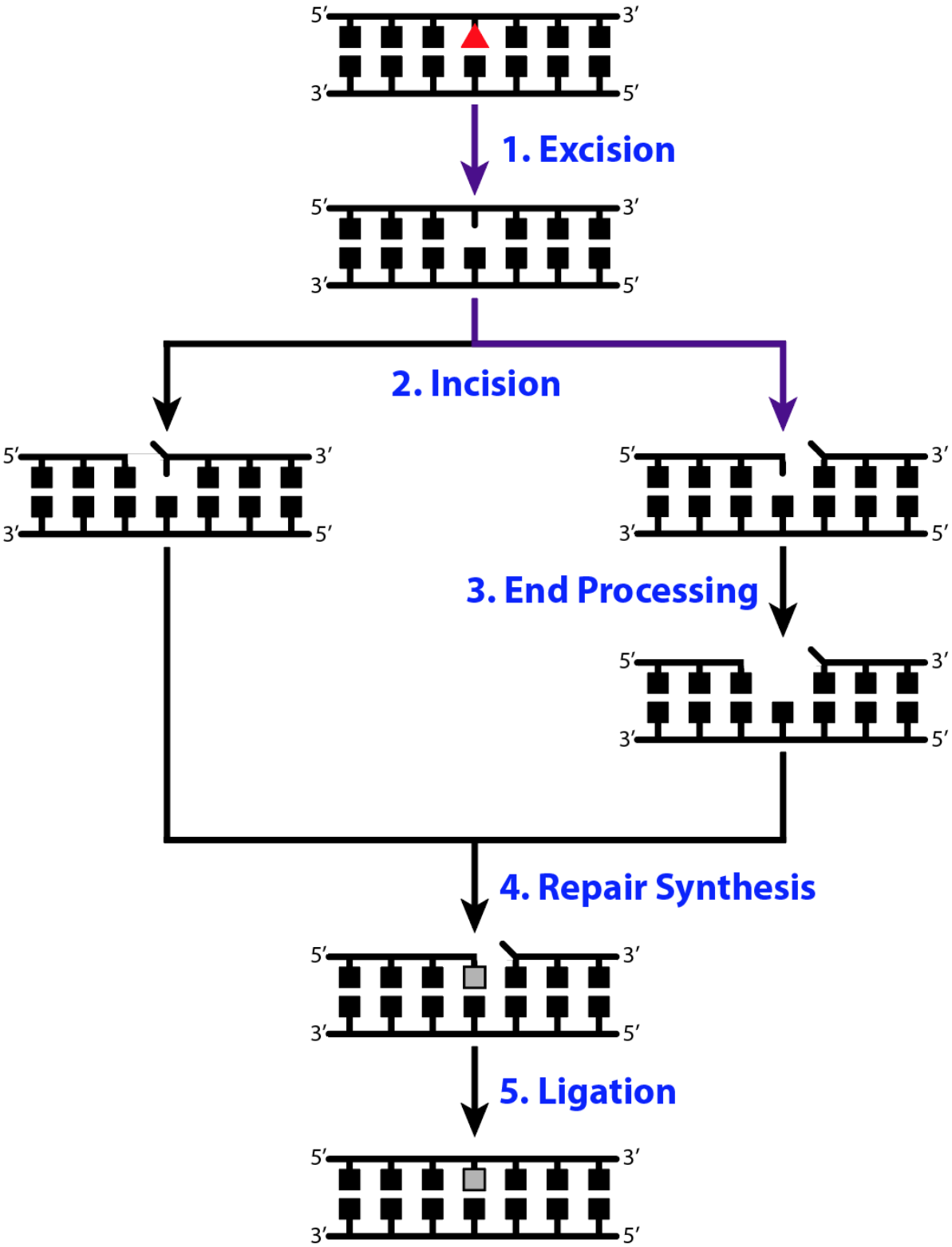


Figure 1.4: Basic steps of the evolutionarily conserved base excision repair pathway.

BER is the main pathway for repair of oxidative DNA damage. The BER pathway consists of 5 main steps: 1) excision; 2) incision; 3) end processing; 4) repair synthesis; and 5) ligation. Short patch BER is depicted here. The red triangle indicates DNA damage and the grey square indicates a repaired base. The purple lines and arrows indicate the steps human NTHL1 and *S. cerevisiae* Ntg1 facilitate.

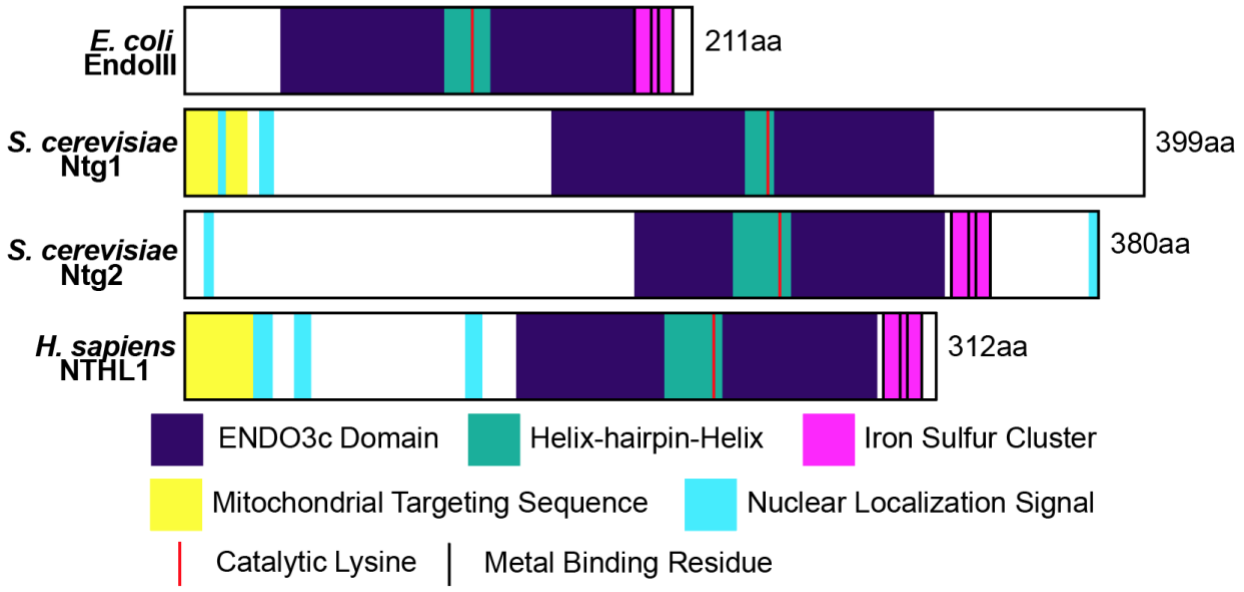


Figure 1.5: Domain structure of evolutionarily conserved Endonuclease III/Nth family proteins.

Domain structure of the Endonuclease III/Nth family proteins from *E. coli*, *S. cerevisiae*, and *H. sapiens*. *S. cerevisiae* contains two NTHL1 orthologs that arose from a whole genome duplication event. ENDO3c Domain (purple), Helix-hairpin-Helix domain (teal), Iron Sulfur Cluster (magenta), Mitochondrial Targeting Sequence (yellow), Nuclear Localization Signal (blue), Catalytic Lysine (red line), Metal Binding Residue (black line).

Chapter 2

Identification of SUMO modification sites in the base excision repair protein, Ntg1

Daniel B. Swartzlander[a,b,1,2], **Annie J. McPherson**[a,b,2], Harry R. Powers[a,3], Kristin L. Limpose[a,c], Emily G. Kuiper[a,d], Natalya P. Degtyareva[a,e], Anita H. Corbett[a,e,4], and Paul W. Doetsch[a,e,f,g,4]

[a]Department of Biochemistry, Emory University School of Medicine, Atlanta, GA 30322

[b]Graduate Program in Genetics and Molecular Biology, Emory University School of Medicine, Atlanta, GA 30322

[c]Graduate Program in Cancer Biology, Emory University School of Medicine, Atlanta, GA 30322

[d]Graduate Program in Biochemistry, Cell and Developmental Biology, Emory University School of Medicine, Atlanta, GA 30322

[e]Winship Cancer Institute, Emory University School of Medicine, Atlanta, GA 30322 [f]Department of Radiation Oncology, Emory University School of Medicine, Atlanta, GA 30322

[g]Department of Hematology and Medical Oncology Emory University School of Medicine, Atlanta, GA 30322

The initial data were collected by DBS. AJMD assembled the figures for the study, wrote and edited the manuscript based on a thesis chapter, and generated data presented in Figures 2.1 and 2.5.

Published in DNA Repair (Amst). 2016 Dec;48:51-62. doi: 10.1016/j.dnarep.2016.10.011. Epub 2016 Oct 31).

2.1 Abstract

DNA damaging agents are a constant threat to genomes in both the nucleus and the mitochondria. To combat this threat, a suite of DNA repair pathways cooperate to repair numerous types of DNA damage. If left unrepaired, these damages can result in the accumulation of mutations which can lead to deleterious consequences including cancer and neurodegenerative disorders. The base excision repair (BER) pathway is highly conserved from bacteria to humans and is primarily responsible for the removal and subsequent repair of toxic and mutagenic oxidative DNA lesions. Although the biochemical steps that occur in the BER pathway have been well defined, little is known about how the BER machinery is regulated. The budding yeast, *Saccharomyces cerevisiae* is a powerful model system to biochemically and genetically dissect BER. BER is initiated by DNA *N*-glycosylases, such as *S. cerevisiae* Ntg1. Previous work demonstrates that Ntg1 is post- translationally modified by SUMO in response to oxidative DNA damage suggesting that this modification could modulate the function of Ntg1. In this study, we mapped the specific sites of SUMO modification within Ntg1 and identified the enzymes responsible for sumoylating/ desumoylating Ntg1. Using a non-sumoylatable version of Ntg1, ntg1ΔSUMO, we performed an initial assessment of the functional impact of Ntg1 SUMO modification in the cellular response to DNA damage. Finally, we demonstrate that, similar to Ntg1, the human homologue of Ntg1, NTHL1, can also be SUMO-modified in response to oxidative stress. Our results suggest that SUMO modification of BER proteins could be a conserved mechanism to coordinate cellular responses to DNA damage.

2.2 Introduction

Genomes in both the nucleus and mitochondria are constantly exposed to various exogenous and endogenous DNA damaging agents (1). A suite of DNA repair pathways cooperate to ensure the efficient repair of numerous types of DNA damage that result from such exposures (2, 3). Oxidative DNA damage, caused by numerous sources including cellular metabolism (4, 5) and exogenous factors (6), is one of the most common forms of DNA damage. Estimates suggest that 90,000 oxidative lesions and 200,000 apurinic/apyrimidinic (AP) sites are generated per human cell per day (7–9). Unrepaired lesions can result in the accumulation of mutations which can trigger deleterious consequences including cancer and neurodegenerative disorders (1–3, 7, 8, 10–16). The base excision repair (BER) pathway is primarily responsible for the removal and repair of toxic and mutagenic oxidative DNA damage (3, 17–19). Numerous studies have defined in detail the biochemical steps that occur in the BER pathway (3, 20), but little is known about how the BER machinery is regulated (21).

BER is initiated by the recognition and hydrolysis of a damaged base by a DNA *N*-glycosylase leaving an AP site (3, 20, 22, 23). The AP site is then further processed to create a nick in the DNA backbone (3, 20, 22, 23). Subsequent steps create a single-strand break that is then filled by a specialized DNA polymerase and sealed by ligase (3, 20, 22, 23). These steps must occur in a sequential manner ensuring that AP sites and single-strand breaks are properly managed to allow repair at the initial site of DNA damage without causing collateral damage via accumulation of BER intermediates (3, 20, 22, 23). The human NTHL1 protein, which is a bifunctional Endonuclease III–like *N*-glycosylase/AP lyase, is responsible for initiating repair of a wide array of oxidative lesions (21, 24–26). As the initiating factor in the BER pathway (3, 21, 22), NTHL1 must be regulated to ensure that repair is rapid, but also regulated to prevent the

accumulation of toxic and mutagenic AP sites and single-strand breaks that are the products of NTHL1 enzymatic activity (21, 24– 26). *N*-glycosylase regulation could occur through a number of distinct mechanisms including modulating protein levels, protein localization, protein-protein interactions, and post-translational modifications (27–35).

Recent discoveries highlight the importance of *N*-glycosylase regulation in cancer (36, 37). Several studies identified mutations in the NTHL1 gene in a recently characterized cancer predisposition syndrome (38–40). These heterozygous loss-of-function mutations in NTHL1 predispose patients to colorectal cancer and other forms of cancer (38–40). Altered NTHL1 function can also result in mislocalization/accumulation of the protein in the cytoplasm of cancer cells in a subset of gastric tumors (36). These studies provide evidence that proper function of NTHL1 is critical to maintain genomic integrity and cellular homeostasis.

Much of the work that has contributed to our knowledge of DNA repair mechanisms has exploited the budding yeast *S. cerevisiae* as DNA repair pathways are conserved through evolution (41). Recent studies of the *S. cerevisiae* orthologues of NTHL1, Ntg1 and Ntg2, reveal that these proteins are post-translationally modified by the Small Ubiquitin-like MOdifier, SUMO (24, 42). The Ntg1 protein is modified in response to DNA damage (24, 42). Sumoylation has the potential to function in a number of regulatory roles including modulating protein-protein interactions and protein activity (27–35). One well-characterized example of SUMO-mediated regulation of the BER pathway is the human thymine DNA glycosylase (TDG), where sumoylation of TDG triggers a conformational change which alters the DNA binding pocket of the enzyme to influence enzyme turnover (43–45). This conformational change in TDG decreases the affinity of TDG for DNA leading to an increase in the off rate and hence an increase in the catalytic efficiency (turnover) of

TDG (43, 44). Similarly, sumoylation could also modulate the function of Ntg1; however, the impact of SUMO modification on Ntg1 function has not yet been explored.

Critical to defining the functional role of SUMO modification of Ntg1 is identifying the SUMO modified sites within Ntg1. In this study, we identify the enzymes that mediate/ regulate sumoylation of Ntg1. We also map the SUMO-modified sites on Ntg1 and perform an initial assessment of the functional importance of sumoylation of Ntg1. In addition, we demonstrate that, similar to Ntg1, human NTHL1 can also be SUMO-modified in response to oxidative stress. Our results suggest that SUMO modification of BER proteins could represent an evolutionarily conserved mechanism by which cells respond to oxidative DNA damage.

2.3 Materials and methods

2.3.1 Strains, plasmids, and media

All haploid *S. cerevisiae* strains and plasmids used in this study are listed in Table 1. *S. cerevisiae* cells were cultured at 25°C, 30°C, or 37°C in YPD medium (1% yeast extract, 2% peptone, 2% dextrose, 0.005% adenine sulfate, and 2% agar for plates) or SD medium (0.17% yeast nitrogen base, 0.5% ammonium sulfate, 2% dextrose, 0.5% adenine sulfate, and 2% agar for plates). In order to introduce plasmids, cells were transformed by a modified lithium acetate method (46).

A centromeric vector (*CEN, URA3*), pRS316 (47) was employed as the backbone for the generation of a construct expressing C-terminally tagged Ntg1-TAP fusion protein (pD0436). The insert was amplified using the primers listed in Table 2 and inserted at the *Not*I restriction site of pRS316 (47). The insert includes the tetracycline repressible promoter (Tet-Off) and the C-terminally tagged *NTG1-TAP* fusion from the DSC0295 strain (24). The *S. cerevisiae* haploid

deletion mutant *ntg1Δ* (DSC0470) generated by dissection of tetrads derived from heterozygous diploid hDNP19 (19), and the SUMO pathway mutant collection (E3 ligase mutant strains, *siz1Δ*, *siz2Δ*, and *siz1Δ/siz2Δ* and desumoylase mutant stains *ulp1-ts* and *ulp2Δ*) were utilized to assess the level of sumoylated wildtype and mutant Ntg1 (19, 48, 49). All lysine to arginine amino acid substitutions (Figure 2.3C) were created by site-directed mutagenesis performed using the QuikChange II Site-Directed Mutagenesis Kit (Stratagene) with the primers listed in Table 2. The resulting plasmids were sequenced to ensure the introduction of the desired mutation and the absence of any additional mutations.

To express recombinant Ntg1, the *NTG1* open reading frame was cloned into pET-15b (Invitrogen) to generate N-terminal His6 epitope tagged His6-Ntg1 (pD0390) (Table 1). Site-directed mutagenesis of *His6-NTG1* was performed at lysines 20, 38, 376, 388, and 396 (lysines to arginines) to create a nonsumoylatable Ntg1 (ntg1ΔSUMO), *His6-Ntg1ΔSUMO* (pD0493), and at lysine 243 (lysine to glutamine), *His6-Ntg1Δcat* (pD0394) (Table 1). Expression vectors were transformed into DE3 cells.

Site-directed mutagenesis at the endogenous NTG1 locus of the wildtype (DSC0367) parent was performed via *delitto perfetto* protocol (50) to generate *ntg1K20,38,376,388,396R*. The resulting variants were then crossed with haploid BER-Nucleotide Excision Repair- (NER-) mutants to create diploids which were then dissected to identify cells with each Ntg1 variant BER*/NER- strain (DSC0367, DSC0369, DSC0371, DSC0561).

NTHL1 was cloned from the RG214598 plasmid (Origene) using the NTHL1-Flag primer pair for the addition of the Flag-tag and cloned into the pcDNA3.1 (+) vector using the *Hin*dIII and *Bam*HI sites.

2.3.2 Exposure to DNA damaging agents

S. cerevisiae cells were grown in 5–35 mL YPD or SD -URA media to either 2 × 107 or 1 × 108 cells/mL, centrifuged, and washed with water. Cells were then resuspended in 5–35 mL water, YPD, or plated onto YPD agar plates containing the appropriate agent: 20 mM hydrogen peroxide (Sigma); or 0.005–0.3% methyl methanesulfonate (MMS) (Sigma). Cells were exposed to agents for 1–2 hours as indicated at 30°C or 37°C.

2.3.3 Immunoblotting Ntg1

The steady-state level of each Ntg1-TAP fusion protein variant was assessed by immunoblotting whole cell lysates with the rabbit polyclonal anti-TAP antibody (1:3,333 dilution, Open Biosystems) to determine the relative level of differentially modified Ntg1 products. An anti-3-phosphoglycerate (PGK) antibody (1:10,000 dilution; Invitrogen) was used as a control determine the relative level of protein lysate loaded into each lane.

The analysis of immunoblots was performed utilizing the ECL Plex immunoblotting detection system (Amersham), the Typhoon Trio variable mode imager (GE Healthcare), and the ImageQuant TL software package (GE Healthcare). To quantify the percentage of modified Ntg1-TAP, the ratio of modified Ntg1 bands to total Ntg1 signal (including modified and unmodified) was determined for wildtype Ntg1 and each lysine to arginine amino acid substitution variant of Ntg1. Previous work demonstrates that modified Ntg1 contains at least one covalently linked SUMO and the size of higher bands is consistent with multiple SUMO additions (24). Standard error of the mean was calculated for each. The two- sample Student’s t-test was employed to test for significance (α=0.05).

2.3.4 Cultured cell lines and cell culture

HT29 colon adenocarcinoma cells were cultured in McCoy's 5A modified media (Corning) and supplemented with 10% FBS, penicillin and streptomycin. Cultured cells were passaged every 3–4 days, or upon 80% confluency.

2.3.5 NTHL1-Flag immunoprecipitation

HT29 colon adenocarcinoma cells were seeded at a density of 1×106 cells in 100 cm2 dishes. Transfection of the NTHL1-Flag construct or empty Flag vector was performed using Lipofectamine3000 (Invitrogen) and a final concentration of 10 μg plasmid/dish. The hydrogen peroxide incubation was performed with a final concentration of 125 μM hydrogen peroxide in sterile PBS for 15 minutes at 37°C. All cells were lysed in NP40 buffer (50 mM Tris pH 8.0, 100 mM NaCl, 32 mM NaF, 0.5% NP40 detergent) supplemented with protease and phosphatase inhibitors (Thermo Scientific), and SENP (de-SUMOylase) SUMO-2 aldehyde inhibitors (Enzo Life Sciences). Antibodies for Flag (Rabbit, 2368; Cell Signaling) or IgG (mouse, ab77118; abcam) were conjugated to Protein G Dynabeads (10007D; Life Sciences) for 2 hours prior to adding lysates. For each sample, 500 μg of total protein was added to the beads and rotated overnight at 4°C. Beads were washed three times in NP40 buffer for 5 minutes each. NTHL1-Flag was eluted from the beads using a 3X Flag® peptide (Sigma) for 2 hours at a working concentration of 100 μg/mL per the manufacturer's instructions.

Following Flag peptide elution, samples were added to Laemmli buffer (50% glycerol, 10% SDS, 100 mM Tris, pH 6.8), boiled for 5 minutes at 95°C, and loaded on 4–12% Bis-Tris gels (Invitrogen). Proteins were transferred onto a polyvinylidene fluoride membrane and blocked with 5% ECL prime (GE Healthcare) in 0.1% PBST for 1 hour at room temperature. Blots were

incubated in primary antibodies overnight at 4°C. All washes were performed in 0.1% PBST at room temperature, and the corresponding horseradish peroxidase-conjugated secondary antibodies were added for 1 hour. Antigen-antibody complexes were detected using Supersignal™ west pico chemiluminescent substrate kit (Thermo Scientific). Antibodies used for western blotting were: NTHL1 (mouse, cat # MAB2675; R&D Systems) and SUMO-2/3 (rabbit, made in Nicholas Seyfried lab, Emory University).

2.3.6 Structural modeling

The Protein Homology/analogY Recognition Engine version 2.0 (Phyre2) server was used to generate a model of Ntg1 based on its E. coli Endonuclease III homolog (PDB ID: 2ABK). The N-terminal and C-terminal domains do not share homology with *E. coli* Endonuclease III but align to other bacterial endonucleases. The N-terminal domain aligns to the restriction endonuclease BsaW*I* (PDB ID: 4ZSF) and the C-terminal domain aligns to the endonuclease Bgl*II* (PDB ID: 1DFM). PyMOL Molecular Graphics System, Version 1.8 Schrödinger LLC was used to model these structures.

2.3.7 Overexpression and purification of the recombinant Ntg1 variants for *in vitro* DNA strand scission assay

To assess the functional consequences of changing five lysines (K20,K38,K376,K388,K396) to arginine within Ntg1, we expressed and purified recombinant protein containing these five amino acid substitutions. We designated this recombinant protein ntg1(K->R)5. As controls, we employed wildtype Ntg1 and a catalytic mutant of Ntg1 (lysine 243 to glutamine) which we term ntg1Δcat. Recombinant Ntg1 was purified as previously described

(51). Briefly, *Escherichia coli* BL21 (DE3) cells containing each variant *His6-Ntg1* plasmids were grown to an OD600 of 0.5–1.0 and expression induced for 4 hours at 25°C. Cells were lysed via sonication and the supernatant was applied to Ni-NTA agarose beads purification (Qiagen) to crudely purify the His6-Ntg1 variants. Crude lysate was eluted through a gravity flow column (BIORAD) and dialyzed. Crudely purified His6-Ntg1 variants were further purified to apparent homogeneity by fast protein liquid chromatography.

2.3.8 Preparation of oligonucleotide and DNA strand scission assay

To assess the functional consequences of changing five lysines (K20,K38,K376,K388,K396) to arginine within Ntg1, we employed an in vitro strand scission assay. An oligonucleotide containing dihydrouracil (DHU) at position 13 (DHU-31mer) was purchased from Midland Certified Reagent Company (Midland, TX, USA). A complementary strand containing a guanine opposite the DHU position was obtained from Eurofins MWG/Operon (Huntsville, AL, USA). The DHU-31mer was 5'-end-labeled with [γ-32P] ATP (Amersham) and T4 polynucleotide kinase (Promega) prior to annealing to the complementary strand (24). Single-stranded DHU-31mer was annealed in a 1:1.6 molar ratio to the appropriate complementary strand, heated to 80°C for 10 minutes and cooled slowly to room temperature.

The AP lyase activity of purified Ntg1 variants (Ntg1, ntg1(K->R)5, and ntg1Δcat) was assayed as previously described (51). Briefly, DNA strand scission assays were carried out in a standard reaction buffer (20 mL) containing 100 mM KCl, 10 mM Tris–HCl, pH 7.5, 1 mM EDTA, 50 fmol of labeled DNA substrate and 20 fmol of Ntg1 protein. Reactions were performed at 37°C for 15 minutes and then stopped by the addition of 10 μL of loading buffer (90% formamide, 1mM EDTA, 0.1% xylene cyanol and 0.1% bromophenol blue) followed by heating at 95°C for 5

minutes. Reaction products were then resolved on a denaturing-urea polyacrylamide gel (15%) and analyzed with a Typhoon Trio variable mode imager (GE Healthcare).

2.3.9 Functional analysis of Ntg1 *in vivo*

To test the sensitivity and the biological function of the Ntg1 complete sumoylation null mutant (ntg1K20,38,376,388,396R), which we term ntg1ΔSUMO (DSC0561), to DNA damaging agents, a serial dilution and spotting assay was employed. Each strain was grown at 30°C to an OD600 of 0.3 – 0.6 in YPD, washed in 5 mL of water, and then diluted to 2×10^7 cells/mL in water. Five-fold serial dilutions of cells were then plated onto plates containing only YPD or YPD with 0.005% MMS. Plates were incubated at 30°C and then analyzed for sensitivity at days 2 and 4.

Growth kinetics experiments were carried out using *S. cerevisiae* cells that express each Ntg1 variant encoded at the endogenous *NTG1* locus in a DNA repair compromised background (DSC0367, DSC0369, DSC0371, and DSC0561). The growth kinetics of four independently isolated ntg1ΔSUMO variants (DSC0561) and four wildtype Ntg1 (DSC0371), all in a DNA repair compromised background, were tested by analyzing growth curves. Ntg1 sumoylation mutants were grown to saturation over 2 days at 30°C in YPD. Cell concentrations were normalized by OD600, and then samples were diluted to an OD600 of 0.05 in 150 μL of YPD medium containing 0, 0.005, or 0.010% MMS and added to the wells of a 96-well microtiter plate. Cell samples were loaded in duplicate, were grown at 30°C with shaking, and absorbance at OD600 was measured every 30 minutes for 48 hours in an ELX808 Ultra microplate reader with KCjunior software (Bio-Tek Instruments, Inc.). The samples for each genotype and duplicate were averaged for every time point and differences between the two genotypes was analyzed by Students t-test.

2.4 Results

2.4.1 Genetic analysis of the Ntg1 sumoylation pathway

Our group has previously shown that in response to cellular exposure to hydrogen peroxide, Ntg1 is post-translationally modified by SUMO (24). As shown in Figure 2.1A, several Ntg1 bands are detected upon exposure to hydrogen peroxide suggesting that Ntg1 could be modified by multiple post-translational modifications, including the possibility for addition of multiple SUMO moieties. To determine whether SUMO modification of BER proteins that initiate repair is conserved, we tested whether we could detect sumoylation of human NTHL1. For this experiment, we transfected HT29 colon adenocarcinoma cells with NTHL1-Flag or an empty Flag vector. Cells were treated with hydrogen peroxide and immunoprecipitated with anti-Flag or with IgG as a control. Total lysate and bound fractions were subjected to immunoblotting to detect NTHL1 and SUMO. As shown in the top panel of Figure 2.1B, we detect NTHL1 in the input and NTHL1 is enriched in the bound fraction, as expected. In samples from cells treated with hydrogen peroxide, a higher molecular weight band of NTHL1 appears, suggesting a post-translational modification. Consistent with SUMO modification, a band of the same molecular weight is recognized by an anti- SUMO-2/3 antibody. The extent of modification of NTHL1 is greatly increased in response to hydrogen peroxide exposure (Figure 2.1B, bottom). We do not detect NTHL1 or SUMO-2/3 in the control IgG immunoprecipitation. Thus, both *S. cerevisiae* Ntg1/Ntg2 (24) and human NTHL1 can be modified by SUMO.

Sumoylation involves a series of conjugations that, in *S. cerevisiae*, are catalyzed by the E1 (Uba2/Aos1 heterodimer), the E2 (Ubc9), and one of four E3 (Siz1, Siz2, Mms21, and Zip3) ligases (52–57). These enzymes catalyze the attachment of the SUMO protein to a substrate lysine

residue through formation of an isopeptide bond (52–57). Sumoylation is a dynamic process that is readily reversible by SUMO proteases, which in *S. cerevisiae* are Ulp1 and Ulp2 (58, 59).

To define the pathway by which Ntg1 is SUMO modified, we examined *S. cerevisiae* cells lacking the E3 ligases, Siz1 and Siz2, as well as Siz1/Siz2 double deletion cells (53, 60). We first examined Ntg1 sumoylation in these *siz1Δ* and *siz2Δ* mutant cells in response to hydrogen peroxide exposure (Figure 2.1C, D). In the *siz1Δ* cells, we detected reduced levels of Ntg1 sumoylation (1.5%) compared to wildtype control cells (4.7%) (Figure 2.1D). The level of Ntg1 sumoylation in the *siz2Δ* cells was largely unchanged (3.2%) as compared to wildtype. In the *siz1Δsiz2Δ* double mutant cells, we could not detect Ntg1 sumoylation (Figure 2.1C, D). These results demonstrate that Siz1 is the primary E3 ligase responsible for hydrogen peroxide–induced sumoylation of Ntg1 while Siz2 could play a minor role in Ntg1 sumoylation.

We next examined Ntg1 sumoylation in cells defective for the SUMO proteases, Ulp1 and Ulp2. The *ULP1* gene is essential so we employed a temperature sensitive mutant, *ulp1-1* (*ulp1-ts*) (61), and shifted cells to 37°C to inactivate Ulp1. The levels of hydrogen peroxide– induced Ntg1 sumoylation in the *ulp1-ts* mutant were significantly higher than in the wildtype control cells. The *ulp1-ts* mutant cells displayed 10.9% monosumoylated Ntg1, contrasting with 4.7% monosumoylated Ntg1 detected in the wildtype cells (Figure 2.1C, D). In contrast to the *ulp1-ts* cells, *ulp2Δ* cells exhibit no detectable change in sumoylation in response to hydrogen peroxide exposure when compared to wildtype (Figure 2.1D). These results suggest that Ulp1 serves as the primary de-sumoylase for Ntg1.

Sumoylation is a dynamic process where only a very small percent of sumoylated product is present at any given time (35). In fact, in wildtype cells, without exogenous exposure to reactive oxygen species (ROS) or oxidative stress, we do not detect modification of Ntg1 (Figure 2.1A).

To determine whether Ntg1 is modified only in response to hydrogen peroxide or is endogenously modified at low levels, we examined Ntg1 sumoylation in the *ulp1-ts* cells and the *ulp2Δ* mutant cells in the absence of any treatment. Loss of Ulp1 function resulted in a dramatic increase in both monosumoylated and multi-modified Ntg1 compared to wildtype (Figure 2.1E, F). In contrast, loss of Ulp2 had no impact on Ntg1 sumoylation levels. These data indicate that Ntg1 can be sumoylated in the absence of exogenous stress.

2.4.2 Identification of Ntg1 sumoylation sites

Sumoylation occurs on lysine residues, typically within SUMO consensus sequences (62, 63). More than two-thirds of known SUMO substrates contain at least one consensus sumoylation motif Ψ-K-x-D/E (where Ψ is a hydrophobic residue, K is the lysine conjugated to SUMO, x is any amino acid, and D/E is an acidic residue) (62, 63). We used freely available search engines to identify predicted sumoylation sites in both NTHL1 and Ntg1 (Figure 2.3A, B). Prediction software identified multiple candidate sumoylation sites in NTHL1 (Figure 2.3A). To identify candidate sumoylation sites in Ntg1, we used a combination of five SUMO prediction programs: SUMOsp 1.0 (64), SUMOsp 2.0/GPS-SUMO (65, 66), SUMOplot (http://www.abgent.com/sumoplot), SUMOpre (67), and PCI-SUMO (68) This analysis identified five putative consensus sumoylation sites (K20, K38, K376, K388, K396) within Ntg1 (Figure 2.2A and Figure 2.3B). Five putative non-consensus sumoylation sites were also identified (Figure 2.2A and Figure 2.3B). Consensus and non-consensus motifs of identified putative sumoylation sites and prediction scores are shown in Figure 2.3B.

We initially tested for SUMO modification within these sites on Ntg1 via mass spectrometry. However, when we analyzed the bacterially expressed Ntg1 through mass

spectrometry the peptides containing the putative SUMO modification sites were not detected. As an alternative approach, to determine which of these putative sites are sumoylated and to generate a form of Ntg1 that cannot be sumoylated, we performed site-directed mutagenesis to create conservative amino acid substitutions of the ten putative sumoylation site lysines to arginines. These substitutions were made in order of predicted site strength for all single sites (Figure 2.3C). In total, we created 25 single and combination lysine to arginine substitutions beginning with a single substitution and proceeding with double, triple, etc. substitutions (Figure 2.3C). We then analyzed the sumoylation status of all the resulting variants of Ntg1 in response to hydrogen peroxide.

The single lysine to arginine substitutions were tested first and the results showed that all Ntg1 variants containing single lysine to arginine substitutions can still be sumoylated (Figure 2.2B). Quantification of the single substitution data showed that one substitution examined (K396R) results in a detectable decrease in the amount of monosumoylated Ntg1, suggesting that K396 could be the primary site of monosumoylation (Figure 2.2C). The finding that no single lysine to arginine substitution leads to a complete loss of SUMO modification supports our earlier results suggesting that Ntg1 is sumoylated at multiple lysines simultaneously. Thus, multiple substitutions are required to produce a variant that cannot be sumoylated. Single substitution of the five putative non-consensus sumoylation sites (K157, 194, 255, 359, 364) did not alter levels of Ntg1 sumoylation, indicating that these sites are not essential for Ntg1 sumoylation. Next, we tested a series of combinations of lysine to arginine Ntg1 variants for sumoylation. The double and triple mutant proteins involving the N- and C-termini, Ntg1$_{K20,38R}$-TAP and Ntg1$_{K20,38,376R}$-TAP, can both still be sumoylated (Figure 2.2B, C). The quadruple mutant protein, Ntg1$_{K20,38,376,388R}$-TAP, shows only a single sumoylated species; while an additional K396 to arginine substitution,

leads to the complete loss of all detectable SUMO-modification of Ntg1 (ntg1ΔSUMO). For the collection of variants, changes in the levels of Ntg1 sumoylation were quantified and are presented in Table 3. These results demonstrate that Ntg1 is sumoylated at any of five consensus sumoylation sites and that all five sites must be simultaneously changed to arginine to generate an Ntg1 variant that cannot be modified by SUMO. Figure 2.3C shows all of the combinations of Ntg1 variants generated and summarizes the total sumoylation loss. Thus, we have identified the five lysine residues within Ntg1 that can be sumoylated.

The lysines within Ntg1 that can be sumoylated reside in the N- and C-terminal domains that are specific to the eukaryotic enzyme and outside of the 307 residues which comprise the evolutionarily conserved catalytic core with homology to the bacterial Endonuclease III protein (69). To provide insight into the location of these lysines within the three-dimensional structure of Ntg1, we generated a homology model of *S. cerevisiae* Ntg1 (Figure 2.4A) using the Protein Homology/analogY Recognition Engine version 2.0 (Phyre2) (70). The predicted model (Figure 2.4A) is based on the structure of the *E. coli* Endonuclease III protein (PDB ID: 2ABK). Tan regions display the high confidence (90%) homology mapping of the region of Ntg1 (amino acids 95–335) with homology to Endonuclease III (Figure 2.2A). The magenta and green regions correspond to the N-terminal (amino acids 1– 94) and C-terminal (amino acids 335–399) domains, respectively (Figure 2.2A). Although the N- and C-terminal domains of Ntg1 do not align to Endonuclease III, they are modeled based on homology to other endonucleases. The N-terminal domain aligns to the restriction endonuclease BsaW*I* (PDB ID: 4ZSF) and the C-terminal domain aligns to the endonuclease Bgl*II* (PDB ID: 1DFM). Based on our homology model (Figure 2.4A), the five lysines that we defined as SUMO modification sites are all surface exposed, consistent with being accessible for modification.

As sumoylation influences DNA binding and turnover of TDG (43–45), we analyzed the proximity of the sumoylation sites to the DNA binding and catalytic centers in our Ntg1 model. The structure of sumoylated-TDG (PDB ID: 1WYW) shows the close proximity of the SUMO modification to the DNA binding and catalytic site of TDG (71), illustrating why sumoylation of TGD might influence TGD catalysis. To assess whether sumoylation could impact DNA binding or catalysis by Ntg1, we superimposed our model of Ntg1 with the structure of Endonuclease III (Figure 2.4A). Previous work implicated K120 and D138 in catalysis and K191 in DNA binding of Endonuclease III (Figure 2.4A) (69). We identified the analogous amino acids in our model of Ntg1 (Figure 2.4A, B). Loop residues in Endo III corresponding to residues 314–318 in Ntg1 are important for DNA binding (Figure 2.4B). Both the catalytic residues and the DNA binding loop in Ntg1 are distant from the sumoylated lysines and extend from the opposite face of the protein. This analysis suggests that sumoylation is unlikely to directly influence Ntg1-mediated catalysis.

2.4.3 *In vitro* functional analysis of the nonsumoylatable Ntg1 variant, ntg1^{K20,38,376,388,396R} (ntg1(K->R)$_5$)

Although changing a lysine residue to an arginine residue conserves the charge and size of the amino acid, such modest changes could induce a conformational change potentially impacting function. To address whether the conservative substitution of the five lysines that constitute the Ntg1 sumoylation sites (K20, K38, K376, K388, K396) impacts the catalytic activity of Ntg1 *in vitro*, we employed an *in vitro* oligonucleotide cleavage assay to compare the enzymatic activity of wildtype Ntg1 to ntg1^{K20,38,376,388,396R}, which we designate ntg1(K->R)$_5$. The oligonucleotide substrate contains dihydrouracil (DHU) which is an Ntg1 substrate (72). As a control, we employed a catalytically inactive Ntg1 (ntg1Δcat) variant created by changing the catalytic lysine

at position 243 to glutamine (73). We incubated purified recombinant His6-Ntg1 variants with the oligonucleotide containing the Ntg1 substrate and detected Ntg1 enzymatic activity as cleavage of the oligonucleotide at the position of the DHU (26). As shown in the cleavage assay presented in Figure 2.5B, His6-ntg1(K->R)$_5$ shows enzymatic activity comparable to wildtype His6-Ntg1 whereas a catalytically inactive form of Ntg1, His6-ntg1Δcat, did not cleave the substrate. We quantitated the results of three independent cleavage experiments (Figure 2.5C). These results confirm that there is no difference in the activity of ntg1(K->R)$_5$ compared to wildtype Ntg1 in this assay. These results indicate that changing the five SUMO modification sites from lysine to arginine does not alter the enzymatic activity of Ntg1.

2.4.4 Functional analysis of Ntg1 *in vivo*

Previous studies showed that Ntg1 is SUMO–modified in response to treatment with hydrogen peroxide (24). To assess whether other types of DNA damage can induce Ntg1 sumoylation, cells were exposed to methyl methanesulfonate (MMS), which induces alkylating DNA damage (74, 75). As shown in Figure 2.6A, Ntg1 is sumoylated in response to treatment with MMS. We exploited this observation to examine how cells that express an Ntg1 variant that cannot be modified by SUMO (ntg1$^{K20,38,\ 376,388,396R}$), which we designate ntg1ΔSUMO, respond to DNA damage. For these experiments, Ntg1 or ntg1ΔSUMO was expressed either in base excision and nucleotide excision repair (NER)- proficient wildtype cells (WT) or repair-deficient (*ntg1Δntg2Δapn1Δrad1Δ*) cells, which lack BER and NER (B-/N-) (19, 76). Cells were exposed to MMS and the growth characteristics of these cells expressing wildtype Ntg1 were compared to those expressing ntg1ΔSUMO. We then examined growth in the absence or presence of MMS (Figure 2.6B). Growth was analyzed at days 2 and 4 following serial dilution and spotting on

plates. As expected (76), the repair proficient (WT) cells grew well under all conditions tested, regardless of which Ntg1 variant was expressed. In contrast, the repair-deficient cells display slow growth in the presence of MMS even with wildtype *NTG1*. The repair-deficient ntg1Δ cells were extremely sensitive to MMS (Figure 2.6B). Surprisingly, the *ntg1ΔSUMO* cells were less sensitive to MMS compared to cells with wildtype NTG1. This result suggests that sumoylation of Ntg1 could be important for coordinating DNA repair with cell cycle progression or DNA damage response.

To further examine the growth of the *ntg1ΔSUMO* cells following treatment with MMS, growth curves were generated for wildtype, repair-deficient cells, Ntg1 in repair-deficient cells, and ntg1ΔSUMO in repair-deficient cells grown in YPD with and without MMS (Figure 2.6C, D). The results indicate that repair-deficient cells that express ntg1ΔSUMO emerge from lag-phase earlier than repair-deficient cells expressing wildtype Ntg1 (Figure 2.6D). As expected, repair-deficient cells expressing either Ntg1 or ntg1ΔSUMO grew equally well in the absence of MMS (Figure 2.6C). These data further suggest that the sumoylation of Ntg1 plays a role in coordinating the growth arrest that occurs in response to DNA damage.

2.5 Discussion

We report here that SUMO modification is a conserved post-translational modification of *S. cerevisiae* Ntg1 and the human orthologue, NTHL1. In *S. cerevisiae*, we identified the two SUMO ligases, Siz1 and Siz2, and the desumoylase, Ulp1, critical for reversible regulation of this modification. We mapped the sites of SUMO modification in Ntg1 and created an Ntg1 that cannot be SUMO modified. Our preliminary analysis of this non-sumoylatable form of Ntg1 reveals that SUMO modification may be important for proper cellular response to DNA damage.

We identified Ulp1 as the primary desumoylase for Ntg1 with little impact of Ulp2. As the primary role of Ulp2 is to remove SUMO from poly(SUMO) chains (55), and we detect no change in SUMO modification of Ntg1 in *ulp2Δ* mutant cells (Figure 2.1C, D), we speculate that Ntg1 could be modified by multiple independent SUMOs rather than a single chain of multiple SUMOs. This model is consistent with our finding that five different lysine residues in Ntg1 can be modified by SUMO. While we cannot rule out the possibility that the Ntg1-TAP used in this study is also modified by other post-translational modifications, the band shifts are consistent with the molecular size and charge of multiple SUMO molecules. Consistent with a possible role for additional post-translational modifications, mass spec analysis reveals that Ntg1 serine 71 is phosphorylated (77). Regardless, the data presented here in combination with our previous publication (24) show that Ntg1 is SUMO modified by at least one SUMO molecule and that there are at least five lysines on Ntg1 that can be SUMO modified. These SUMO molecules could coordinate other post-translational modifications.

Our data show that both hydrogen peroxide and MMS can induce sumoylation of Ntg1 (24). Like hydrogen peroxide, treatment with MMS can cause oxidative stress and generate ROS (78). Therefore, we cannot yet clearly distinguish whether hydrogen peroxide and MMS trigger sumoylation of Ntg1 through the same or distinct mechanisms. Further work will be required to determine how the sumoylation machinery responds to DNA damage and/or oxidative stress. Regulation could occur through activation of SUMO E3 ligases or through inhibition of the Ulp1 desumoylase. Further analysis will be required to dissect this mechanism.

Sumoylation can influence numerous functions of a protein including catalytic activity, localization, stability, and/or protein-protein interactions (79). As sumoylation plays a role in regulating the binding capabilities of TDG by modulating the interaction with DNA, sumoylation

could also impact the DNA binding ability of Ntg1. However, based on homology modeling and mapping of the sumoylation sites, our model suggests that sumoylation at any of the five sites we identified likely does not directly interfere with the DNA binding to Ntg1. Consistent with this result, none of the Ntg1 lysine to arginine substitutions altered catalytic activity of the recombinant protein *in vitro* (Figure 2.5). With respect to localization, Ntg1 sumoylation at K20 and K38 on Ntg1 are within or just adjacent to the consensus organelle targeting localization sequences. The N-terminus of Ntg1 contains a mitochondrial targeting sequence (MTS) at amino acids 1–26 and a classical bipartite nuclear localization signal (cNLS) at amino acids 14–17 and 31–37 (51). The proximity of the sumoylation sites, specifically K20 and K38, to these localization signals suggests a potential role for sumoylation in regulating subcellular localization of Ntg1. In fact, our previous biochemical fractionation studies showed that sumoylated Ntg1 is detected only in the nucleus (24). Consistent with a conserved regulatory model, human NTHL1 also contains putative SUMO sites at K56 and K60 that overlap a predicted cNLS at amino acids 56–60 (Figure 2.3A). Another possible function of sumoylation is to modulate protein-protein interactions (79). Little is known about the interacting partners of Ntg1. One high-throughput yeast two-hybrid study identified two DNA damage response proteins, Rad59 and Rfc2, as physical interactors of Ntg1 (80). Rad59 is involved in double-strand break repair (81), and Rfc2 is part of the ATPase clamp loader for the proliferating cell nuclear antigen (PCNA) processivity factor for DNA polymerases (82). As both of these proteins are implicated in DNA damage response, they could mediate crosstalk between BER and the DNA damage response pathway. A critical next step in understanding the functional impact of sumoylation on Ntg1 is to identify SUMO-dependent interacting proteins.

Our data (Figure 2.6B, C, D) show cells expressing ntg1ΔSUMO display more rapid growth compared to cells expressing wildtype Ntg1 in a DNA repair deficient background in

response to MMS. Alkylation damage induced by MMS can be mutagenic and lead to cytotoxic blockage of replication forks (75, 83, 84). One possibility is that sumoylation of Ntg1 is required for proper checkpoint activation or maintenance. Cell cycle checkpoints are activated by sensor proteins, such as Rad9 (85), detecting an increase in DNA damage and initiating a signal cascade that ultimately leads to activation of Rad53, the protein kinase responsible for cell cycle arrest (86, 87). Activation of Rad53 is critical for stabilization of replication forks and activating the DNA repair pathway (81). Interestingly, improper activation of Rad53 results in an increased resistance to MMS via engagement of translesion synthesis (TLS) (88). Further investigation of this potential connection between the DNA checkpoint protein, Rad53, and the BER protein, Ntg1, could reveal a novel DNA damage response activator.

A number of studies have identified roles for SUMO in modulating DNA repair (89–98). The work presented here suggests SUMO-mediated regulation could extend to the evolutionarily conserved BER pathway. Indeed, regulation of the initial step of the BER pathway could be crucial to ensure genome integrity.

2.6 Figures and tables

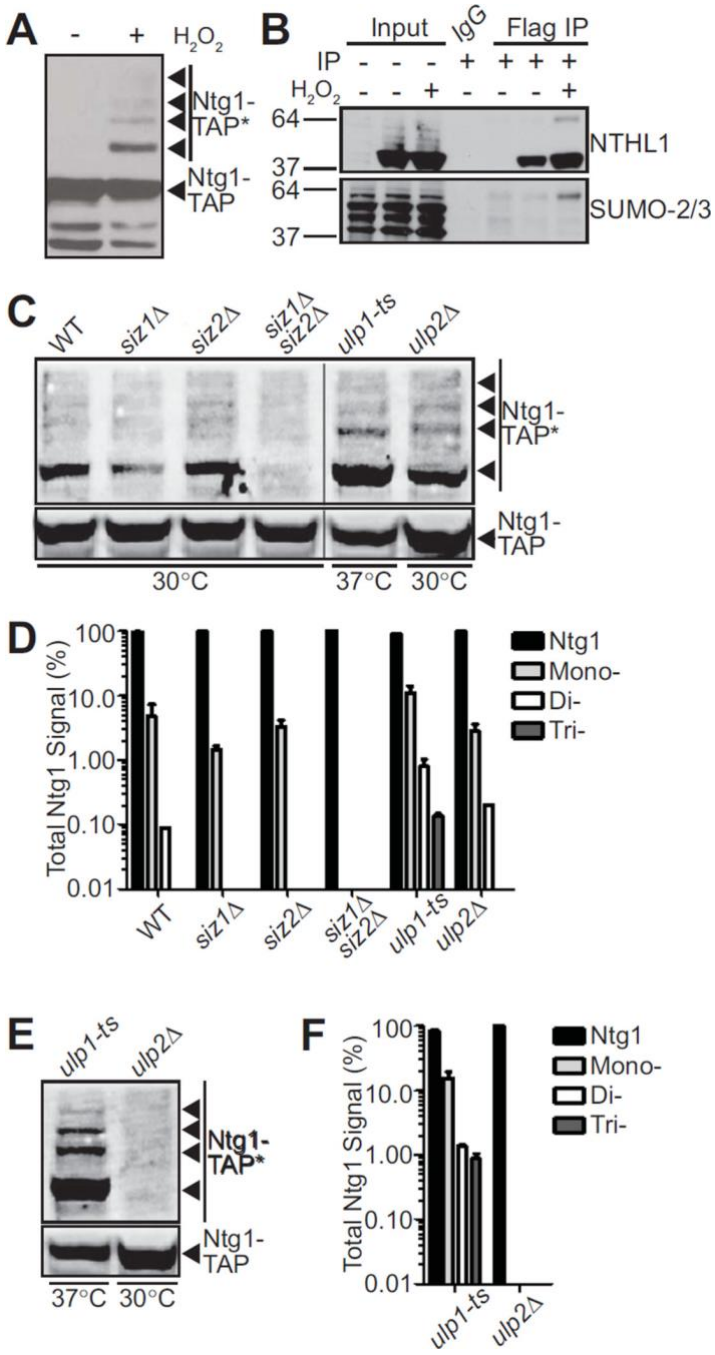


Figure 2.1: Sumoylation of Ntg1 is conserved and mediated by Siz1/2

A. Wildtype *S. cerevisiae* cells expressing Ntg1-TAP were exposed to 0 (−) or 20 mM (+) H2O2 for 1 hour at 30°C. Cells were pelleted, lysed, and immunoblotted to detect TAP- tagged Ntg1. Bands corresponding to post-translationally modified Ntg1 including SUMO- modified

Ntg1 (24) are indicated by Ntg1-TAP*. B. Colon adenocarcinoma cells (HT29) were transfected with NTHL1-Flag or empty Flag vector and treated with 0 (−) or 125 μM (+) H2O2 for 15 minutes at 37°C. Cells were lysed, immunoprecipitated with Flag antibodies and both the Input and Flag IP fractions were subjected to immunoblotting. An IgG bead alone immunoprecipitation was included as a control. The blot was probed with NTHL1 and SUMO-2/3 antibodies as indicated. C. Wildtype (WT), *siz1Δ*, *siz2Δ*, *siz1Δsiz2Δ*, *ulp1-ts*, or *ulp2Δ* cells were transformed with a plasmid expressing Ntg1-TAP. Cells were (C) exposed to 20 mM hydrogen peroxide or (E) not treated. Cells were incubated at 30°C except ulp1-ts cells which were shifted to the non-permissive temperature of 37°C. Each sample was lysed, immunoblotted, and bands were quantified. Nonadjacent lanes in the same image are separated by a black line. D. The data from (C) were quantitated. The total amount of Ntg1- TAP including unmodified and modified Ntg1-TAP was set to 100% (Ntg1) and the fraction of signal present in bands (Total Ntg1 Signal %) corresponding to the size consistent with Mono-, Di-, and Tri-sumoylation is plotted on a log scale. Results shown are the average of two independent experiments. Error bars represent SEM. E. To examine sumoylation of Ntg1 in the absence of oxidative damage, *ulp1-ts* and *ulp2Δ* cells expressing Ntg1-TAP were analyzed to detect any modified Ntg1 species (Ntg1-TAP*). F. The data from (E) were quantitated. The total amount of Ntg1-TAP including unmodified and modified Ntg1-TAP was set to 100% (Ntg1) and the fraction of signal present in bands (Total Ntg1 Signal %) corresponding to the size consistent with Mono-, Di-, and Tri-sumoylation is plotted on a log scale. Results shown are the average of two independent experiments. Error bars represent SEM.

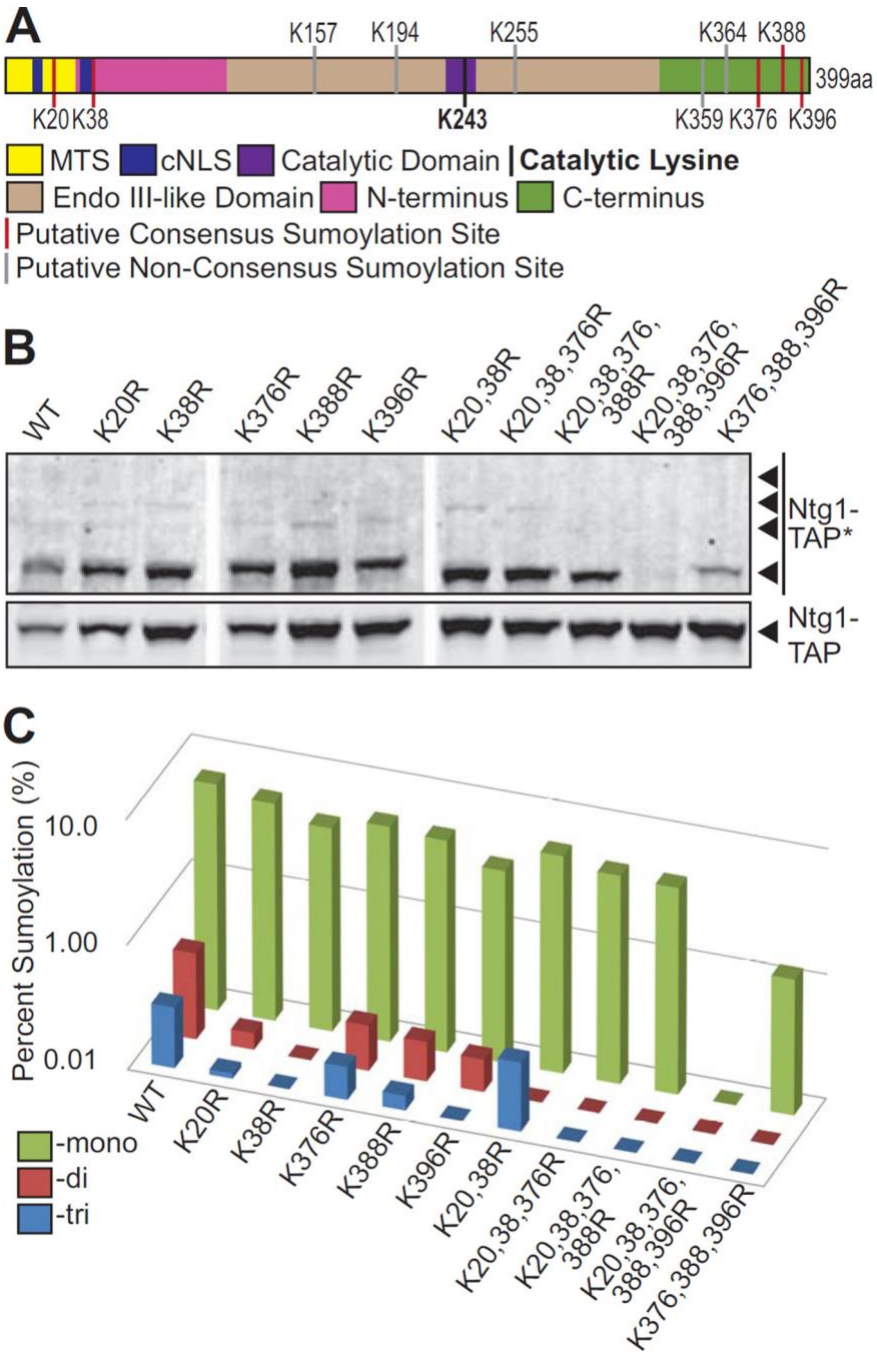


Figure 2.2: Identification of sumoylation sites in Ntg1

A. A domain schematic of Ntg1 is shown with the following functional motifs/domains indicated: The Mitochondrial Targeting Sequence (MTS) in yellow, the classical Nuclear Localization Signal (cNLS) in dark blue, the Catalytic Domain in purple. The Catalytic Lysine, K243, is depicted as a black bar. The central region of Ntg1 that is homologous to *E. coli*

Endonuclease III is shown in tan (amino acids 95–335) while the non-conserved N- and C-terminal domains are indicted in magenta (amino acids 1–94) and green (amino acids 336–399), respectively. Putative Consensus Sumoylation Sites are shown as red bars and Putative Non-Consensus Sumoylation Sites are shown as grey bars. B. A series of Ntg1 variants with candidate SUMO modification sites altered from lysine to arginine were generated and expressed in temperature sensitive *ulp1* cells. Cells were treated with 20 mM hydrogen peroxide for 1 hour at 30°C, lysed, and immunoblotted to detected Ntg1-TAP and modified Ntg1-TAP (Ntg1-TAP*). Nonadjacent lanes in the same image are separated by white space. C. Results from (B) were quantitated. For each Ntg1 variant, the percent of total Ntg1-TAP signal present in the band corresponding to the size of Mono-, Di-, and Tri- sumoylation (indicated as Percent Sumoylation) is plotted on a log scale.

A

Human NTHL1							
Software	Ψ-K-x-D/E	GPS-SUMO		PCI-SUMO		SUMOplot	Combined
Non-consensus lysine	Motif	Score	P-value	Confidence	Motif?	Score	
56	VK$_{56}$RP	1.795	0.81	0.2403	No	0.82	0.51
60	RK$_{60}$AQ	0.751	0.882	0.1582	No		0.38
75	EK$_{75}$GE	3.818	0.085			0.5	0.75
83	LK$_{83}$VP	1.985	0.411	0.7957	No	0.8	0.52
105	NK$_{105}$KD	2.408	0.306				0.67
126	PK$_{126}$VR	2.07	0.628				0.33
257	TK$_{257}$SP	2.294	0.615				0.33

B

Saccharomyces cerevisiae Ntg1							
Software	Ψ-K-x-D/E	SUMOsp 2.0		PCI-SUMO		SUMOplot	Combined
Lysine	Motif	Score	Confidence	Sumo?	Confidence	Score	
Consensus							
20	VK$_{20}$TE	3.199	High	Yes	0.8014	0.93	**0.94**
38	IK$_{38}$QE	5.825	High	Yes	0.8771	0.94	**0.96**
376	VK$_{376}$HE	2.711	High	Yes	0.2574	0.93	**0.85**
388	VK$_{388}$LE	2.801	High	Yes	0.3171	0.93	**0.86**
396	VK$_{396}$VE	2.697	High			0.93	**0.97**
Non-consensus							
157	TK$_{157}$DE	2.779	Medium	No	0.2127		**0.53**
194	TK$_{194}$LD	2.75	Medium	Yes	0.6845		**0.75**
255	GK$_{255}$IE			Yes	0.538	0.67	**0.48**
359	VK$_{359}$YL			Yes	0.5946		**0.40**
364	GK$_{364}$RE			No	0.2233	0.67	**0.35**

C

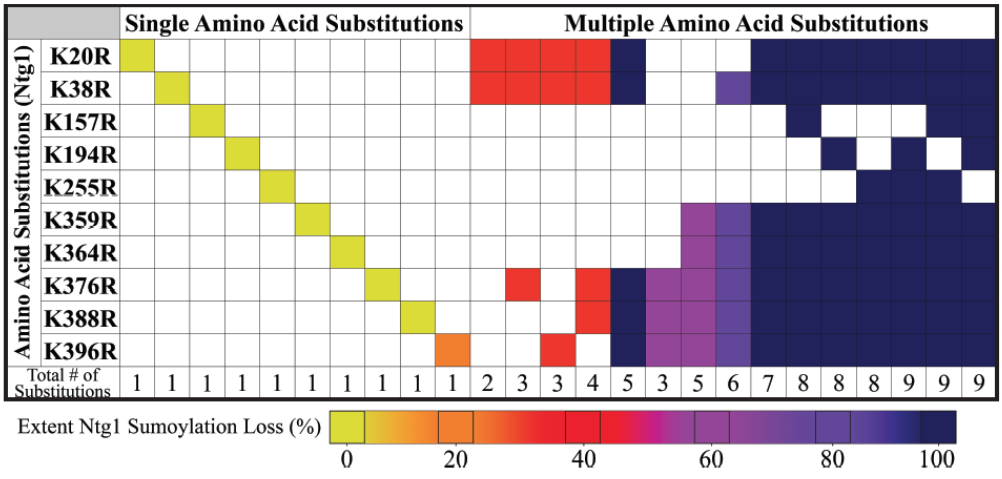


Figure 2.3: Putative SUMO site analysis in humans and *S. cerevisiae* and graph of sumoylation loss by amino acid substitution. A. A lists the putative sumoylation sites on human NTHL1 as K56, 60, 74, 83, 105, 126, and 257. This table also lists the sequence motif and software

utilized, and the score or confidence of the potential for this site to be sumoylated. The combined across the three SUMO prediction software are also shown. B. A list of the putative consensus sumoylation sites on *S. cerevisiae* Ntg1 as K20, 38, 376, 388, and 396. The putative non-consensus sites are K157, 194, 255, 359, and 364. This table also lists the sequence motif and software utilized, and the score or confidence of the potential for this site to be sumoylated. The combined across the three SUMO prediction software are also shown. C. This image depicts the systematic strategy for the putative SUMO site variants in singles, doubles, triples and so on. The color gradient indicates the extent of Ntg1 sumoylation loss. Yellow indicates 0% loss of sumoylation, red/purple indicate 50% loss of sumoylation, and navy indicates 100% loss of sumoylation.

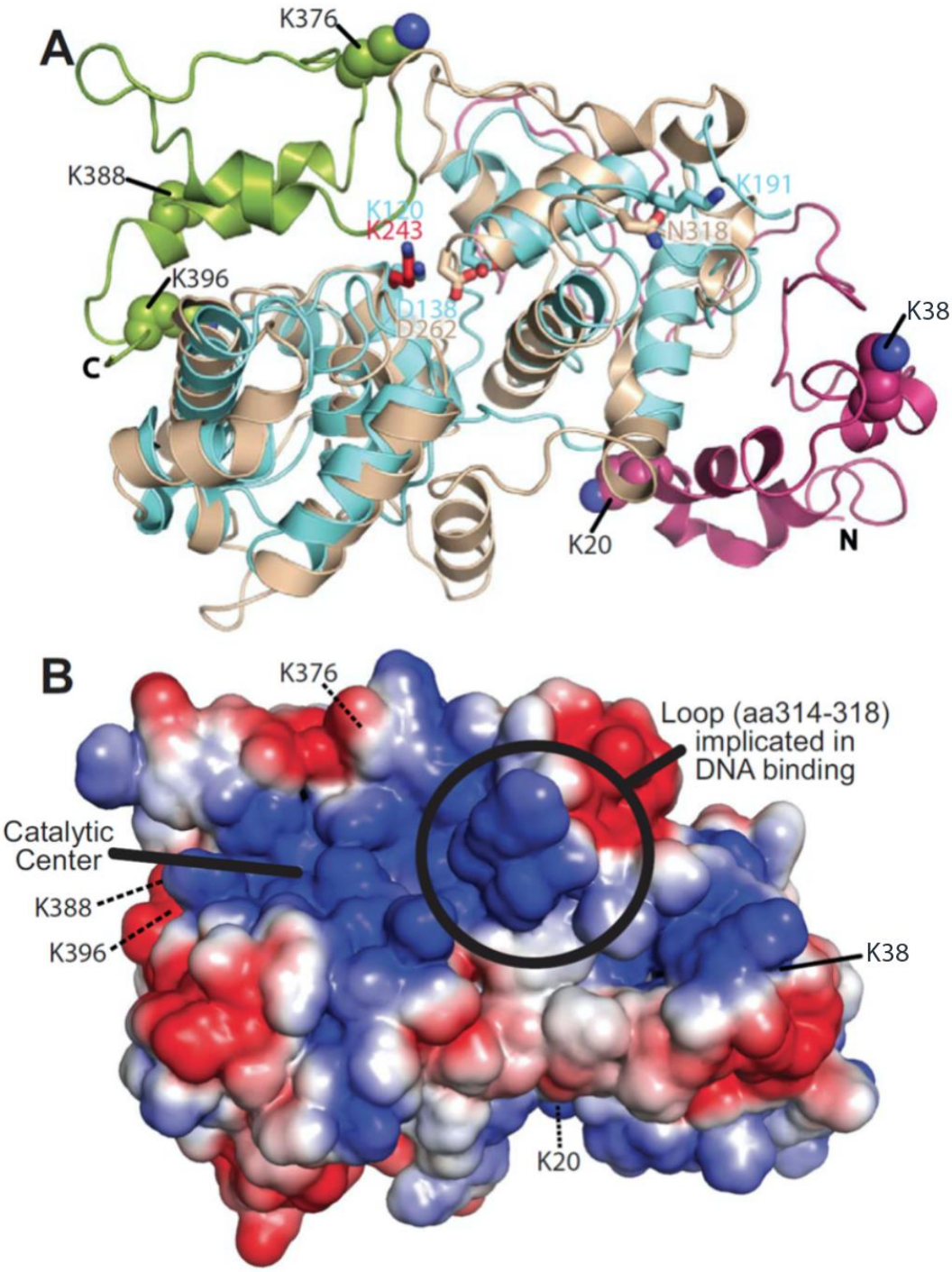


Figure 2.4: Homology model of Ntg1

A. A homology model of Ntg1 shown as a ribbon diagram was generated as described in Materials and Methods. The model is overlaid on the *E. coli* Ntg1 homologue, Endonuclease III, structure (cyan, PDB ID: 2ABK). The Ntg1 catalytic domain (amino acids 95–335; tan), N-

terminal domain (amino acids 1–94; magenta), C-terminal domain (amino acids 336–399; green), catalytic amino acid of Ntg1 (K243, red) and Endonuclease III (K120, blue) are shown in addition to Endonuclease III amino acids D138, important for catalysis, and K191, implicated in DNA binding (69), and the corresponding amino acids in Ntg1 (D262 and N318, respectively). The five consensus sumoylation sites (K20, K38, K376, K388, and K396) are shown as balls and indicated by the labeling. B. An electrostatic model of Ntg1 is shown based on the homology model. Positive and negative residues are colored in blue and red, respectively. White indicates neutral residues. The loop containing residues 314–318, indicated by a circle, has been implicated in DNA binding by Endo III (69). The catalytic center is indicated by a bold black line and the five consensus sumoylation sites are labeled and indicated by black lines. Residues 20, 376, 388, and 396 are located on the back face of the model and are indicated by black dotted lines.

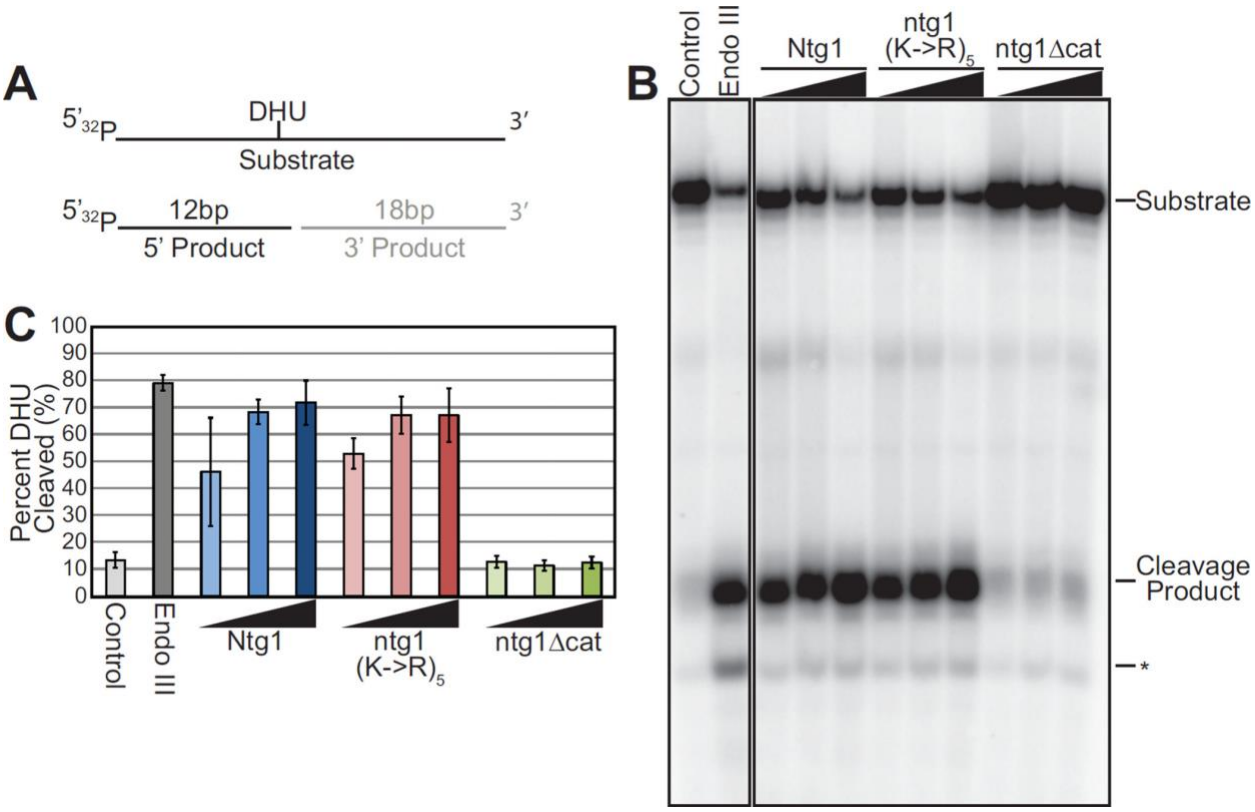


Figure 2.5: Functional analysis of Ntg1 variant

A. A schematic of the substrate employed for the *in vitro* cleavage assay, which contains dihydrouracil (DHU) embedded in a 31mer oligo, illustrating the substrate and expected products of the cleavage reaction is shown. B. Recombinant *E. coli* Endonuclease III (Endo III), and Ntg1 variants, His6-Ntg1, His6-ntg1(K->R)5, catalytically inactive ntg1 (His6-ntg1Δcat), were employed for the *in vitro* cleavage assay. Increasing amounts of recombinant protein (5–50 ng) were added to radioactively-labeled substrate. Oligonucleotide Cleavage Products were electrophoresed and subjected to phosphorimager analysis. The Control lane shows the substrate with no added protein. The positive control is addition of 50 ng of *E. coli* Endo III. The position of the labeled product generated by cleavage (Cleavage Product) is indicated. Random degradation product is indicated by an asterisk (*). Nonadjacent lanes in the same image are separated by black lines. Results shown in (B) are representative of three independent experiments. C. Quantification

of Cleavage Product generated for each Ntg1 variant from three independent experiments. Results are shown as Percent DHU Cleaved. Error bars represent standard deviation in the data.

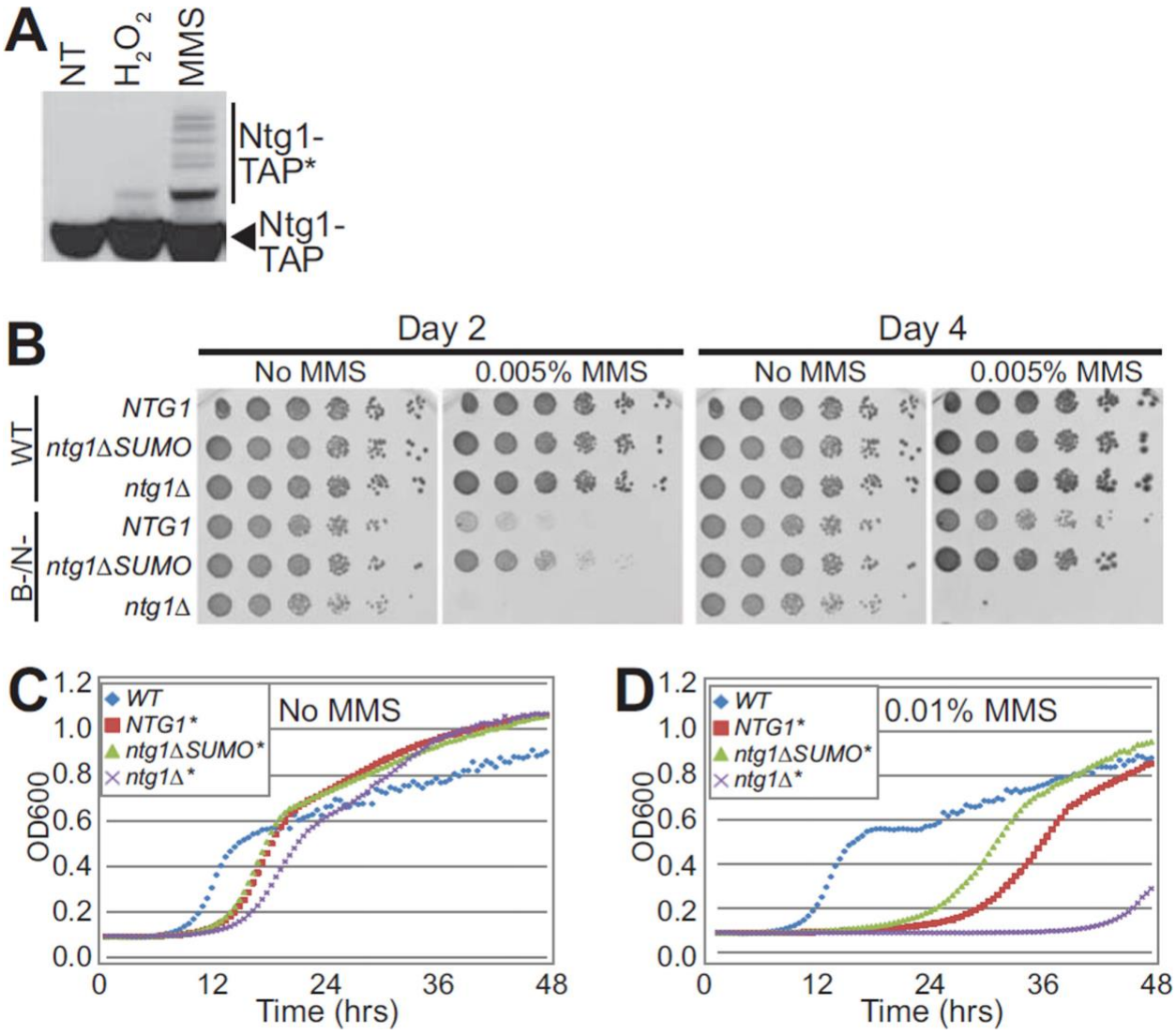


Figure 2.6: Functional analysis of ntg1ΔSUMO in DNA damage pathways

A. Wildtype cells expressing Ntg1-TAP were exposed to hydrogen peroxide (H2O2), methyl methanesulfonate (MMS), or were not treated (NT) and lysed. Lysate was subjected to immunoblotting to detect Ntg1-TAP and modified forms of Ntg1-TAP (Ntg1-TAP*). B. Cells with either a full complement of wildtype (WT) DNA repair pathways or deficient in both base excision repair and nucleotide excision repair (B-/N-) were employed. As described in Materials and Methods, the genotype for B-/N- cells (DSC0369) is *ntg1Δntg2Δapn1Δrad1Δ*. Both the WT and B-/N- cells were engineered to express ntg1ΔSUMO and compared to cells expressing NTG1 or

lacking Ntg1 (*ntg1Δ*). Cultures were 5-fold serially diluted and spotted onto rich media or rich media containing 0.005% MMS and incubated at 30°C for 4 days. Pictures were taken at Day 2 and Day 4. C/D. The same samples as shown in (B) with either intact DNA repair pathways (WT) (blue diamond) or deficient in base excision repair and nucleotide excision repair (B-/N-), denoted by an *, contain wildtype *NTG1* (red square), or ntg1ΔSUMO (green triangle), or lack Ntg1 (ntg1Δ) (purple X) at the endogenous NTG1 locus. The genotype for B-/N- is *ntg1Δntg2Δapn1Δrad1Δ* (DSC0369). Cells were grown in liquid culture with No MMS (C) or with 0.01% MMS (D) for 48 hours. OD600 readings were taken every 30 minutes and plotted vs time. Results shown in (B, C, and D) are representative of at least three independent experiments.

Table 2.1: Strains and plasmids

Strain or Plasmid	Description	References
DSC0295	MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0; Tet-Off C-terminally TAP-tagged Ntg1	(24)
YSC1178-7499106 (DSC0297)	MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0; C- terminally TAP-tagged Ntg1	Open Biosystems
BY4147 (DSC0313)	MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0	Open Biosystems
DSC0470	MATa ntg1::hphMX4, his7-1, lys2Δ5'::LEU-lys2Δ3', ade5-1, trp1-289, ura3-52	This study
EJY341 (DSC0527)	MATa trp1-Δ1 ura3-52 his3-Δ200 leu2-3,112 lys2-801 [cir°]	(99)
EJY342 (DSC0528)	MATa trp1-Δ1 ura3-52 his3-Δ200 leu2-3,112 lys2-801 siz1Δ::LEU2 [cir°]	(99)
EJY343 (DSC0529)	MATa trp1-Δ1 ura3-52 his3-Δ200 leu2-3,112 lys2-801 siz2Δ::TRP1 [cir°]	(99)
EJY344 (DSC0530)	MATa trp1-Δ1 ura3-52 his3-Δ200 leu2-3,112 lys2-801 siz1Δ::LEU2 siz2Δ::TRP1 [cir°]	(99)
MHY1488 (DSC0534)	MATa ulp1Δ::HIS3 LEU2::ulp1-333	(58)
EJY447 (DSC0535)	MATa trp1-Δ1 ura3-52 his3-Δ200 leu2-3,112 lys2-801 ulp2Δ::kanMX [cir°]	(60)
GBY5 (DSC0536)	MATa smt3-allR::TRP1	(55)
DSC0537	MATa ntg1::hphMX4, his7-1, lys2Δ5'::LEU- lys2Δ3', ade5-1, trp1-289, ura3-52, pD0436	This study
DSC0538	MATa trp1-Δ1 ura3-52 his3-Δ200 leu2-3,112 lys2-801 [cir°], pD0436	This study
DSC0539	MATa smt3-allR::TRP1, pD0436	This study
DSC0540	MATa ulp1Δ::HIS3 LEU2::ulp1-333, pD0436	This study
hDNP19	MATa/MATα rad1::kanMX/RAD1 ntg1::hphMX4/NTG1 ntg2::BSD/NTG2 apn1::TRP1/APN1 DSF1::URA3/DSF1 his7-1/his7-1 lys2Δ5′::LEU-lys2Δ3′/lys2Δ5′::LEU- lys2Δ3′ ade5-1/ade5-1 trp1-289/trp1-289 ura3- 52/ura3-52	(19)
DSC0367	MATa his7-1 lys2Δ5′::LEU-lys2Δ3′ ade5-1 trp1- 289 ura3-52	(51)
DSC0369	MATa ntg1::hphMX4 rad1::kanMX ntg2::BSD apn1::TRP1 his7-1 lys2Δ5′::LEU-lys2Δ3′ ade5-1 trp1-289 ura3-52	(51)
DSC0371	MATa rad1::kanMX ntg2::BSD apn1::TRP1 his7-1 lys2Δ5′::LEU-lys2Δ3′ ade5-1 trp1-289 ura3-52	(51)

Strain or Plasmid	Description	References
DSC0561	MATa ntg1k20,38,376,388,396R rad1::kanMX ntg2::BSD apn1::TRP1 DSF1::URA3 his7-1 lys2Δ5′::LEU-lys2Δ3′ ade5-1 trp1-289 ura3-52	This study
DSC0549	MATa ntg1$_{k396R}$ rad1::kanMX ntg2::BSD apn1::TRP1 DSF1::URA3 his7-1 lys2Δ5′::LEU- lys2Δ3′ ade5-1 trp1-289 ura3-52	This study
DSC0551	MATa ntg1$_{k20,38R}$ rad1::kanMX ntg2::BSD apn1::TRP1 DSF1::URA3 his7-1 lys2Δ5′::LEU- lys2Δ3′ ade5-1 trp1-289 ura3-52	This study
DSC0558	MATa ntg1$_{k20,38,376,388R}$ rad1::kanMX ntg2::BSD apn1::TRP1 DSF1::URA3 his7-1 lys2Δ5′::LEU- lys2Δ3′ ade5-1 trp1-289 ura3-52	This study
DSC0561	MATa ntg1$_{k20,38,376,388,396R}$ rad1::kanMX ntg2::BSD apn1::TRP1 DSF1::URA3 his7-1 lys2Δ5′::LEU-lys2Δ3′ ade5-1 trp1-289 ura3-52	This study
DSC0555	MATa ntg1$_{k376,388,396R}$ rad1::kanMX ntg2::BSD apn1::TRP1 DSF1::URA3 his7-1 lys2Δ5′::LEU- lys2Δ3′ ade5-1 trp1-289 ura3-52	This study
pD0390	pET-15b His6-NTG1	(51)
pD0394	pET -15b His6-NTG1$_{\Delta cat}$	(51)
pD0493	pET -15b His6-NTG1$_{(K->R)5}$	This study
pD0436	Tet-Off NTG1-TAP, CEN, URA3, amp$_R$	This study
pD0437	Tet-Off ntg1$_{K20R}$-TAP, CEN, URA3, amp$_R$	This study
pD0438	Tet-Off ntg1$_{K38R}$-TAP, CEN, URA3, amp$_R$	This study
pD0444	Tet-Off ntg1$_{K376R}$-TAP, CEN, URA3, amp$_R$	This study
pD0445	Tet-Off ntg1$_{K388R}$-TAP, CEN, URA3, amp$_R$	This study
pD0446	Tet-Off ntg1$_{K396R}$-TAP, CEN, URA3, amp$_R$	This study
pD0447	Tet-Off ntg1$_{K20,38R}$-TAP, CEN, URA3, amp$_R$	This study
pD0448	Tet-Off ntg1$_{K20,38,376R}$-TAP, CEN, URA3, amp$_R$	This study
pD0449	Tet-Off ntg1$_{K20,38,396R}$-TAP, CEN, URA3, amp$_R$	This study
pD0450	Tet-Off ntg1$_{K20,38,376,388R}$-TAP, CEN, URA3, amp$_R$	This study
pD0451	Tet-Off ntg1$_{K20,38,376,388,396R}$-TAP, CEN, URA3, amp$_R$	This study
pD0452	Tet-Off ntg1$_{K376,388,396R}$-TAP, CEN, URA3, amp$_R$	This study

Table 2.2: Plasmid construction primers

Primer Purpose	Primer Name	Sequence (5' - 3')
pD0436	tetNtg1Cla- F1	GAATCGATTGCAGTTTCATTTGATGCTCGATGAG
	His-Ntg1Cla-R1	GAATCGATGTATTCTGGGCCTCCATGTCGC
K20R	K20R2-F	CAATTCTGAGGAAAAGACCGCTGGTAAGGACTGAAACTGG
	K20R2-R	CCAGTTTCAGTCCTTACCAGCGGTCTTTTCCTCAGAATTG
K38R	K38R-F	GGACCAAAATCAGACAAGAAGAGGTTGTCCCTCAACCCGTG
	K38R-R	CACGGGTTGAGGGACAACCTCTTCTTGTCTGATTTTGGTCC
K157R	K157R-F	GATGCTATCATCGCAAACAAGAGATGAAGTTACCGCAATGGC
	K157R-R	GCCATTGCGGTAACTTCATCTCTTGTTTGCGATGATAGCATC
K194R	K194R-F	CCGTTTTACAAATCAATGAGACCAGATTAGACGAATTGATTCATTCAG
	K194R-R	CTGAATGAATCAATTCGTCTAATCTGGTCTCATTGATTTGTAAAACGG
K255R	K255R-F	CATTACAAAAGGCATGGGGCAGGATTGAAGGTATCTGCGTTGACG
	K255R-R	CGTCAACGCAGATACCTTCAATCCTGCCCCATGCCTTTTGTAATG
K359R	K359R-F	GCAAAATATCATGAGTTATCCAAAGTGGGTGAGATACCTGGAAGG
	K359R-R	CCTTCCAGGTATCTCACCCACTTTGGATAACTCATGATATTTTGC
K364R	K364R-F	TACCTGGAAGGAAGACGTGAACTGAACGTGGAGGCGG
	K364R-R	CCGCCTCCACGTTCAGTTCACGTCTTCCTTCCAGGTA
K376R	K376R-F	CGTGGAGGCGGAAATCAATGTTAGACACGAGGAGAAAACAG
	K376R-R	CTGTTTTCTCCTCGTGTCTAACATTGATTTCCGCCTCCACG
K388R	K388R-F	CGAGGAGAAAACAGTTGAAGAAACTATGGTCAGACTGGAAAATG
	K388R-R	CATTTTCCAGTCTGACCATAGTTTCTTCAACTGTTTTCTCCTCG
K396R	K396R-F	GGAAAATGATATTTCTGTTAGAGTAGAGGACGGTCGACGG
	K396R-R	CCGTCGACCGTCCTCTACTCTAACAGAAATATCATTTTCC
NTHL1	NTHL1- Flag-F	ACACTGGCGGCCGTTACTAGTGGATCCT
	NTHL1- Flag-R	ACGACTCACTATAGGGAGACCCAAGCTT

Table 2.3: Quantification of mono-, di-, tri- sumoylation of Ntg1 variants

Potential Sumo Modification	Fraction of Modified Ntg1 (%)										
	WT	K20R	K38R	K376R	K388R	K396R	K20, 38R	K20,38, 376R	K20,38, 376, 388R	K20,38, 376,388, 396R	K376, 388, 396R
Ntg1-Tri	0.31	0.11	0.02	0.18	0.13	0.00	0.36	0.02	0.00	0.00	0.00
Ntg1-Di	0.49	0.14	0.00	0.23	0.21	0.18	0.00	0.05	0.00	0.00	0.00
Ntg1-Mono	6.52	5.53	4.27	5.30	4.98	3.45	5.50	4.66	4.51	0.04	1.24

2.7 Acknowledgments and funding sources

2.7.1 Acknowledgments

We would like to thank Christine Dunham for her help with the generation of the protein homology model. We would also like to acknowledge members of the Corbett and Doetsch laboratories for helpful discussions and advice.

This study was supported in part by the Emory Integrated Genomics Core (EIGC), which is subsidized by the Emory University School of Medicine and is one of the Emory Integrated Core Facilities. Additional support was provided by the National Center for Advancing Translational Sciences of the National Institutes of Health under Award Number UL1TR000454. The content is solely the responsibility of the authors and does not necessarily reflect the official views of the National Institutes of Health.

2.7.2 Funding sources

This work was supported by the National Institutes of General Medical Sciences [RO1 GM05872816; T32 GM008490 22; and F31 GM115178 01] and the National Institutes of Health [NIH ES011163].

2.8 References

1. Altieri F, Grillo C, Maceroni M, Chichiarelli S. DNA damage and repair: from molecular mechanisms to health implications. Antioxid Redox Signal. 2008; 10(5):891–937. [PubMed: 18205545]

2. Fortini P, Pascucci B, Parlanti E, D'Errico M, Simonelli V, Dogliotti E. The base excision repair: mechanisms and its relevance for cancer susceptibility. Biochimie. 2003; 85(11):1053–1071. [PubMed: 14726013]

3. Boesch P, Weber-Lotfi F, Ibrahim N, Tarasenko V, Cosset A, Paulus F, et al. DNA repair in organelles: Pathways, organization, regulation, relevance in disease and aging. Biochim Biophys Acta. 2011; 1813(1):186–200. [PubMed: 20950654]

4. Kryston TB, Georgiev AB, Pissis P, Georgakilas AG. Role of oxidative stress and DNA damage in human carcinogenesis. Mutation research. 2011; 711(1–2):193–201. [PubMed: 21216256]

5. Pan Y. Mitochondria, reactive oxygen species, and chronological aging: a message from yeast. Experimental gerontology. 2011; 46(11):847–852. [PubMed: 21884780]

6. de Gruijl FR, van Kranen HJ, Mullenders LH. UV-induced DNA damage, repair, mutations and oncogenic pathways in skin cancer. J Photochem Photobiol B. 2001; 63(1–3):19–27. [PubMed: 11684448]

7. Fraga CG, Shigenaga MK, Park JW, Degan P, Ames BN. Oxidative damage to DNA during aging: 8-hydroxy-2'-deoxyguanosine in rat organ DNA and urine. Proc Natl Acad Sci U S A. 1990; 87(12): 4533–4537. [PubMed: 2352934]

8. Nakamura J, Swenberg JA. Endogenous apurinic/apyrimidinic sites in genomic DNA of mammalian tissues. Cancer Res. 1999; 59(11):2522–2526. [PubMed: 10363965]

9. Salmon TB, Evert BA, Song B, Doetsch PW. Biological consequences of oxidative stress-induced DNA damage in Saccharomyces cerevisiae. Nucleic Acids Res. 2004; 32(12):3712–3723. [PubMed: 15254273]

10. Stefl S, Nishi H, Petukh M, Panchenko AR, Alexov E. Molecular mechanisms of disease-causing missense mutations. J Mol Biol. 2013; 425(21):3919–3936. [PubMed: 23871686]

11. Evert BA, Salmon TB, Song B, Jingjing L, Siede W, Doetsch PW. Spontaneous DNA damage in Saccharomyces cerevisiae elicits phenotypic properties similar to cancer cells. J Biol Chem. 2004; 279(21):22585–22594. [PubMed: 15020594]

12. Beckman KB, Ames BN. Oxidative decay of DNA. J Biol Chem. 1997; 272(32):19633–19636. [PubMed: 9289489]

13. Wallace DC. Diseases of the mitochondrial DNA. Annu Rev Biochem. 1992; 61:1175–1212. [PubMed: 1497308]

14. Wei YH. Oxidative stress and mitochondrial DNA mutations in human aging. Proc Soc Exp Biol Med. 1998; 217(1):53–63. [PubMed: 9421207]

15. Cooke MS, Evans MD, Dizdaroglu M, Lunec J. Oxidative DNA damage: mechanisms, mutation, and disease. FASEB J. 2003; 17(10):1195–1214. [PubMed: 12832285]

16. Lenaz G. Role of mitochondria in oxidative stress and ageing. Biochim Biophys Acta. 1998; 1366(1–2):53–67. [PubMed: 9714734]

17. Muftuoglu M, de Souza-Pinto NC, Dogan A, Aamann M, Stevnsner T, Rybanska I, et al. Cockayne syndrome group B protein stimulates repair of formamidopyrimidines by NEIL1 DNA glycosylase. J Biol Chem. 2009; 284(14):9270–9279. [PubMed: 19179336]

18. Larsen NB, Rasmussen M, Rasmussen LJ. Nuclear and mitochondrial DNA repair: similar pathways? Mitochondrion. 2005; 5(2):89–108. [PubMed: 16050976]

19. Degtyareva NP, Chen L, Mieczkowski P, Petes TD, Doetsch PW. Chronic oxidative DNA damage due to DNA repair defects causes chromosomal instability in Saccharomyces cerevisiae. Mol Cell Biol. 2008; 28(17):5432–5445. [PubMed: 18591251]

20. Sentürker S, Auffret van der Kemp P, You HJ, Doetsch PW, Dizdaroglu M, Boiteux S. Substrate specificities of the Ntg1 and Ntg2 proteins of Saccharomyces cerevisiae for oxidized DNA bases are not identical. Nucleic Acids Res. 1998; 26(23):5270–5276. [PubMed: 9826748]

21. Swartzlander, DB.; Bauer, NC.; Corbett, AH.; Doetsch, PW. Chapter 5 – Regulation of base excision repair in eukaryotes by dynamic localization strategies. In: Doetsch, PW., editor. Mechanisms of DNA Repair. Progress in Molecular Biology and Translational Science. Vol. 110. Oxford, UK: Academic Press; 2012. p. 93-121.

22. Nilsen H, Krokan HE. Base excision repair in a network of defense and tolerance. Carcinogenesis. 2001; 22(7):987–998. [PubMed: 11408341]

23. Memisoglu A, Samson L. Base excision repair in yeast and mammals. Mutat Res. 2000; 451(1–2): 39–51. [PubMed: 10915864]

24. Griffiths LM, Swartzlander D, Meadows KL, Wilkinson KD, Corbett AH, Doetsch PW. Dynamic compartmentalization of base excision repair proteins in response to nuclear and mitochondrial oxidative stress. Mol Cell Biol. 2009; 29(3):794–807. [PubMed: 19029246]

25. Ikeda S, Kohmoto T, Tabata R, Seki Y. Differential intracellular localization of the human and mouse endonuclease III homologs and analysis of the sorting signals. DNA Repair (Amst). 2002; 1(10):847–854. [PubMed: 12531031]

26. You HJ, Swanson RL, Harrington C, Corbett AH, Jinks-Robertson S, Sentürker S, et al. Saccharomyces cerevisiae Ntg1p and Ntg2p: broad specificity N-glycosylases for the repair of

oxidative DNA damage in the nucleus and mitochondria. Biochemistry. 1999; 38(35):11298–11306. [PubMed: 10471279]

27. Bettermann K, Benesch M, Weis S, Haybaeck J. SUMOylation in carcinogenesis. Cancer Lett. 2012; 316(2):113–125. [PubMed: 22138131]

28. Cremona CA, Sarangi P, Zhao X. Sumoylation and the DNA damage response. Biomolecules. 2012; 2(3):376–388. [PubMed: 24926426]

29. Bartek J, Hodny Z. SUMO boosts the DNA damage response barrier against cancer. Cancer Cell. 2010; 17(1):9–11. [PubMed: 20129245]

30. Johnson ES. Protein modification by SUMO. Annu Rev Biochem. 2004; 73:355–382. [PubMed: 15189146]

31. Dery U, Masson JY. Twists and turns in the function of DNA damage signaling and repair proteins by post-translational modifications. DNA Repair (Amst). 2007; 6(5):561–577. [PubMed: 17258515]

32. Dou H, Huang C, Van Nguyen T, Lu LS, Yeh ET. SUMOylation and de-SUMOylation in response to DNA damage. FEBS Lett. 2011; 585(18):2891–2896. [PubMed: 21486569]

33. Thompson LH. Recognition, signaling, and repair of DNA double-strand breaks produced by ionizing radiation in mammalian cells: The molecular choreography. Mutation Research/Reviews in Mutation Research. 2012; 751(2):158–246.

34. Kirkin V, Dikic I. Role of ubiquitin- and Ubl-binding proteins in cell signaling. Curr Opin Cell Biol. 2007; 19(2):199–205. [PubMed: 17303403]

35. Bergink S, Jentsch S. Principles of ubiquitin and SUMO modifications in DNA repair. Nature. 2009; 458(7237):461–467. [PubMed: 19325626]

36. Goto M, Shinmura K, Igarashi H, Kobayashi M, Konno H, Yamada H, et al. Altered expression of the human base excision repair gene NTH1 in gastric cancer. Carcinogenesis. 2009; 30(8):1345–1352. [PubMed: 19414504]

37. Koketsu S, Watanabe T, Nagawa H. Expression of DNA repair protein: MYH, NTH1, and MTH1 in colorectal cancer. Hepatogastroenterology. 2004; 51(57):638–642. [PubMed: 15143881]

38. Short E, Thomas LE, Hurley J, Jose S, Sampson JR. Inherited predisposition to colorectal cancer: towards a more complete picture. J Med Genet. 2015; 52(12):791–796. [PubMed: 26297796]

39. Rivera B, Castellsague E, Bah I, van Kempen LC, Foulkes WD. Biallelic NTHL1 Mutations in a Woman with Multiple Primary Tumors. The New England journal of medicine. 2015; 373(20): 1985–1986. [PubMed: 26559593]

40. Weren RD, Ligtenberg MJ, Kets CM, de Voer RM, Verwiel ET, Spruijt L, et al. A germline homozygous mutation in the base-excision repair gene NTHL1 causes adenomatous polyposis and colorectal cancer. Nature genetics. 2015; 47(6):668–671. [PubMed: 25938944]

41. Bauer NC, Corbett AH, Doetsch PW. The current state of eukaryotic DNA base damage and repair. Nucleic Acids Res. 2015; 43(21):10083–10101. [PubMed: 26519467]

42. Hannich JT, Lewis A, Kroetz MB, Li SJ, Heide H, Emili A, et al. Defining the SUMO-modified proteome by multiple approaches in Saccharomyces cerevisiae. J Biol Chem. 2005; 280(6):4102–4110. [PubMed: 15590687]

43. Hardeland U, Steinacher R, Jiricny J, Schar P. Modification of the human thymine-DNA glycosylase by ubiquitin-like proteins facilitates enzymatic turnover. EMBO JOURNAL. 2002; 21(6):1456–1464. [PubMed: 11889051]

44. Smet-Nocca C, Wieruszeski JM, Leger H, Eilebrecht S, Benecke A. SUMO-1 regulates the conformational dynamics of thymine-DNA glycosylase regulatory domain and competes with its DNA binding activity. BMC Biochem. 2011; 12:4. [PubMed: 21284855]

45. Steinacher R, Schar P. Functionality of human thymine DNA glycosylase requires SUMO-regulated changes in protein conformation. Curr Biol. 2005; 15(7):616–623. [PubMed: 15823533]

46. Ito H, Fukuda Y, Murata K, Kimura A. Transformation of intact yeast cells treated with alkali cations. J Bacteriol. 1983; 153(1):163–168. [PubMed: 6336730]

47. Sikorski RS, Hieter P. A system of shuttle vectors and yeast host strains designed for efficient manipulation of DNA in Saccharomyces cerevisiae. Genetics. 1989; 122(1):19–27. [PubMed: 2659436]

48. Parikh SS, Mol CD, Slupphaug G, Bharati S, Krokan HE, Tainer JA. Base excision repair initiation revealed by crystal structures and binding kinetics of human uracil-DNA glycosylase with DNA. Embo J. 1998; 17(17):5214–5226. [PubMed: 9724657]

49. Tang CH, Wei W, Liu L. Regulation of DNA repair by S-nitrosylation. Biochim Biophys Acta. 2012; 1820(6):730–735. [PubMed: 21571039]

50. Storici F, Resnick MA. The delitto perfetto approach to in vivo site-directed mutagenesis and chromosome rearrangements with synthetic oligonucleotides in yeast. Methods Enzymol. 2006; 409:329–345. [PubMed: 16793410]

51. Swartzlander DB, Griffiths LM, Lee J, Degtyareva NP, Doetsch PW, Corbett AH. Regulation of base excision repair: Ntg1 nuclear and mitochondrial dynamic localization in response to genotoxic stress. Nucleic Acids Res. 2010; 38(12):3963–3974. [PubMed: 20194111]

52. Sampson DA, Wang M, Matunis MJ. The small ubiquitin-like modifier-1 (SUMO-1) consensus sequence mediates Ubc9 binding and is essential for SUMO-1 modification. J Biol Chem. 2001; 276(24):21664–21669. [PubMed: 11259410]

53. Johnson ES, Gupta AA. An E3-like factor that promotes SUMO conjugation to the yeast septins. Cell. 2001; 106(6):735–744. [PubMed: 11572779]

54. Johnson ES, Blobel G. Ubc9p is the conjugating enzyme for the ubiquitin-like protein Smt3p. J Biol Chem. 1997; 272(43):26799–26802. [PubMed: 9341106]

55. Bylebyl GR, Belichenko I, Johnson ES. The SUMO isopeptidase Ulp2 prevents accumulation of SUMO chains in yeast. The Journal of biological chemistry. 2003; 278(45):44113–44120. [PubMed: 12941945]

56. Tatham MH, Jaffray E, Vaughan OA, Desterro JM, Botting CH, Naismith JH, et al. Polymeric chains of SUMO-2 and SUMO-3 are conjugated to protein substrates by SAE1/SAE2 and Ubc9. J Biol Chem. 2001; 276(38):35368–35374. [PubMed: 11451954]

57. Cheng CH, Lo YH, Liang SS, Ti SC, Lin FM, Yeh CH, et al. SUMO modifications control assembly of synaptonemal complex and polycomplex in meiosis of Saccharomyces cerevisiae. Genes Dev. 2006; 20(15):2067–2081. [PubMed: 16847351]

58. Li SJ, Hochstrasser M. A new protease required for cell-cycle progression in yeast. Nature. 1999; 398(6724):246–251. [PubMed: 10094048]

59. Li SJ, Hochstrasser M. The yeast ULP2 (SMT4) gene encodes a novel protease specific for the ubiquitin-like Smt3 protein. Mol Cell Biol. 2000; 20(7):2367–2377. [PubMed: 10713161]

60. Schwienhorst I, Johnson ES, Dohmen RJ. SUMO conjugation and deconjugation. Molecular & general genetics : MGG. 2000; 263(5):771–786. [PubMed: 10905345]

61. Soustelle C, Vernis L, Freon K, Reynaud-Angelin A, Chanet R, Fabre F, et al. A new Saccharomyces cerevisiae strain with a mutant Smt3-deconjugating Ulp1 protein is affected in DNA replication and requires Srs2 and homologous recombination for its viability. Molecular and cellular biology. 2004; 24(12):5130–5143. [PubMed: 15169880]

62. Vertegaal AC, Andersen JS, Ogg SC, Hay RT, Mann M, Lamond AI. Distinct and overlapping sets of SUMO-1 and SUMO-2 target proteins revealed by quantitative proteomics. Mol Cell Proteomics. 2006; 5(12):2298–2310. [PubMed: 17000644]

63. Rodriguez MS, Dargemont C, Hay RT. SUMO-1 conjugation in vivo requires both a consensus modification motif and nuclear targeting. J Biol Chem. 2001; 276(16):12654–12659. [PubMed: 11124955]

64. Xue Y, Zhou F, Fu C, Xu Y, Yao X. SUMOsp: a web server for sumoylation site prediction. Nucleic Acids Res. 2006; 34:W254–W257. Web Server issue. [PubMed: 16845005]

65. Ren J, Gao X, Jin C, Zhu M, Wang X, Shaw A, et al. Systematic study of protein sumoylation: Development of a site-specific predictor of SUMOsp 2.0. Proteomics. 2009; 9(12):3409–3412.

66. Zhao Q, Xie Y, Zheng Y, Jiang S, Liu W, Mu W, et al. GPS-SUMO: a tool for the prediction of sumoylation sites and SUMO-interaction motifs. Nucleic Acids Res. 2014; 42:W325–W330. Web Server issue. [PubMed: 24880689]

67. Xu J, He Y, Qiang B, Yuan J, Peng X, Pan XM. A novel method for high accuracy sumoylation site prediction from protein sequences. BMC Bioinformatics. 2008; 9:8. [PubMed: 18179724]

68. Green, JR.; Dmochowski, GM.; Golshani, A. Can Med Biol Eng Conf. Vancouver, BC: 2006. Prediction of protein sumoylation sites via parallel cascade identification.

69. Thayer MM, Ahern H, Xing D, Cunningham RP, Tainer JA. Novel DNA binding motifs in the DNA repair enzyme endonuclease III crystal structure. EMBO J. 1995; 14(16):4108–4120. [PubMed: 7664751]

70. Kelley LA, Mezulis S, Yates CM, Wass MN, Sternberg MJE. The Phyre2 web portal for protein modeling, prediction and analysis. Nat Protocols. 2015; 10(6):845–858. [PubMed: 25950237]

71. Baba D, Maita N, Jee JG, Uchimura Y, Saitoh H, Sugasawa K, et al. Crystal structure of thymine DNA glycosylase conjugated to SUMO-1. Nature. 2005; 435(7044):979–982. [PubMed: 15959518]

72. Dizdaroglu M, Bauche C, Rodriguez H, Laval J. Novel substrates of Escherichia coli nth protein and its kinetics for excision of modified bases from DNA damaged by free radicals. Biochemistry. 2000; 39(18):5586–5592. [PubMed: 10820032]

73. Augeri L, Lee Y-M, Barton AB, Doetsch PW. Purification, characterization, gene cloning, and expression of Saccharomyces cerevisiae redoxyendonuclease, a homolog of Escherichia coli Endonuclease III. Biochemistry. 1997; 36(4):721–729. [PubMed: 9020769]

74. Boiteux S, Guillet M. Abasic sites in DNA: repair and biological consequences in Saccharomyces cerevisiae. DNA Repair. 2004; 3(1):1–12. [PubMed: 14697754]

75. Sedgwick B, Bates PA, Paik J, Jacobs SC, Lindahl T. Repair of alkylated DNA: recent advances. DNA Repair. 2007; 6(4):429–442. [PubMed: 17112791]

76. Swanson RL, Morey NJ, Doetsch PW, Jinks-Robertson S. Overlapping specificities of base excision repair, nucleotide excision repair, recombination, and translesion synthesis pathways for DNA base damage in Saccharomyces cerevisiae. Mol Cell Biol. 1999; 19(4):2929–2935. [PubMed: 10082560]

77. Gnad F, Gunawardena J, Mann M. PHOSIDA 2011: the posttranslational modification database. Nucleic Acids Res. 2011; 39:D253–D260. (Database issue). [PubMed: 21081558]

78. Rowe LA, Degtyareva N, Doetsch PW. DNA damage-induced reactive oxygen species (ROS) stress response in Saccharomyces cerevisiae. Free Radic Biol Med. 2008; 45(8):1167–1177. [PubMed: 18708137]

79. Jentsch S, Muller S. Regulatory Functions of Ubiquitin and SUMO in DNA Repair Pathways. Subcell Biochem. 2010; 54:184–194. [PubMed: 21222283]

80. Ho Y, Gruhler A, Heilbut A, Bader GD, Moore L, Adams SL, et al. Systematic identification of protein complexes in Saccharomyces cerevisiae by mass spectrometry. Nature. 2002; 415(6868): 180–183. [PubMed: 11805837]

81. Weinert TA, Kiser GL, Hartwell LH. Mitotic checkpoint genes in budding yeast and the dependence of mitosis on DNA replication and repair. Genes & development. 1994; 8(6):652–665. [PubMed: 7926756]

82. Noskov V, Maki S, Kawasaki Y, Leem SH, Ono B, Araki H, et al. The RFC2 gene encoding a subunit of replication factor C of Saccharomyces cerevisiae. Nucleic Acids Res. 1994; 22(9): 1527–1535. [PubMed: 8202350]

83. Boiteux S, Laval J. Mutagenesis by alkylating agents: coding properties for DNA polymerase of poly (dC) template containing 3-methylcytosine. Biochimie. 1982; 64(8–9):637–641. [PubMed: 6814512]

84. Larson K, Sahm J, Shenkar R, Strauss B. Methylation-induced blocks to in vitro DNA replication. Mutation research. 1985; 150(1–2):77–84. [PubMed: 4000169]

85. Prakash L. Lack of chemically induced mutation in repair-deficient mutants of yeast. Genetics. 1974; 78(4):1101–1118. [PubMed: 4376097]

86. Schwartz MF, Duong JK, Sun Z, Morrow JS, Pradhan D, Stern DF. Rad9 phosphorylation sites couple Rad53 to the Saccharomyces cerevisiae DNA damage checkpoint. Mol Cell. 2002; 9(5): 1055–1065. [PubMed: 12049741]

87. Toh GW, Lowndes NF. Role of the Saccharomyces cerevisiae Rad9 protein in sensing and responding to DNA damage. Biochemical Society transactions. 2003; 31(Pt 1):242–246. [PubMed: 12546694]

88. Conde F, Ontoso D, Acosta I, Gallego-Sanchez A, Bueno A, San-Segundo PA. Regulation of tolerance to DNA alkylating damage by Dot1 and Rad53 in Saccharomyces cerevisiae. DNA Repair (Amst). 2010; 9(10):1038–1049. [PubMed: 20674515]

89. Sarangi P, Bartosova Z, Altmannova V, Holland C, Chavdarova M, Lee SE, et al. Sumoylation of the Rad1 nuclease promotes DNA repair and regulates its DNA association. Nucleic Acids Res. 2014; 42(10):6393–6404. [PubMed: 24753409]

90. Hang LE, Lopez CR, Liu X, Williams JM, Chung I, Wei L, et al. Regulation of Ku-DNA association by Yku70 C-terminal tail and SUMO modification. J Biol Chem. 2014; 289(15): 10308–10317. [PubMed: 24567323]

91. Altmannova V, Eckert-Boulet N, Arneric M, Kolesar P, Chaloupkova R, Damborsky J, et al. Rad52 SUMOylation affects the efficiency of the DNA repair. Nucleic Acids Res. 2010; 38(14):4708– 4721. [PubMed: 20371517]

92. Hardeland U, Steinacher R, Jiricny J, Schar P. Modification of the human thymine-DNA glycosylase by ubiquitin-like proteins facilitates enzymatic turnover. EMBO J. 2002; 21(6):1456–1464. [PubMed: 11889051]

93. Sacher M, Pfander B, Hoege C, Jentsch S. Control of Rad52 recombination activity by double-strand break-induced SUMO modification. Nat Cell Biol. 2006; 8(11):1284–1290. [PubMed: 17013376]

94. Steinacher R, Schar P. Functionality of human thymine DNA glycosylase requires SUMO-regulated changes in protein conformation. Curr Biol. 2005; 15(7):616–623. [PubMed: 15823533]

95. Bergink S, Ammon T, Kern M, Schermelleh L, Leonhardt H, Jentsch S. Role of Cdc48/p97 as a SUMO-targeted segregase curbing Rad51–Rad52 interaction. Nat Cell Biol. 2013; 15(5):526–532. [PubMed: 23624404]

96. Papouli E, Chen S, Davies AA, Huttner D, Krejci L, Sung P, et al. Crosstalk between SUMO and ubiquitin on PCNA is mediated by recruitment of the helicase Srs2p. Mol Cell. 2005; 19(1):123– 133. [PubMed: 15989970]

97. Pfander B, Moldovan GL, Sacher M, Hoege C, Jentsch S. SUMO-modified PCNA recruits Srs2 to prevent recombination during S phase. Nature. 2005; 436(7049):428–433. [PubMed: 15931174]

98. Torres-Rosell J, Sunjevaric I, De Piccoli G, Sacher M, Eckert-Boulet N, Reid R, et al. The Smc5– Smc6 complex and SUMO modification of Rad52 regulates recombinational repair at the ribosomal gene locus. Nat Cell Biol. 2007; 9(8):923–931. [PubMed: 17643116]

99. Chen XL, Reindle A, Johnson ES. Misregulation of 2 microm circle copy number in a SUMO pathway mutant. Molecular and cellular biology. 2005; 25(10):4311–4320. [PubMed: 15870299]

Web References

100. http://www.abgent.com/sumoplot Date Last used: 8/4/16

Chapter 3

A *Saccharomyces cerevisiae* model for overexpression of Ntg1, a base excision DNA repair protein, reveals novel genetic interactions

Annie J. McPherson-Davie[1,2], Ziad M. Jowhar[1,3], Paul W. Doetsch[4], and Anita H. Corbett[1†]

[1]Department of Biology, Emory University, Atlanta, GA, 30322, USA.

[2]Genetics and Molecular Biology Graduate Program, Emory University, Atlanta, GA, 30322, USA.

[3]Initiative for Maximizing Student Development, Emory University, Atlanta, GA, 30322, USA.

[4]Mutagenesis and DNA Repair Regulation Group, National Institutes of Environmental Health Sciences, Division of Intramural Research, Genome Integrity & Structural Biology Laboratory, Durham, NC, 27709, USA.

† To whom correspondence should be addressed

The contribution by AJMD includes writing and editing the manuscript, image creation, and all data presented in Figures 3.1-4.

Submitted to DNA Repair (Amst).

3.1 Abstract

The base excision repair (BER) pathway repairs oxidative DNA damage, a very common and detrimental form of damage to the genome. Although the biochemical steps BER have been well characterized, little is understood about how the pathway is regulated. Such regulation is critical, as cells must be poised to respond rapidly to DNA damage while avoiding activity of repair proteins that can produce DNA damage as intermediates in the repair pathway. Indeed, recent work reveals that overexpression of the human BER protein, NTHL1, a DNA N-glycosylase, can cause genomic instability and early cellular hallmarks of cancer (Limpose *et al*., *NAR* 2018). We developed a *Saccharomyces cerevisiae* model to explore how overexpression of NTHL1 may impair cellular function. Overexpression of Ntg1, the budding yeast orthologue of NTHL1, impairs cell growth. To dissect mechanisms underlying this growth defect, we overexpressed either wild-type Ntg1 or a catalytically inactive variant of Ntg1 ($ntg1_{catdead}$). Consistent with results obtained with NTHL1, both variants of Ntg1 impair cell growth, but only the wild-type protein causes accumulation of double-strand breaks and chromosome loss. We took advantage of the budding yeast system to screen a panel of DNA repair mutants for resistance/sensitivity to overexpression of wild-type Ntg1 or $ntg1_{catdead}$. This analysis identified several cellular pathways that protect cells from Ntg1-induced damage, including nucleotide excision repair (NER). The homologous recombination pathway is critical to counter overexpression of wild-type Ntg1, but not $ntg1_{catdead}$, consistent with the finding that overexpression of wild-type Ntg1 causes accumulation of double-strand breaks. Finally, we identified a link to sumoylation and employed a variant of Ntg1 that cannot be modified to SUMO to probe how this post-translational modification could contribute to regulation of Ntg1 function. This study describes a budding yeast system to understand how cells regulate and respond to demands of dysregulation of the BER pathway.

3.2 Introduction

Within both the nucleus and the mitochondria, DNA is susceptible to damage [1-6]. There are many sources of both exogenous and endogenous DNA damage [1, 2, 7-9]. An abundant endogenous source of DNA damage is reactive oxygen species (ROS) which are a byproduct of energy production within the cell [1, 2, 10-14]. ROS can cause oxidative damage throughout the cell and can induce many types of DNA damage including 8-oxo-2'deoxyguanosine [1, 10, 15, 16]. These damages must be rapidly and efficiently repaired to maintain genomic and genetic stability [10, 17-19]. To combat the myriad of damages that occur, cells employ a battery of DNA repair pathways [1, 2, 9, 10, 17, 20-22].

Base excision repair (BER) is the major repair pathway for oxidative DNA damage [1, 10, 17, 23, 24]. BER is initiated by an N-glycosylase detecting a non-helix distorting base damage [24-26]. The N-glycosylase flips out the base and cleaves the base from the backbone, resulting in an abasic site25,28,29. The DNA backbone on the 5'-side of the abasic site is then cleaved by an apurinic/apyrimidinic (AP) endonuclease or on the 3'-side by a N-glycosylase with AP lyase function [24, 27]. If cleavage occurs on the 3'-side of the abasic site, an AP endonuclease must cleave the backbone and expose the preceding hydroxyl, to allow DNA synthesis to occur [24, 27]. After synthesis, the strand is ligated back together to complete the repair process [10, 24].

The BER proteins must be available for rapid deployment in response to DNA damage but must also be precisely regulated to prevent promiscuous damage to the genome [28-31]. One major protein that initiates DNA repair in the BER pathway is the *N*-glycoslyase NTHL1, an evolutionarily conserved member of the endonuclease III family [32-34]. NTHL1 is a bifunctional DNA N-glycosylase with associated apurinic/apyrimidinic (AP) lyase function [35, 36]. As

NTHL1 initiates DNA repair by introducing a nick in the DNA backbone, its activity must be tightly regulated to prevent the accumulation of spurious nicks in the genomic material [30].

As with a number of other DNA repair pathway proteins, mutations in the NTHL1 gene have been linked to cancer [37-40]. In addition, consistent with the concept that precise regulation of NTHL1 function is critical, recent work has revealed that NTHL1 expression is misregulated in a variety of types of cancer [41-43]. Furthermore, a recent study demonstrated that overexpression of NTHL1 in non-cancerous human bronchial epithelial cells results in loss of genetic information and early hallmarks of cancer including genomic instability [44]. Interestingly, overexpression of a catalytically inactive form of NTHL1 induced similar phenotypes as wild-type NTHL1, suggesting multiple modes by which overexpression of this BER protein can induce DNA damage/genome instability. Overexpression of NTHL1 in this cell model triggered replication stress signaling. Studies using a reporter system revealed a decrease in homologous recombination in these cells overexpressing NTHL1. This work provided important insight into how overexpression of a BER protein could contribute to genomic instability but could not broadly define the spectrum of pathways cells may employ to respond to such damage [44].

In the present study, we sought to develop budding yeast as a model to investigate how cells respond to damage induced by dysregulation of early steps in BER. To do so, we first established that overexpression of S. cerevisiae Ntg1, which is the budding yeast orthologue of NTHL1 [45], causes a growth defect when overexpressed in S. cerevisiae. Consistent with the studies of NTHL1 [44], overexpression of catalytically inactive Ntg1 (ntg1catdead) also impairs cell growth. While overexpression of wild-type Ntg1 causes an increase in double-strand breaks and chromosome loss, overexpression of catalytically inactive ntg1catdead does not. These results suggest that the growth defects seen in cells that overexpress Ntg1 or ntg1catdead are triggered, at

least in part, by separate mechanisms. We then took advantage of the budding yeast system to probe genetic interactions with BER by screening a panel from the gene deletion collection for deletion mutants that show either enhanced sensitivity or resistance to overexpression of Ntg1 and/or ntg1catdead. We identified a number of pathways that show such genetic interactions, including the homologous recombination pathway, nucleotide excision repair, and SUMO-mediated DNA damage response. As previous work demonstrated that Ntg1 is sumoylated [46], we expanded our analysis of the sumoylation pathway and also explored whether SUMO modification of Ntg1 impacts overexpression of Ntg1. By taking advantage of these genetic approaches, we have identified pathways of interest that could be explored to better understand the mechanism by which overexpression of human NTHL1 could contribute to cancer phenotypes.

3.3 Materials and methods

3.3.1 Strains, plasmids, and media

All haploid *S. cerevisiae* strains and plasmids used in this study are listed in Table 1. *S. cerevisiae* cells were cultured at 25°C, 30°C, or 37°C in YPD medium (1% yeast extract, 2% peptone, 2% dextrose, 0.005% adenine sulfate, and 2% agar for plates) or SD medium (0.17% yeast nitrogen base, 0.5% ammonium sulfate, 0.5% adenine sulfate, and 2% agar for plates) or SD medium (0.17% yeast nitrogen base, 0.5% ammonium sulfate, 0.005% adenine sulfate, and 2% agar for plates) with either 2% dextrose, 3% raffinose, 2% galactose, 3% glycerol, or both 3% raffinose and 2% galactose. To introduce plasmids, cells were transformed by a modified lithium acetate method [47].

A galactose inducible 2μ vector (*pGAL1*, *URA3*), pPS293 (Addgene plasmid #8851; http://n2t.net/addgene:8851; RRID: Addgene 8851) was employed to express C-terminally

epitope-tagged Ntg1-2xmyc (pAC3425). We expressed wild-type Ntg1, a catalytically inactive Ntg1 (ntg1$_{catdead}$; *ntg1^{K243Q}*) [46, 48, 49], a nonsumoylatable Ntg1 (ntg1ΔSUMO; *ntg1^{K20,38,376,388,396R}*) [46]. This Ntg1 variant was also generated in the catalytically inactive form ntg1ΔSUMO$_{catdead}$ (*ntg1^{K20,38,376,388,396R,K243Q}*). The resulting plasmids were sequenced to ensure the correct desired sequence and the absence of any additional mutations.

The *S. cerevisiae* strain expressing endogenous Rad52-YFP [50] was provided by Rodney Rothstein. This strain was employed to assay double-strand break formation. The *S. cerevisiae* reporter strain containing a chromosome fragment with *SUP11* [51] was provided by Munira Basrai. This strain was employed to quantify chromosome loss in response to overexpression of Ntg1. The *S. cerevisiae* haploid deletion collection was utilized to explore the genetic interactions with overexpression of Ntg1. The SUMO pathway mutant collection (E3 ligase mutant strains, *siz1Δ*, *siz2Δ*, and *siz1Δ/siz2Δ*, and desumoylase mutant stains *ulp1-1* and *ulp2Δ*) were utilized to assess the impact of global sumoylation loss or accumulation [52-54].

3.3.2 *S. cerevisiae* growth assays

S. cerevisiae cells containing *URA3* plasmids were grown in media lacking uracil plus 3% raffinose overnight. The next day, the OD600 of the cultures was measured and the cultures were diluted down to the lowest OD600. The cultures were then serially diluted 5-fold and 2.5 μL of each dilution was spotted onto plates lacking uracil containing either 2% glucose or 2% galactose. Plates were grown at 30°C and pictures were taken of the plates on day 2.

Liquid growth curves were collected for three independently isolated colonies per sample. Cells were grown overnight at 30°C in media lacking uracil with 3% raffinose. Cell concentrations were normalized by OD600, and then samples were diluted to an OD600 of 0.05 in 150 μL of

media lacking uracil with either 2% glucose or 2% galactose and placed in the wells of a 96-well microtiter plate. Cell samples were loaded in triplicate, were grown at 30°C with shaking, and absorbance at OD600 was measured every 30 minutes for 24 hours in an ELX808 Ultra microplate reader with KCjunior software (Bio-Tek Instruments, Inc.).

3.3.3 Cell viability assay

Cells were grown overnight in media lacking uracil plus 3% raffinose. The next day, cultures were harvested at mid-log phase (OD600 of 0.3-0.6) and equalized, then diluted 1:10 into media lacking uracil plus 2% galactose. Cells were grown in galactose overnight at 30°C. The next morning, OD600 was measured, cells were equalized again, diluted 1:1000 and an estimated 100 cells were plated onto plates lacking uracil plus 2% glucose and the plates were incubated for five days at 30°C. On day five, colonies were counted. The number of colonies grown on the plates containing Vector alone was considered the total number of live cells plated and therefore considered 100% survival. We then divided the number of colonies grown per plasmid by the total number of live cells plated and converted to a percentage. These experiments were conducted in both biological and technical triplicate. Standard deviations in the biological replicates are shown.

3.3.4 Quantification of cell growth

Cells were grown, spotted, and imaged as described above. In order to better compare differences in growth, each strain and corresponding plasmid is given a value between 1 and 10 based on growth. A value of 10 corresponds to the growth of the control Vector for each mutant. A spot with large full colonies is given the value of 2 points, while spots with smaller colonies are given the value of 1. Spots with no growth are given a value of 0. This scoring approach is used

for each of the five spots of serially diluted cells, yielding the range of 1-10 for each sample analyzed as we only analyzed samples where some growth could be detected in the most concentrated spot. These values, which are averages of the value obtained for biological triplicates, are displayed as a heat map.

3.3.5 Double-strand break formation assay

Cells containing an endogenous Rad52-YFP [50] (generously provided by Rodney Rothstein) were grown overnight in media lacking uracil plus 3% raffinose. The next day, cells were diluted and grown until mid-log phase (OD600 0.3-0.6). At mid-log phase, the cells were dosed with 2% galactose and control samples with Vector alone were incubated in the presence of 2% galactose and 0.3% MMS. These cultures were incubated at 30°C for 2 hours. 30 minutes before treatment ended, DAPI was added to the culture to allow visualization of the nucleus. After treatment, the MMS was inactivated with 10% Sodium Thiosulfate and fixed with 4% formaldehyde. Cells were immobilized in agarose and 15 YFP and DAPI images were captured at 0.2-µm intervals along the z-axis with an oil immersion 100x objective on a Confocal Olympus FV1000 Upright microscope quantitated with FIJI. Images were analyzed for foci in at least three consecutive z-planes. For each sample, 300 individual cells were analyzed, and the data are represented as a percentage. The results shown are the average of three independent experiments. Merged images were created in FIJI Is Just ImageJ (FIJI) [55].

3.3.6 Chromosome loss assay

An *S. cerevisiae* reporter strain containing a chromosome fragment with *SUP11* [51, 56] (generously provided by Munira Basrai) plus test plasmids were grown overnight in media lacking

uracil and histidine with limited adenine plus 3% raffinose. The next day, cultures were harvested and diluted down to an OD600 of 1 and then were diluted 1:1000 in water. Then 10, 30, and 50 μL of diluted cells were plated onto plates lacking uracil and containing limited adenine with either 2% glucose or 3% raffinose and 2% galactose and grown for 5 days at 30°C. Plates were then moved to 4°C for 3 days to enrich the red pigment. For each sample, 300 colonies were counted and any colony exhibiting red pigment were noted. The number of colonies with pigment was divided by the number of colonies counted. These data were generated in triplicate.

3.4 Results

3.4.1 Overexpression of Ntg1 impairs *S. cerevisiae* cell growth.

Our previous study demonstrates that overexpression of human NTHL1 causes early cancer phenotypes, including genomic instability, in mammalian cells [44], but defining the pathways that cells employ to combat this damage is challenging in mammalian cells. Thus, we assessed whether a budding yeast model could be employed to define the pathways that contribute to phenotypes observed upon overexpression of the BER protein, NTHL1. To analyze the budding yeast counterpart of NTHL1, Ntg1 [45], we first tested whether overexpression of Ntg1 in budding yeast causes a growth phenotype. In parallel with studies of NTHL1 [44], we examined overexpression of both wild-type Ntg1 and a catalytically dead variant of Ntg1 (K243Q) [48] we term ntg1$_{catdead}$ [46, 49]. We expressed these Ntg1 proteins from a galactose-inducible plasmid [57] to allow regulated expression. Wild-type *S. cerevisiae* cells were transformed and grown on control glucose plates or galactose plates to induce expression of Ntg1 or ntg1$_{catdead}$. As controls, we employed Vector alone and a subunit of the RNA exosome, Rrp44 [58], which impairs cell growth when overexpressed. The Rrp44 subunit of the RNA exosome is a 3'-5' exonuclease/endonuclease

that mediates RNA decay and processing [58, 59]. We reasoned that overexpression of this RNA processing factor impairs cell growth through mechanisms distinct from Ntg1 overexpression and thus could serve as a control for efficient galactose-mediated induction under different growth conditions.

Cells that overexpress Ntg1 show a mild impairment of growth when compared to control cells with Vector alone (Figure 3.1A). Cells that express ntg1^{catdead} also show slow growth (Figure 3.1A) compared to the Vector control cells. To provide a quantitative measure of growth defects in cells that overexpress Ntg1 or ntg1^{catdead}, we conducted a liquid growth assay. As shown in Figure 3.1B, cells that overexpress either Ntg1 or ntg1^{catdead} show slower growth than control cells with Vector alone. Notably, overexpression of ntg1^{catdead} causes slower growth than overexpression of Ntg1. The difference in growth between cells overexpressing wild-type Ntg1 and ntg1^{catdead} could be explained by a difference in the level of expression achieved. To assess whether Ntg1 and ntg1^{catdead} are expressed at similar levels, wild-type cells containing a galactose inducible myc-tagged Ntg1 or ntg1^{catdead} were grown in media containing galactose, samples were collected at the indicated time points, and analyzed by immunoblotting. As shown in Figure 3.1C, Ntg1 and ntg1^{catdead} are expressed at approximately equal levels.

The growth assays cannot distinguish whether the delay in growth caused by overexpression of Ntg1 and ntg1^{catdead} is due to a cytostatic effect, where cell growth is arrested, or a cytotoxic effect, where cell viability is lost. To assess whether Ntg1 overexpression is cytostatic or cytotoxic, we conducted a viability test by inducing the expression of Ntg1 or ntg1^{catdead} with galactose overnight. After induction, cells were plated on plates containing glucose to determine the number of viable cells present in the culture. Colony forming units were counted, averaged, and compared to cells expressing the control Vector alone. The percent viability of cells

expressing Vector control was set to 100%. The percent viability of colonies overexpressing Ntg1 (40%, p-value of 0.0004) or ntg1$_{catdead}$ (59%, p-value of 0.0001), is significantly decreased when compared to Vector (Figure 3.1D). The viability of colonies overexpressing Ntg1 is significantly different (p-value of 0.0236) from the percent viability for cells overexpressing ntg1$_{catdead}$. Thus, overexpression of Ntg1 and ntg1$_{catdead}$ impair cell growth and induce cell death with a greater effect in cells overexpressing catalytically active Ntg1 as compared to a catalytically inactive variant of Ntg1.

3.4.2 Overexpression of Ntg1 causes DNA double-strand breaks and chromosome loss in *S. cerevisiae.*

Previous work showed that overexpression of NTHL1 causes accumulation of double-strand breaks [44]. To test for induction of double-strand breaks in the budding yeast model overexpressing Ntg1, we employed an *S. cerevisiae* reporter system expressing Rad52-YFP which forms foci at sites of double-strand breaks [50]. We overexpressed Ntg1 or ntg1$_{catdead}$ in these Rad52-YFP cells and analyzed the number of foci that form compared to control cells with Vector alone (Figure 3.2A). In Figure 3.2B, we quantified the percent of cells with foci. In cells with control Vector, this value was 0.11%, with cells expressing Ntg1 at 5% (p-value of 0.016), and ntg1$_{catdead}$ at 1.3% (p-value of 0.258). The difference in double-strand break foci detected in cells that overexpress ntg1$_{catdead}$ is statistically different from that of foci produced in cells that overexpress wild-type Ntg1 (p-value of 0.03), consistent with a requirement for the catalytic activity of Ntg1 to cause accumulation of double-strand breaks.

As double-strand breaks can lead to loss of genetic material and previous analysis of overexpression of NTHL1 showed an accumulation of micronuclei [44], we extended this analysis

to analyze loss of whole chromosomes. To test for chromosome loss, we employed a colorimetric sectoring assay[98,99]. We overexpressed Ntg1 or ntg1^{catdead} in an *S. cerevisiae* reporter strain containing a covering chromosome fragment that when lost results in cells producing a red pigment [51, 56]. Figure 3.2C shows the percent of colonies that lost the chromosome fragment during the first division resulting in a half-sectored colony. We quantified the percent of half-sectored colonies in cells with control Vector (0.2%), Ntg1 (1.8%, p-value of 0.037) and ntg1^{catdead} (0.2%). The difference in colonies that half-sectored in cells that overexpress Ntg1 is statistically different from that of half-sectoring colonies produced in ntg1^{catdead} (p-value of 0.043). Thus, overexpression of Ntg1 but not ntg1^{catdead} results in an increase in both double-strand breaks and whole chromosome loss.

3.4.3 Interplay of base excision repair with DNA damage response pathways.

As we have established that overexpression of yeast Ntg1 causes a growth phenotype and exhibits similar DNA damage phenotypes as detected for overexpression of NTHL1 [44], we next exploited the yeast deletion collection to interrogate pathways involved in DNA damage response for genetic interactions with overexpression of Ntg1. For these studies, we employed a serial dilution and spotting assay to assess relative effects on cell growth. As described in Materials and Methods, this serial dilution assay employs 5-fold serial dilutions. This approach means that any change in growth between adjacent spots of cells on the plate reflects a five-fold change in growth. We selected a panel of deletion mutants from the yeast deletion collection and overexpressed either Ntg1 or ntg1^{catdead}. We used the Vector alone as the control as well as overexpression of Rrp44. We reasoned that by setting growth in cells expressing Vector alone to a standard for each mutant analyzed, we could account for any differences in growth between wild-type cells and the deletion

mutants analyzed. We employed the Rrp44 control to ensure that galactose induction is functional in each of the mutants analyzed. This approach could be employed to identify pathways that interact with overexpression of Ntg1, but also to infer which pathways might be employed for cells to respond to Ntg1-induced damage by identifying those pathways that when impaired make cells more susceptible to overexpression of Ntg1 and/or ntg1$_{catdead.}$ Such an analysis is readily performed using the budding yeast system.

Figure 3.3 shows examples of the serial dilution growth assays (Figure 3.3A,B) that were performed as well as a heat map that summarizes the complete set of deletion mutants analyzed (Figure 3.3C). Results of the serial dilution growth assay for a deletion mutant, *Δtel1*, that shows no genetic interaction with overexpression of either Ntg1 or ntg1$_{catdead}$ (Figure 3.3A) and a deletion mutant, *Δrad51* (Figure 3.3B), that is sensitive to overexpression of wild-type Ntg1. Figures 3.3A and 3B show wild-type cells on the top panel and the deletion mutant cells on the bottom panel. The *Δtel1* cells show comparable growth to wild-type with the Vector alone control with no change in cell growth that shows any detectable difference from the wild-type control cells for either Ntg1 or ntg1$_{catdead}$. Figure 3.3B shows a change observed in the growth of *Δrad2* cells that overexpress Ntg1, as evidenced by the lack of growth detected in the fourth spot in the *Δrad2* cells as compared to wild-type cells, with no detectable change for overexpression of ntg1$_{catdead}$ compared to the wild-type control.

We employed these 5-fold serial dilution growth assays to develop a semi-quantitative scale, allowing us to analyze and compare results for a panel of deletion mutants. If growth was detected in all spots, this was scored a 10 on the growth scale. Growth in four spots was scored an 8, continuing to growth in a single spot set to 2. We employed a scale of 1-10 to allow for some subjective analysis of how much growth was evident in the most dilute spot where growth was

detected. For all mutants analyzed, we used conditions where growth could be detected in all five spots in the Vector control sample, providing a comparable scale for all deletion mutants regardless of any growth defect present in the mutant cells. In addition, we focused our conclusions on mutants where overexpression of the galactose-inducible control, Rrp44, did not markedly change growth as compared to the wild-type cells as a control to ensure that any differences detected were not due to altered galactose-mediated induction. All results presented in Figure 3.3C represent the consensus result obtained from three independent growth assays.

The heat map shown in Figure 3.3C summarizes the data from a panel of deletion mutants analyzed. In this heat map, the growth of cells with Vector control was set to 10 (bright blue). All deletion mutants should be compared to the top row which shows the effect of overexpressing Ntg1 (7/8, dark blue), ntg1$_{catdead}$ (5/6, gold) or the control Rrp44 (4/5, yellow gold) on this growth scale in wild-type cells. Results are shown for 22 different deletion mutants, which were divided into categories based on function. We focused on these 22 mutants to provide a representation of different DNA repair relevant pathways and illustrate how this system could be used to survey a variety of these pathways.

Based on the data summarized in Figure 3.3C, overexpression of both Ntg1 and ntg1$_{catdead}$ causes slow growth comparable to that detected in wild-type cells for several of the mutants analyzed, including *Δapn1*, *Δyku70*, *Δyku80*, *Δtel1*, *Δexo1* and *Δcrt1*. Several mutants are sensitive to overexpression of both Ntg1 and ntg1$_{catdead}$ with the most striking example the nucleotide excision repair pathway mutant, represented by *Δrad1*. We also identified pathways that are more sensitive to overexpression of wild-type Ntg1 than ntg1$_{catdead}$. For example, both *Δrad51* and *Δrad50* cells show a stronger growth defect when Ntg1 is overexpressed as compared to wild-type cells, but no significant change as compared to wild-type cells with overexpression of ntg1$_{catdead}$.

As both Rad50 and Rad51 are key components of homologous recombination [60], this result suggests that cells with impaired homologous recombination are sensitive to overexpression of catalytically active Ntg1. Interestingly, we did not identify any mutants that show sensitivity to overexpression of ntg1$_{catdead}$ without a concomitant effect for Ntg1. This finding is consistent with the model that overexpression of Ntg1/NTHL1 can impair cell growth and cause replication stress through at least two mechanisms, one of which results from the catalytic activity of Ntg1/NTHL1 and one that results from altered protein/protein interactions [44]. Catalytically active Ntg1/NTHL1 could mediate both of these effects, while ntg1$_{catdead}$ would only impact protein/protein interactions. The broad pathways that show sensitivity to overexpression of Ntg1 and/or ntg1$_{catdead}$ are homologous recombination and nucleotide excision repair.

Among the pathways that display some resistance to overexpression of Ntg1 and/or ntg1$_{catdead}$, *Δmlp2* cells are more resistant to overexpression of ntg1$_{catdead}$ than to overexpression of the control Rrp44 (Figure 3.3C). The Mlp2 protein is a component of the SUMO-mediated DNA damage response pathway [61], raising the possibility that SUMO modification of a protein could modulate the repair pathways by which cells respond to overexpression of Ntg1.

3.4.4 SUMO modification modulates the effect of Ntg1 overexpression.

A potential link to SUMO modification in cellular responses to overexpression to Ntg1 is intriguing because previous work demonstrated that Ntg1 is sumoylated, raising the possibility that sumoylation of Ntg1 could regulate Ntg1-mediated DNA damage [46]. To further investigate a potential role for sumoylation in cellular response to overexpression of Ntg1, we employed a yeast mutant lacking both SUMO E3 ligases, Siz1 and Siz2 (*Δsiz1/Δsiz2*) [53, 54]. The results of this analysis show that overexpression of both Ntg1 and ntg1$_{catdead}$ in cells lacking global

sumoylation causes a modest decrease in cell growth on the galactose plates relative to the wild-type control cells (Figure 3.4A). We also tested overexpression of Ntg1 in yeast cells with impaired desumoylase activity, by employing a temperature sensitive *ULP1* mutant (*ulp1-1*) [52]. The *ulp1-1* cells grow poorly on galactose plates; however, overexpression of Ntg1 impairs cell growth relative to Vector alone (Figure 3.4B). These data further suggest that SUMO-mediated interactions can contribute to the growth phenotypes caused by overexpression of Ntg1.

As Ntg1 can be modified by SUMO on multiple lysine residues (Figure 3.4C) [46], we directly tested whether sumoylation of Ntg1 impacts the overexpression phenotype. To address this question, we exploited a nonsumoylatable Ntg1 variant (ΔSUMO) where five lysines are changed to the conserved but nonsumoylatable residue arginine (K->R) [46]. Overexpression of this nonsumoylatable Ntg1 causes more of a growth defect than wild-type Ntg1 (Figure 3.4D). In contrast, a variant of Ntg1 that lacks SUMO modification and catalytic activity ($\Delta SUMO_{catdead}$) does not impair cell growth to the same extent as overexpression of $ntg1_{catdead}$ (Figure 3.4D). These results suggest that SUMO modification of Ntg1 contributes to the overexpression phenotype of Ntg1.

3.5 Discussion

In this study, we established a budding yeast model that can be used to explore how cells respond to overexpression of Ntg1 and potentially extended to define how overexpression of NTHL1 could contribute to cancer phenotypes. We present evidence that overexpression of catalytically active Ntg1 causes an increase in both double-strand breaks and chromosome loss. In contrast, while overexpression of a catalytically inactive form of Ntg1 also impairs cells growth when overexpressed, no increase in double-strand breaks or chromosome loss is evident. These

results support a model where overexpression of Ntg1 can alter cell physiology through at least two mechanisms, one that depends on the enzymatic activity and one that does not. We exploited this yeast genetics system to define cellular pathways that display genetic interactions with overexpression of Ntg1. As illustrated in Figure 3.5, these genetic interactions identify potential interactions with DNA repair pathways that are involved in NER (Rad1), double-strand break repair (Rad50, Rad51), and SUMO-mediated DNA damage response (Mlp2).

The evolutionarily conserved BER proteins human NTHL1 and *S. cerevisiae* Ntg1 are functionally and mechanistically similar [10]. As described here, the similarities also extend to consequences of overexpression. Overexpression of NTHL1 results in an increase in DNA damage, DNA double-strand breaks, and micronuclei formation [44]. Here, utilizing budding yeast, we show that overexpression of Ntg1 in yeast also causes double-strand breaks and chromosome loss. In humans and yeast, overexpression of both NTHL1 and Ntg1 genetically interact with double-strand break repair pathways homologous recombination and non-homologous end joining [44].

As with overexpression of human NTHL1, results from this budding yeast model suggest that overexpression of Ntg1 can impact cellular physiology through at least two pathways. First, overexpression of catalytically active Ntg1 can lead to biological endpoints that likely result from accumulation of nicks in the DNA backbone. These endpoints include double-strand breaks and chromosome loss. However, a catalytically inactive Ntg1 variant ($ntg1_{catdead}$) can also impair cell growth and previous work has demonstrated that this Ntg1 variant is not able to introduce nicks into the DNA backbone even in the presence of DNA damage [49]. As overexpression of Ntg1 even in the absence of catalytic activity can impair yeast cell growth, we speculate that Ntg1 can interact with other proteins to alter key protein-protein interactions. This model was also suggested

for the NTHL1 protein [44], building on evidence that NTHL1 physically interacts with components of the NER pathway [9, 62], including XPG [63]. We did not detect genetic interactions between overexpression of either wild-type Ntg1 or ntg1$_{catdead}$ and the yeast mutant *Δrad2* (Figure 3.3A). *RAD2* encodes the budding yeast orthologue of XPG [64]. While this result may seem counterintuitive, if overexpression of the Ntg1 protein sequesters Rad2, cells lacking Rad2 would not be subject to any additional biological effects if Rad2 were already absent. Thus, the lack of genetic interaction with *Δrad2* remains consistent with a model where overexpression of Ntg1 could sequester components of the NER pathways, supporting the possibility of crosstalk between these two pathways [30].

An ideal cross-species system that would take advantage of the yeast genetics approach, but employ human NTHL1, would be overexpression of NTHL1 in budding yeast. We attempted this approach but could not drive high expression of NTHL1 in *S. cerevisiae* despite evaluating a number of promoters and plasmids. Low levels of an epitope-tagged NTHL1 could be detected when a similar galactose-inducible approach to that described here for Ntg1 was employed (data not shown). However, no effect on cell growth was detected and some mutants that are sensitive to overexpression of Ntg1 were not sensitive to the levels of NTHL1 expression that could be achieved. As an alternative approach to define pathways sensitive to overexpression of NTHL1, we have developed doxycycline-inducible, non-tumorigenic human cell lines where expression of NTHL1 can be modulated. These systems have been established both in human bronchial epithelial cells (HBEC) cells [65] and in a breast epithelial cell line (MCF10A) [66]. A powerful approach will be to employ the yeast system described here to develop hypotheses that can be tested in these cell culture systems and eventually in mouse models.

Taken together, results from both budding yeast and human cells suggest that regulation of BER is critically important to maintain genome integrity. Either too much or too little BER activity can be detrimental to cells, demonstrating that cells require a carefully regulated level of BER activity [67].

3.6 Figures and tables

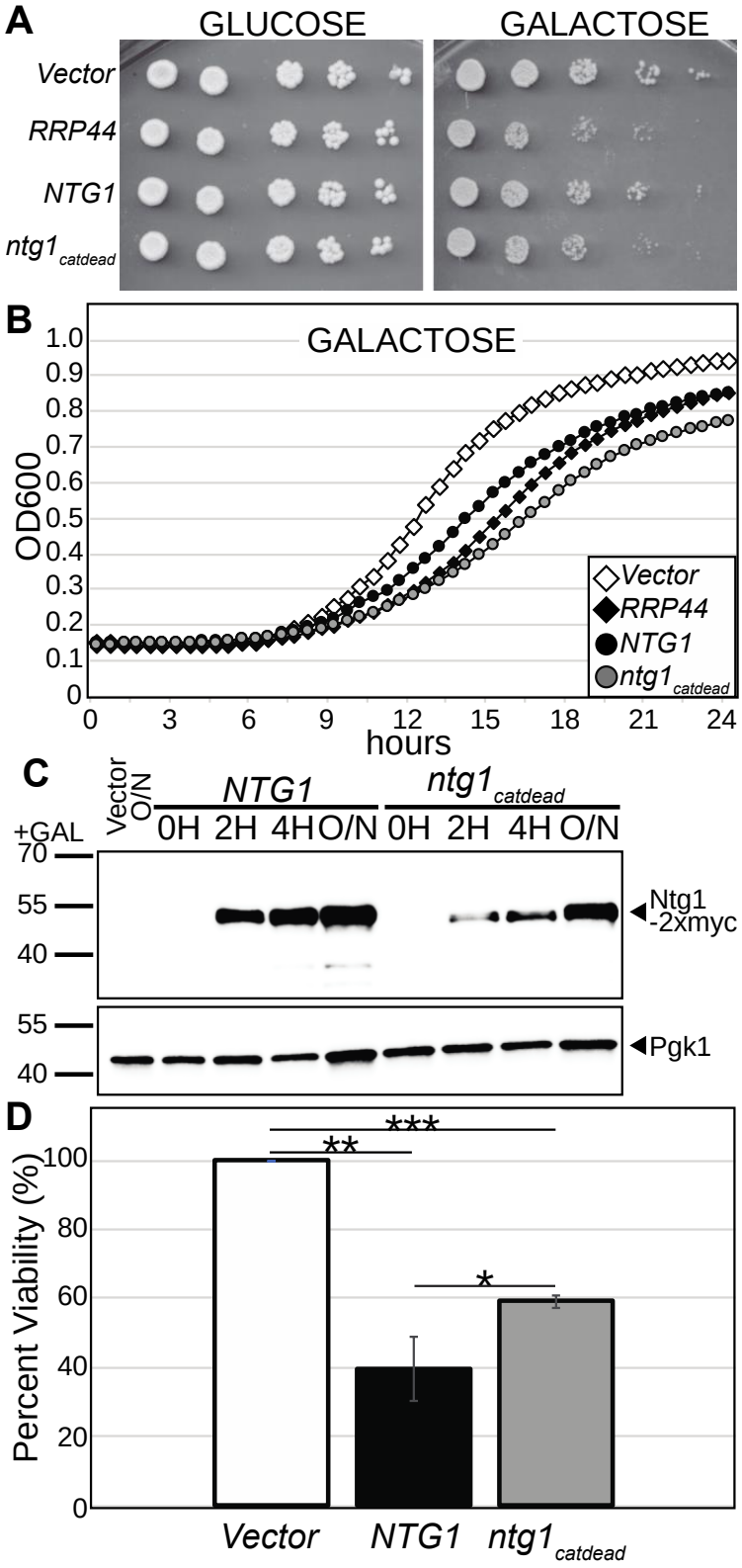


Figure 3.1: Overexpression of Ntg1 impairs *S. cerevisiae* cell growth.

Wild-type cells were transformed with negative control galactose-inducible Vector (Vector), positive control galactose-inducible *RRP44*-2xmyc (*RRP44*), galactose-inducible *NTG1*-2xmyc (*NTG1*), or galactose-inducible catalytically inactive *ntg1*$_{catdead}$-2xmyc (ntg1$_{catdead}$). Cultures were grown overnight in media lacking uracil with raffinose at 30°C. A) Overnight cultures were 5-fold serially diluted and spotted on plates lacking uracil with glucose or galactose and the plates were incubated at 30°C. Pictures were taken on day 2. B) Overnight cultures were diluted into media lacking uracil with galactose and quantitative growth curve analysis was performed. OD600 readings were taken every 30 minutes and plotted vs time. Each culture appears on the graph as follows: Vector (white diamond), Rrp44 (black diamond), wild-type Ntg1 (black circle), and ntg1$_{catdead}$ (grey circle). C) Overnight cultures were diluted into media containing galactose and samples were collected at 0, 2, 4 hours and overnight (ON). Cells were lysed and lysate was subjected to immunoblotting to detect Ntg1-2xmyc and ntg1$_{catdead}$-2xmyc (Ntg1-2xmyc) and Pgk1 (Pgk1) serves as a loading control. Results shown in (A, B, and C) are representative of at least three independent experiments. D) Overnight cultures were diluted into media lacking uracil with galactose for 16 hours and plated on plates lacking uracil plus glucose and incubated for 4 days. At the end of 4 days, the colony forming units were counted. The viability of each sample was normalized to control Vector and is expressed as a percentage. The white circles denote the average percent viability. The * indicates a p-value of < 0.05 and ** is a p-value of < 0.005.

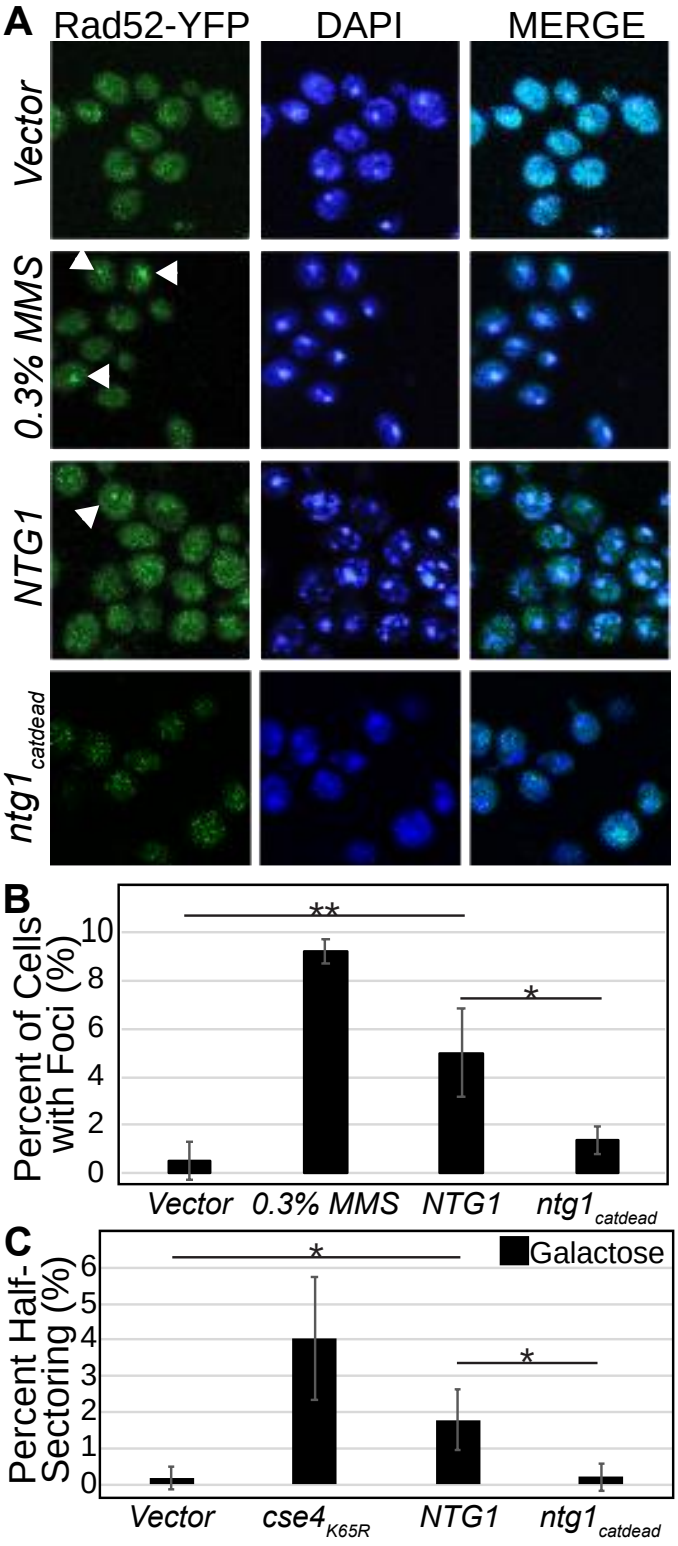


Figure 3.2: Overexpression of Ntg1 causes DNA double-strand breaks and chromosome loss in *S. cerevisiae*, however, overexpression of ntg1$_{catdead}$ does not cause chromosome loss.

Cells expressing Rad52-YFP [50] were transformed with negative control galactose-inducible Vector (Vector), galactose-inducible NTG1-2xmyc (NTG1), or galactose-inducible ntg1$_{catdead}$-2xmyc (ntg1$_{catdead}$). Cultures were grown overnight in media lacking uracil with raffinose. Samples were dosed with galactose and, as a positive control, Vector was dosed with galactose and 0.3% MMS (0.3% MMS). These cultures were incubated at 30°C for two hours. DAPI staining was utilized to visualize chromatin within the nucleus. MMS was inactivated and cells were fixed. A) Representative images from each sample show YFP (YFP), DAPI (DAPI), and merged images (MERGE). B) For each sample, 300 cells were analyzed for foci and the data are presented as the percent of cells analyzed that show foci. The results shown are the average of three independent experiments. Error bars represent standard deviation. The * indicates a p-value of < 0.05 and ** is a p-value of < 0.01.C) Cells containing a colorimetric reporter chromosome to measure chromosome loss (YPH1018 [51]), were transformed with negative control galactose-inducible Vector (Vector), positive control galactose-inducible 6His-3HA-cse4K65R (*cse4K65R*), galactose-inducible NTG1-2xmyc (NTG1), or galactose-inducible ntg1$_{catdead}$-2xmyc (ntg1catdead). Cultures were grown overnight in media lacking uracil with limited adenine and raffinose at 30°C. Cells were plated on plates lacking uracil with limited adenine and either glucose, or raffinose and galactose and the plates were incubated for 5 days at 30°C. C. C) For each sample, 300 cells were analyzed for chromosome loss and the data are represented as the percentage of cells showing chromosome loss in this assay as indicated by the number of half-sectoring colonies present in each sample, indicative of chromosome loss that occurs at the first cell division. An * indicates a p-value of < 0.05. The results shown are the average of three independent experiments.

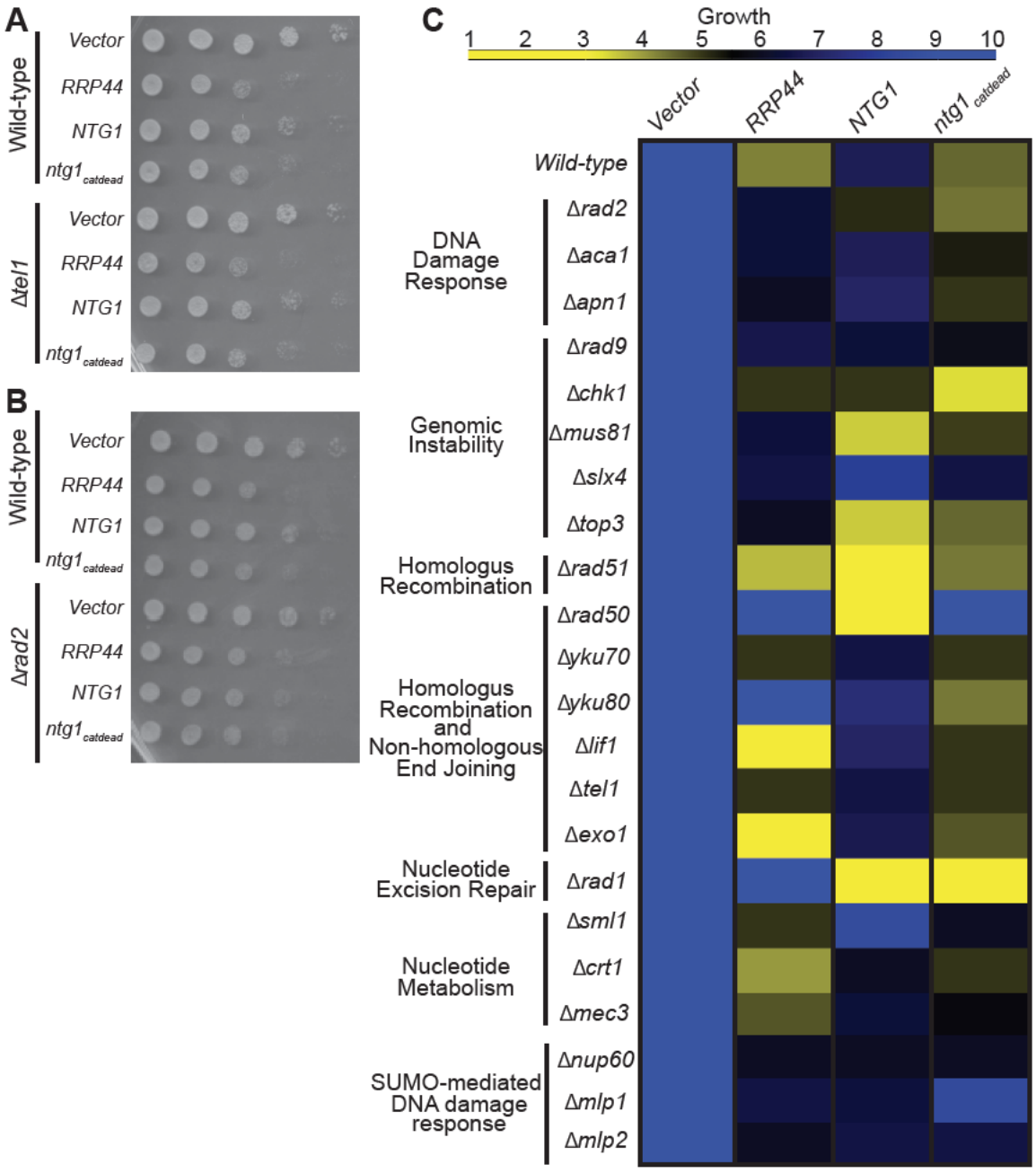


Figure 3.3: Interplay of base excision repair with DNA damage response pathways.

Wild-type or a panel of S. cerevisiae deletion mutant cells were transformed with a negative control galactose-inducible Vector (Vector), positive control galactose-inducible *RRP44*-2xmyc (*RRP44*), galactose-inducible *NTG1*-2xmyc (*NTG1*), or galactose-inducible *ntg1*$_{catdead}$-2xmyc (*ntg1*$_{catdead}$). Cultures were grown overnight in media lacking uracil with raffinose. Overnight

cultures were 5-fold serially diluted and spotted on plates lacking uracil with glucose or galactose and the plates were incubated at 30°C. Pictures were taken on day 2. Representative images from the growth assays that were employed to assess pathway interaction are shown for a mutant that (A) shows no change in growth compared to wild-type cells (*Δtel1*) and for (B) a deletion mutant (*Δrad51*) that shows increased sensitivity to overexpression of Ntg1 but not ntg1$_{catdead}$ (*Δrad51*). C) This approach was employed to generate a heat map for a panel of deletion mutants involved in DNA damage response pathways. As described in Materials and Methods, the scale for the heat map, which is shown at the top, ranges from 1 to 10. A score of 10 means growth is detected in all five of the serially diluted spots. A score of 1 means only poor growth was detected in the most concentrated spot. “Normal growth”, which is growth detected across the serial dilution in all spots, corresponding to a score of 10, is indicated by a bright blue color on the scale. Each mutant is normalized to the respective growth of the Vector expressed in that mutant background. Only those mutants that show no change in sensitivity to overexpression of the control RNA exosome subunit, Rrp44, were considered as showing a genetic interaction with Ntg1 and/or Ntg1$_{catdead}$. The data compiled in the heat map are representative of at least three independent experiments for each deletion mutant.

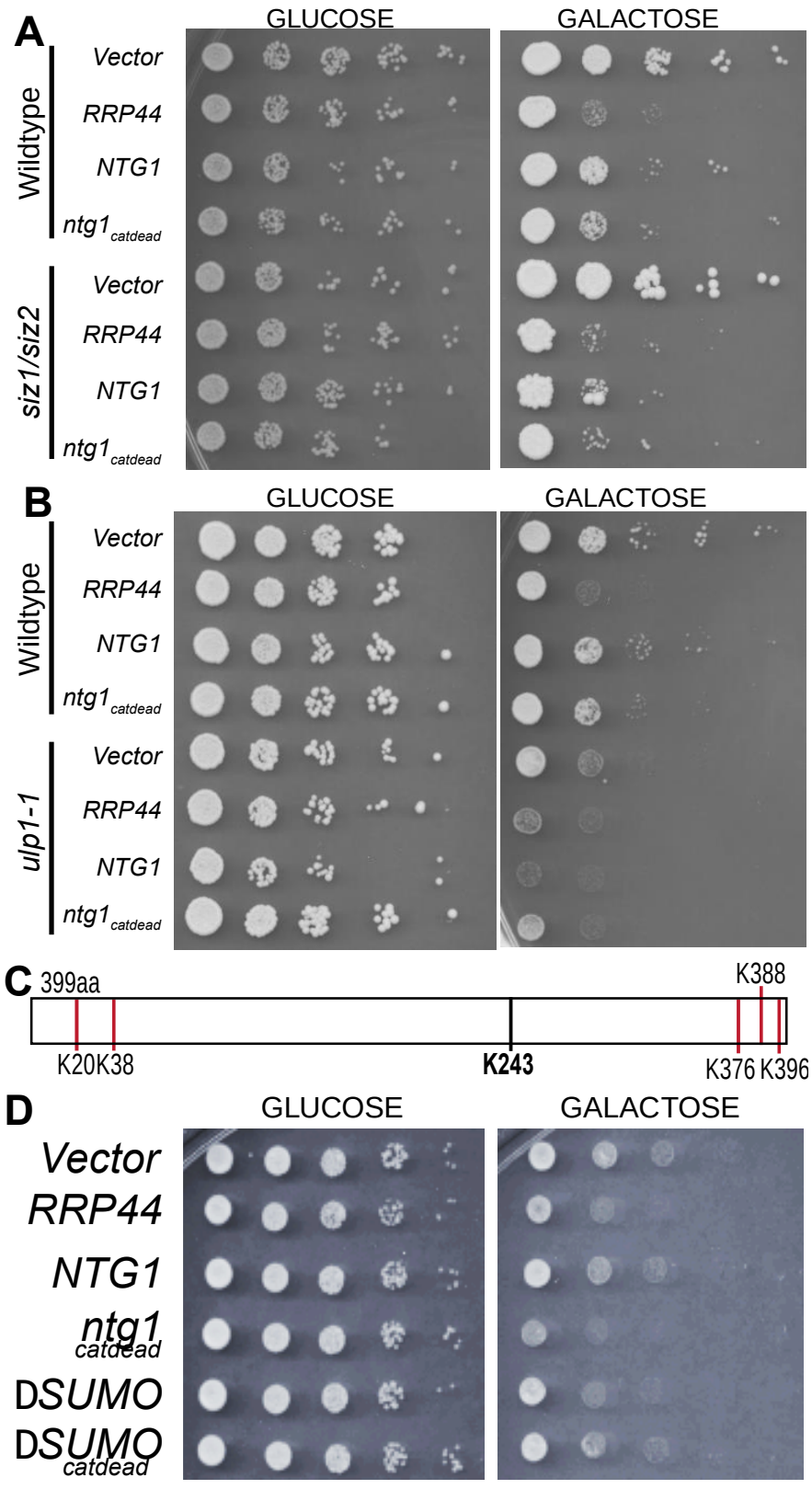


Figure 3.4: SUMO modification modulates the effect of Ntg1 overexpression.

Either (A) *Δsiz1/2* or (B) *ulp1-1* cells were transformed with negative control galactose-inducible Vector (Vector), positive control galactose-inducible R*RP44*-2xmyc (*RRP44*), galactose-inducible *NTG1*-2xmyc (*NTG1*), or galactose-inducible *ntg1*$_{catdead}$-2xmyc (*ntg1*$_{catdead}$). Cultures were grown overnight in media lacking uracil with raffinose at 30°C. Overnight cultures were 5-fold serially diluted and spotted on plates lacking uracil with glucose or galactose and the plates were incubated at 30°C. Pictures were taken on day 2. Results shown are representative of at least three independent experiments. C) Schematic depicting the five sumoylation sites and catalytic lysine of Ntg1 [46]. The SUMO modification sites are denoted by a red line (K20, 38, 376, 388, and 396), and the catalytic lysine is indicated by a black line (K243). D) Wild-type yeast were transformed with galactose-inducible plasmids, negative control Vector (Vector), positive control *RRP44*-2xmyc (*RRP44*), *NTG1*-2xmyc (*NTG1*), ntg1$_{catdead}$-2xmyc (*ntg1*$_{catdead}$) or a nonsumoylatable ntg1 variant with or without catalytic activity, *ntg1K20,38,376,388,396R* (*ΔSUMO*), or *ntg1K20,38,376,388,396R*$_{catdead}$ (*ΔSUMO*$_{catdead}$). Cultures were grown overnight in media lacking uracil with raffinose at 30°C. Overnight cultures were 5-fold serially diluted and spotted on plates lacking uracil with either glucose or galactose and the plates were incubated at 30°C. Pictures were taken on day 2. Results shown in (D) are representative of at least three independent experiments.

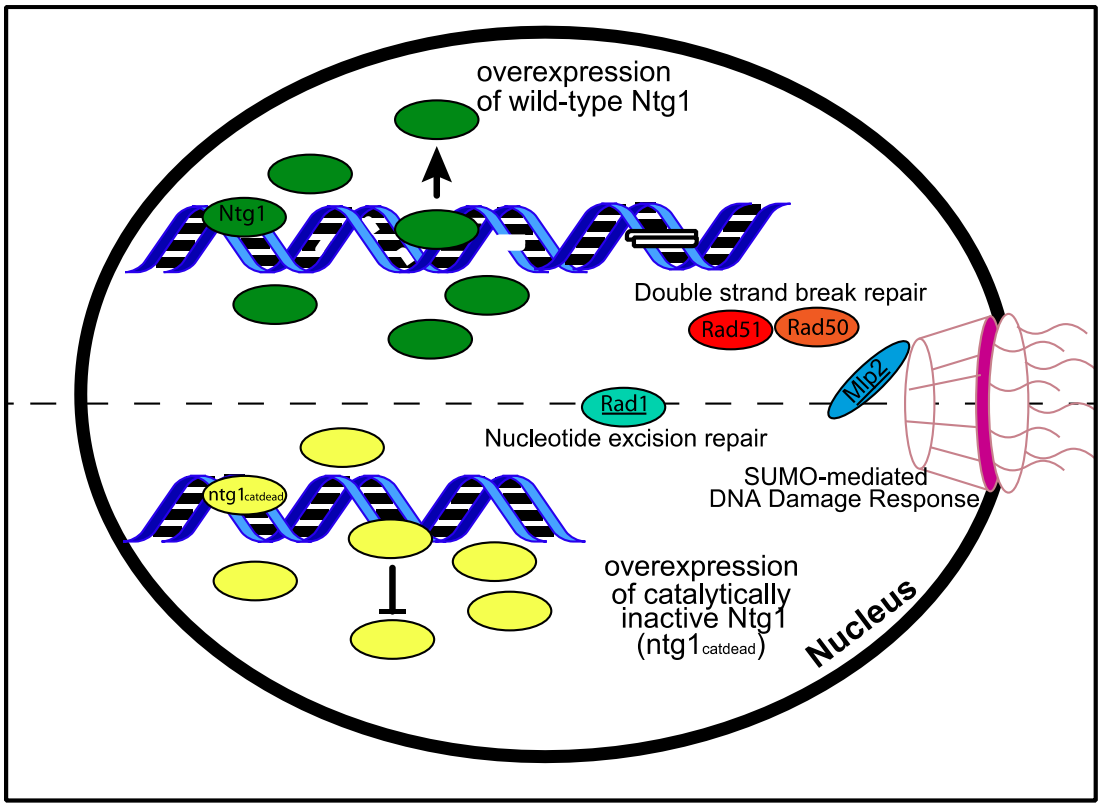


Figure 3.5: Model summarizing genetic interactions identified upon Ntg1 overexpression.

A model illustrates the functional consequences of overexpression of wild-type Ntg1 (top) and catalytically inactive Ntg1, ntg1$_{catdead}$. Cells lacking either Rad1 or Mlp2 (underlined) are hypersensitive to overexpression of both Ntg1 proteins. In contrast, only wild-type Ntg1 overexpression is toxic to cells lacking components of double-strand break repair (Rad50 or Rad51). These findings suggest that catalytically active Ntg1 has the potential to contribute to cellular phenotypes by introducing nicks into the DNA backbone that could accumulate and, ultimately, require cells to deploy homologous recombination to repair the damage. In contrast, overexpression of both wild-type and ntg1$_{catdead}$ could lead associated proteins to be sequestered and thus decrease the capacity of other repair pathways, including nucleotide excision repair and the SUMO-dependent pathways.

Table 3.1: Strains and plasmids

***S. cerevisiae* strains**			
Corbett Lab (ACY Strain #)	**Standard Name**	**Genotype**	**Source/Reference**
ACY402	BY4741	*MAT**a** his3Δ1 leu2Δ0 met15Δ0 ura3 Δ0*	Brachmann *et al.* [68]
ACY2638	rad52-YFP	*MAT***a** *ADE2 bar1::LEU2 RAD52-YFP*	Mayolo *et al.* [50]
ACY2639	YPH1018	*Mat**α** ura3-52 lys2-801 ade2-101 trp1Δ63 his3Δ200 leu2Δ1 CFIII (CEN3L.YPH278) HIS3 SUP11*	Au *et al.* [51]
ACY2640	YMB3468	*Mat**α** ura3-52 lys2-801 ade2-101 trp1Δ63 his3Δ200 leu2Δ1 (pSB817-pGAL1/10-MYC-cse4 K16R, 2μ, URA3) CFIII (CEN3L.YPH278) HIS3 SUP11*	Au *et al.* [51]
ACY2641	YOK262	*MAT***a** *trp1-Δ1 ura3-52 his3-Δ200 leu2-3,112 lys2-801*	Dohmen *et al.* [69]
ACY2642	EJY344	*MAT***a** *siz1Δ::LEU2 siz2Δ::TRP1*	Johnson *et al.* [54]
ACY2643	MHY1488	*MAT***a** *ulp1Δ::HIS3 LEU2::ulp1-333*	Li *et al.* [52]
ACY2644	EJY447	*MAT***a** *trp1-Δ1 ura3-52 his3-Δ200 leu2-3,112lys2-801 ulp2Δ::kanMX* [cir°]	Schwienhorst *et al.* [70]
Plasmids			
Corbett Lab (pAC#)	**Standard Name**	**Description**	**Source/Reference**
pAC18	pPS293	*pGAL1 URA3 2μ vector*	Gift of Pamela Silver
pAC3294		*pGAL1-Rrp44-myc*	This study
pAC3425		*pGAL1-Ntg1-2xmyc*	This study
pAC3535		*pGAL1-ntg1-catdead-2xmyc*	This study
pAC3536		*pGAL1-ntg1Δsumo-2xmyc*	This study
pAC3537		*pGAL1-ntg1Δsumo-catdead-2xmyc*	This study

3.7 Acknowledgments and funding sources

3.7.1 Acknowledgments

This study was supported in part by the Emory Integrated Genomics Core (EIGC), which is subsidized by the Emory University School of Medicine and is one of the Emory Integrated Core Facilities.. The content is solely the responsibility of the authors and does not necessarily reflect the official views of the National Institutes of Health.

3.7.2 Funding sources

This research project was supported in part by the Emory University Integrated Cellular Imaging Microscopy Core. Additional support was provided by the Georgia Clinical & Translational Science Alliance of the National Institutes of Health under Award Number UL1TR002378. ZJ was supported by the Emory IMSD grant (5R25GM125598). AMD was supported by NIH F31 (GM115178). PWD was supported by US National Institute of Health Intramural Research Program Project Z1AES103328.

3.8 References

[1] S. Maynard, S.H. Schurman, C. Harboe, N.C. de Souza-Pinto, V.A. Bohr, Base excision repair of oxidative DNA damage and association with cancer and aging, Carcinogenesis, 30 (2009) 2-10.

[2] F. Altieri, C. Grillo, M. Maceroni, S. Chichiarelli, Comprehensive Invited Review DNA Damage and Repair: From Molecular Mechanisms to Health Implications, Antioxidants \& Redox Signaling, 10 (2008) 891--937.

[3] C. Haag-Liautard, N. Coffey, D. Houle, M. Lynch, B. Charlesworth, P.D. Keightley, Direct Estimation of the Mitochondrial DNA Mutation Rate in Drosophila melanogaster, PLoS Biol, 6 (2008) 204.

[4] J. Miquel, An integrated theory of aging as the result of mitochondrial-DNA mutation in differentiated cells *, in, 1991, pp. 99--117.

[5] C. Richter, J.-W. Park, B.N. Amest, Normal oxidative damage to mitochondrial and nuclear DNA is extensive (8-hydroxydeoxyguanosine/aging/cancer/mutation), in, 1988, pp. 6465--6467.

[6] J.H.J. Hoeijmakers, DNA Damage, Aging, and Cancer, New England Journal of Medicine, 361 (2009) 1475--1485.

[7] J.F. Turrens, Mitochondrial formation of reactive oxygen species, The Journal of Physiology, 552 (2003) 335--344.

[8] L.H.F. Mullenders, Solar UV damage to cellular DNA: from mechanisms to biological effects, Photochemical & photobiological sciences : Official journal of the European Photochemistry Association and the European Society for Photobiology, 17 (2018) 1842-1852.

[9] N. Chatterjee, G.C. Walker, Mechanisms of DNA damage, repair, and mutagenesis, Environmental and molecular mutagenesis, 58 (2017) 235-263.

[10] N.C. Bauer, A.H. Corbett, P.W. Doetsch, The current state of eukaryotic DNA base damage and repair, Nucleic Acids Res, 43 (2015) 10083-10101.

[11] P.A. Riley, Free radicals in biology: oxidative stress and the effects of ionizing radiation, Int J Radiat Biol, 65 (1994) 27-33.

[12] T. Finkel, N.J. Holbrook, Oxidants, oxidative stress and the biology of ageing, Nature, 408 (2000) 239--247.

[13] J. Limón-Pacheco, M.E. Gonsebatt, The role of antioxidants and antioxidant-related enzymes in protective responses to environmentally induced oxidative stress, Mutation research, 674 (2009) 137-147.

[14] T. Lindahl, Instability and decay of the primary structure of DNA, Nature, 362 (1993) 709--715.

[15] C.G. Fraga, M.K. Shigenaga, J.-W. Park, P. Degant, B.N. Amest, Oxidative damage to DNA during aging: 8-Hydroxy-2'-deoxyguanosine in rat organ DNA and urine (cancer/mutation/endogenous DNA adducts/8-hydroxyguanine/oxygen radicals), in, 1990, pp. 4533--4537.

[16] K.B. Beckman, B.N. Ames, Oxidative Decay of DNA, Journal of Biological Chemistry, 272 (1997) 19633--19636.

[17] P. Boesch, F. Weber-Lotfi, N. Ibrahim, V. Tarasenko, A. Cosset, F. Paulus, R.N. Lightowlers, A. Dietrich, DNA repair in organelles: Pathways, organization, regulation, relevance in disease and aging, Biochimica et Biophysica Acta (BBA) - Molecular Cell Research, 1813 (2011) 186--200.

[18] Y.H. Wei, Oxidative Stress and Mitochondrial DNA Mutations in Human Aging, Experimental Biology and Medicine, 217 (1998) 53--63.

[19] P.A. Jeggo, L.H. Pearl, A.M. Carr, DNA repair, genome stability and cancer: a historical perspective, Nat Rev Cancer, 16 (2016) 35-42.

[20] V.A. Bohr, Repair of oxidative DNA damage in nuclear and mitochondrial DNA, and some changes with aging in mammalian cells,, Free Radical Biology and Medicine, 32 (2002) 804--812.

[21] T.W. O'Rourke, N.A. Doudican, M.D. Mackereth, P.W. Doetsch, G.S. Shadel, Mitochondrial dysfunction due to oxidative mitochondrial DNA damage is reduced through cooperative actions of diverse proteins., Molecular and cellular biology, 22 (2002) 4086--4093.

[22] Y.J. Kim, D.M. Wilson, 3rd, Overview of base excision repair biochemistry, Curr Mol Pharmacol, 5 (2012) 3-13.

[23] S.S. Wallace, Base excision repair: a critical player in many games, DNA Repair (Amst), 19 (2014) 14-26.

[24] Y.-J.a. Kim, Overview of Base Excision Repair Biochemistry COPING WITH ENDOGENOUS DNA DAMAGE, in, 2012.

[25] T. Lindahl, New class of enzymes acting on damaged DNA, Nature, 259 (1976) 64--66.

[26] T. Lindahl, DNA glycosylases, endonucleases for apurinic/apyrimidinic sites, and base excision-repair., Progress in nucleic acid research and molecular biology, 22 (1979) 135--192.

[27] S. Boiteux, M. Guillet, Mini review Abasic sites in DNA: repair and biological consequences in Saccharomyces cerevisiae, DNA Repair, 3 (2004) 1--12.

[28] T. Saha, J.K. Rih, R. Roy, R. Ballal, E.M. Rosen, Transcriptional Regulation of the Base Excision Repair Pathway by BRCA1, Journal of Biological Chemistry, 285 (2010) 19092--19105.

[29] N.s. Knudsen, S.D. Andersen, A. Ltzen, F.C. Nielsen, L.J. Rasmussen, Nuclear translocation contributes to regulation of DNA excision repair activities, DNA Repair, 8 (2009) 682--689.

[30] K.L. Limpose, A.H. Corbett, P.W. Doetsch, BERing the burden of damage: Pathway crosstalk and posttranslational modification of base excision repair proteins regulate DNA damage management, DNA Repair (Amst), 56 (2017) 51-64.

[31] M. Christmann, B. Kaina, Epigenetic regulation of DNA repair genes and implications for tumor therapy, Mutation research, 780 (2019) 15-28.

[32] V. Bandaru, S. Sunkara, S.S. Wallace, J.P. Bond, A novel human DNA glycosylase that removes oxidative DNA damage and is homologous to Escherichia coli endonuclease VIII, in, 2002, pp. 517--529.

[33] R. Aspinwall, D.G. Rothwell, T. Roldan-Arjona, C. Anselmino, C.J. Ward, J.P. Cheadle, J.R. Sampson, T. Lindahl, P.C. Harris, I.D. Hickson, Cloning and characterization of a functional human homolog of Escherichia coli endonuclease III., Proceedings of the National Academy of Sciences of the United States of America, 94 (1997) 109--114.

[34] K. Imai, A.H. Sarker, K. Akiyama, S. Ikeda, M. Yao, K. Tsutsui, T. Shohmori, S. Seki, Genomic structure and sequence of a human homologue (NTHL1/NTH1) of Escherichia coli endonuclease III with those of the adjacent parts of TSC2 and SLC9A3R2 genes, Gene, 222 (1998) 287-295.

[35] S. Ikeda, T. Biswas, R. Roy, T. Izumi, I. Boldogh, A. Kurosky, A.H. Sarker, S. Seki, S. Mitra, Purification and characterization of human NTH1, a homolog of Escherichia coli endonuclease III. Direct identification of Lys-212 as the active nucleophilic residue, J Biol Chem, 273 (1998) 21585-21593.

[36] D.R. Marenstein, M.K. Chan, A. Altamirano, A.K. Basu, R.J. Boorstein, R.P. Cunningham, G.W. Teebor, Substrate Specificity of Human Endonuclease III (hNTH1), Journal of Biological Chemistry, 278 (2003) 9005--9012.

[37] R.D. Weren, M.J. Ligtenberg, C.M. Kets, R.M. de Voer, E.T. Verwiel, L. Spruijt, W.A. van Zelst-Stams, M.C. Jongmans, C. Gilissen, J.Y. Hehir-Kwa, A. Hoischen, J. Shendure, E.A. Boyle, E.J. Kamping, I.D. Nagtegaal, B.B. Tops, F.M. Nagengast, A. Geurts van Kessel, J.H. van Krieken, R.P. Kuiper, N. Hoogerbrugge, A germline homozygous mutation in the base-excision repair gene NTHL1 causes adenomatous polyposis and colorectal cancer, Nat Genet, 47 (2015) 668-671.

[38] V. Larouche, A. Akirov, Co-occurrence of breast cancer and neuroendocrine tumours: New genetic insights beyond Multiple Endocrine Neoplasia syndromes, 2 (2019) e00092.

[39] M. Terradas, P.M. Munoz-Torres, Contribution to colonic polyposis of recently proposed predisposing genes and assessment of the prevalence of NTHL1- and MSH3-associated polyposes, 40 (2019) 1910-1923.

[40] A. Tubbs, A. Nussenzweig, Endogenous DNA Damage as a Source of Genomic Instability in Cancer, Cell, 168 (2017) 644-656.

[41] M. Goto, K. Shinmura, H. Igarashi, M. Kobayashi, H. Konno, H. Yamada, M. Iwaizumi, S. Kageyama, T. Tsuneyoshi, S. Tsugane, H. Sugimura, Altered expression of the human base excision repair gene NTH1 in gastric cancer, Carcinogenesis, 30 (2009) 1345--1352.

[42] S. Koketsu, T. Watanabe, H. Nagawa, Expression of DNA repair protein: MYH, NTH1, and MTH1 in colorectal cancer., Hepato-gastroenterology, 51 (2004) 638--642.

[43] D.G. Albertson, Gene amplification in cancer, Trends Genet, 22 (2006) 447-455.

[44] K.L. Limpose, K.S. Trego, Z. Li, S.W. Leung, A.H. Sarker, J.A. Shah, S.S. Ramalingam, E.M. Werner, W.S. Dynan, P.K. Cooper, A.H. Corbett, P.W. Doetsch, Overexpression of the base excision repair NTHL1 glycosylase causes genomic instability and early cellular hallmarks of cancer, Nucleic Acids Research, 46 (2018) 4515--4532.

[45] L. Eide, M. Bjrs, M. Pirovano, I. Alseth, K.G. Berdal, E. Seeberg, Base excision of oxidative purine and pyrimidine DNA damage in Saccharomyces cerevisiae by a DNA glycosylase with sequence similarity to endonuclease III from Escherichia coli., Proceedings of the National Academy of Sciences of the United States of America, 93 (1996) 10735--10740.

[46] D.B. Swartzlander, A.J. McPherson, H.R. Powers, K.L. Limpose, E.G. Kuiper, N.P. Degtyareva, A.H. Corbett, P.W. Doetsch, Identification of SUMO modification sites in the base excision repair protein, Ntg1, DNA Repair, 48 (2016) 51--62.

[47] H. Ito, Y. Fukuda, K. Murata, A. Kimura, Transformation of intact yeast cells treated with alkali cations., Journal of bacteriology, 153 (1983) 163--168.

[48] L. Augeri, K.K. Hamilton, A.M. Martin, P. Yohannes, P.W. Doetsch, Purification and properties of yeast redoxyendonuclease, Methods in Enzymology, 234 (1994) 102--115.

[49] D.B. Swartzlander, L.M. Griffiths, J. Lee, N.P. Degtyareva, P.W. Doetsch, A.H. Corbett, Regulation of base excision repair: Ntg1 nuclear and mitochondrial dynamic localization in response to genotoxic stress, Nucleic Acids Res, 38 (2010) 3963-3974.

[50] A.A. de Mayolo, I. Sunjevaric, R. Reid, U.H. Mortensen, R. Rothstein, M. Lisby, The rad52-Y66A allele alters the choice of donor template during spontaneous chromosomal recombination, DNA Repair, 9 (2010) 23--32.

[51] W.-C. Au, M.J. Crisp, S.Z. Deluca, O.J. Rando, M.A. Basrai, Altered Dosage and Mislocalization of Histone H3 and Cse4p Lead to Chromosome Loss in Saccharomyces cerevisiae, Genetics, 179 (2008) 263--275.

[52] S.-J. Li, M. Hochstrasser, A new protease required for cell-cycle progression in yeast, Nature, 398 (1999) 246--251.

[53] E.S. Johnson, Protein Modification by SUMO, Annual Review of Biochemistry, 73 (2004) 355--382.

[54] E.S. Johnson, A.A. Gupta, An E3-like factor that promotes SUMO conjugation to the yeast septins., Cell, 106 (2001) 735--744.

[55] J. Schindelin, I. Arganda-Carreras, E. Frise, V. Kaynig, M. Longair, T. Pietzsch, S. Preibisch, C. Rueden, S. Saalfeld, B. Schmid, J.-Y. Tinevez, D.J. White, V. Hartenstein, K. Eliceiri, P. Tomancak, A. Cardona, Fiji: an open-source platform for biological-image analysis, Nature Methods, 9 (2012) 676--682.

[56] P. Hieter, C. Mann, M. Snyder, R.W. Davis, Mitotic Stability of Yeast Chromosomes: A Colony Color Assay That Measures Nondisjunction and Chromosome Loss, in, 1985, pp. 381--392.

[57] M. Johnston, A model fungal gene regulatory mechanism: the GAL genes of Saccharomyces cerevisiae, Microbiological reviews, 51 (1987) 458-476.

[58] C. Schneider, J.T. Anderson, D. Tollervey, The exosome subunit Rrp44 plays a direct role in RNA substrate recognition, Mol Cell, 27 (2007) 324-331.

[59] C. Schneider, E. Leung, J. Brown, D. Tollervey, The N-terminal PIN domain of the exosome subunit Rrp44 harbors endonuclease activity and tethers Rrp44 to the yeast core exosome, Nucleic Acids Res, 37 (2009) 1127-1140.

[60] W.D. Wright, S.S. Shah, W.-D. Heyer, Homologous recombination and the repair of DNA double-strand breaks., The Journal of biological chemistry, 293 (2018) 10524--10535.

[61] C. Strambio-de-Castillia, G. Blobel, M.P. Rout, Proteins Connecting the Nuclear Pore Complex with the Nuclear Interior, Journal of Cell Biology, 144 (1999) 839--855.

[62] G. Spivak, Nucleotide excision repair in humans, DNA Repair (Amst), 36 (2015) 13-18.

[63] K.S. Trego, T. Groesser, A.R. Davalos, A.C. Parplys, W. Zhao, M.R. Nelson, A. Hlaing, B. Shih, B. Rydberg, J.M. Pluth, M.S. Tsai, J.H. Hoeijmakers, P. Sung, C. Wiese, J. Campisi, P.K. Cooper, Non-catalytic Roles for XPG with BRCA1 and BRCA2 in Homologous Recombination and Genome Stability, Mol Cell, 61 (2016) 535-546.

[64] A. O'Donovan, A.A. Davies, J.G. Moggs, S.C. West, R.D. Wood, XPG endonuclease makes the 3′ incision in human DNA nucleotide excision repair, Nature, 371 (1994) 432-435.

[65] R.D. Ramirez, S. Sheridan, L. Girard, M. Sato, Y. Kim, J. Pollack, M. Peyton, Y. Zou, J.M. Kurie, J.M. Dimaio, S. Milchgrub, A.L. Smith, R.F. Souza, L. Gilbey, X. Zhang, K. Gandia, M.B. Vaughan, W.E. Wright, A.F. Gazdar, J.W. Shay, J.D. Minna, Immortalization of human bronchial epithelial cells in the absence of viral oncoproteins, Cancer Res, 64 (2004) 9027-9034.

[66] H.D. Soule, T.M. Maloney, S.R. Wolman, W.D. Peterson, Jr., R. Brenz, C.M. McGrath, J. Russo, R.J. Pauley, R.F. Jones, S.C. Brooks, Isolation and characterization of a spontaneously immortalized human breast epithelial cell line, MCF-10, Cancer Res, 50 (1990) 6075-6086.

[67] J. Hassall, Old Nursery Stories and Rhymes, Blackie \& Son, 1904.

[68] C.B. Brachmann, A. Davies, G.J. Cost, E. Caputo, J. Li, P. Hieter, J.D. Boeke, Designer deletion strains derived from Saccharomyces cerevisiae S288C: a useful set of strains and plasmids for PCR-mediated gene disruption and other applications, Yeast (Chichester, England), 14 (1998) 115-132.

[69] R.J. Dohmen, R. Stappen, J.P. McGrath, H. Forrová, J. Kolarov, A. Goffeau, A. Varshavsky, An essential yeast gene encoding a homolog of ubiquitin-activating enzyme, J Biol Chem, 270 (1995) 18099-18109.

[70] I. Schwienhorst, E.S. Johnson, R.J. Dohmen, SUMO conjugation and deconjugation, Molecular & general genetics : MGG, 263 (2000) 771-786.

Chapter 4: Conclusions and future directions

4.1 Summary

DNA is under a constant barrage of DNA damaging agents from sources both exogenous and endogenous to the cell[1,7]. These damages can result in anything from silent mutations to genetic instability and can lead to tumorigenesis[7,36]. These assaults threaten the genetic integrity of the nuclear and mitochondrial DNA[4,5]. Cells have developed six major evolutionarily conserved repair pathways to combat these assaults[4]. One such repair pathway is BER which is the major pathway for repair of oxidative DNA damage[18,44,46,49]. Much is known about the biochemical mechanism by which BER is able to repair the DNA. The five major steps are: 1) excision; 2) incision; 3) end processing; 4) repair synthesis; and 5) ligation[18]. Many of these initial studies into the biochemical mechanism have been conducted in budding yeast, *S. cerevisiae*. Loss of DNA repair components often lead to tumorigenesis. Despite the importance of DNA repair pathways, not much is known about how these pathways are regulated. Moreover, how these repair pathways intersect potentially with each other is also unknown.

The work described in Chapter 2 of this dissertation defines the SUMO modification sites in the Ntg1 *S. cerevisiae* BER protein. I first show that Ntg1 can be multi-sumoylated in the presence of oxidative stress, and present evidence that the human counterpart to Ntg1, NTHL1, is also sumo-modified in the presence of oxidative stress. I identify the sumoylation sites on Ntg1 as lysines 20, 38, 376, 388, and 396, with lysine 396 being the major site of monosumoylation. All of these sites can be SUMO-modified, and not until all sites are changed to arginine can sumoylation of Ntg1 be ablated. We created a non-sumoylatable Ntg1 variant, ntg1ΔSUMO. Cells that express ntg1ΔSUMO as the only copy of Ntg1 in an otherwise BER and NER deficient background are less sensitive to MMS, which causes alkylation damage, compared to cells that

express wild-type Ntg1 in the same background. This finding suggests that Ntg1 sumoylation plays a role in communicating with cell cycle arrest factors in response to MMS-induced damage.

In Chapter 3, I show that overexpression of Ntg1 causes a slow growth phenotype, DSBs, chromosome loss, and cell death. Overexpression of catalytically inactive Ntg1, despite having a stronger growth phenotype, does not cause DSBs and chromosome loss. These data further the idea first established in our lab that Ntg1 functions in at least two distinct ways, one that requires catalytic activity, and one that is independent of the catalytic activity of Ntg1. I also developed a screen in budding yeast to quickly identify players of the DNA damage response (DDR) pathway that genetically interact with overexpression of Ntg1. This system accurately identified some previously known interactions and suggests a few novel genetic interactions. This system has already shown its value in this sampling of DDR genes, and will provide an interesting platform to jumpstart research into genetic interactions with BER in future studies. Ultimately, the research that stems from this genetic screen can help researchers better understand how NTHL1 interacts with the DDR pathways to create better personalized cancer therapies. We also investigated the role SUMO-mediated interactions have on the growth phenotype displayed in cells overexpressing Ntg1.

4.2 Sumoylation as a means of regulating BER

Several BER proteins are SUMO modified, such as yeast Ntg1 and Ntg2, and human PARP1, TDG, and PCNA[90,100,101]. However, very little is known about how the SUMO-modification impacts the function of repair proteins. Nuclear Ntg1 is SUMO modified endogenously, but the steady-state levels of SUMO modified Ntg1 increases in response to oxidative stress[90,102]. This increase in the SUMO modification in response to cellular stress

suggests that sumoylation of Ntg1 is important for Ntg1 regulation. Sumoylation can affect a protein in a number of ways, such as decreasing protein degradation, increasing enzyme turnover, and modulating protein-protein interactions[100,103].

Modification by a SUMO molecule can physically block the addition of the degradation signaling molecule, ubiquitin[100]. This could provide increased stability of a protein and allow Ntg1/NTHL1 time to remove lesions[100]. Regulating the rate at which an *N*-glycosylase, as the initiator of the repair pathway, processes DNA lesions is emerging as a critical way in which to regulate BER in higher organisms[85,88]. Sumoylation could potentially be a mechanism by which this regulation is achieved, as the enzymatic turnover of the human TDG glycosylase is increased when TDG is sumoylated[104–106]. While the SUMO molecule is not required for NTHL1 homodimerization, XPG, or APE binding *in vitro*, in an *in vivo* system a SUMO molecule could block NTHL1 binding with these proteins, and/or encourage other protein pairing[85,88]. The N-terminus of NTHL1 is a likely location for these protein-protein interactions to occur and Ntg1 has two confirmed SUMO sites at the N-terminus (K20 and K38) and NTHL1 has a number of putative SUMO sites within the 101aa N-terminus (K56, K60, K75, and K85)[102]. Additionally, complete loss of Ntg1 sumoylation, in an otherwise BER and NER deficient background, results in cells that are less sensitive to damage induced by MMS[102]. This suggests that Ntg1 SUMO modification is involved in signaling cell cycle arrest, possibly through protein-protein interaction. This is a promising area of potential regulation in which more studies should be conducted.

As levels of SUMO modification increase on Ntg1 in response to oxidative stress, we speculate that there may be an additional role for SUMO modification of Ntg1 that we have not yet defined[90,102]. A major role for SUMO modification is to regulate protein-protein interactions so SUMO modification may alter the protein binding partners of Ntg1. We performed a mass

spectrometry screen to compare interacting partners for Ntg1 and ntg1ΔSUMO in the absence and presence of oxidative stress. This analysis led to the identification of several candidate proteins that could show SUMO-dependent interactions with Ntg1. Among the proteins identified in this screen was Cst6, a basic leucine zipper (bZIP) transcription factor from the ATF/CREB family involved in stress-responsive regulatory networks[107]. In the mass spectrometry studies, interaction with Cst6 was completely abolished when Ntg1 could not be SUMO modified. While validating this interaction and others identified in the screen is challenging due to the lack of antibodies available for budding yeast proteins, such studies provide a platform for future research. We could employ epitope-tagged proteins to explore and validate interaction, but the low levels of the endogenous proteins for many of the proteins under investigation, including Ntg1, made us cautious about using such an overexpression system. Future studies could explore these interactions and exploit the ntg1ΔSUMO variant that we created for both genetic and biochemical approaches to better define the role of SUMO modification in responding to oxidative damage. In addition, future studies could map the sites of SUMO modification on human NTHL1 to perform similar functional studies in the mammalian system.

4.3 Regulation of NTHL1 and Ntg1 expression and activity

The mRNA levels transcribed from endogenous *NTG1* and *NTG2*, and the observed protein levels of endogenous NTHL1 in non-transformed human cells suggest these proteins are likely to be present at low levels in the absence of a cellular insult that causes DNA damage[69,78,87,94]. However, in response to DNA damage, these proteins need to be rapidly deployed to allow repair before damage becomes fixed in the genome[1,2,26,108]. This situation suggests a requirement of a high level of regulation of protein activity. Along the same lines, a large number of vastly diverse

tumors have increased *NTHL1* mRNA and NTHL1 protein expression and transient overexpression of NTHL1 in HBEC cells result in DNA damage and early hallmarks of cancer[91,94,109–112]. *In vitro* experiments show NTHL1 has 100-fold less catalytic activity compared to EndIII[70,113]. Our studies in yeast show that overexpression of Ntg1 results in a growth defect, DSBs, chromosome loss, and increased cell death[101].

However, very few of the elements that are involved in the critical regulation of NTHL1 at the expression and activity level have been identified. As there are so many levels and methods by which protein expression and activity can be regulated, this area is particularly of interest and may tie into the role of regulation by SUMO modification as mentioned above. An interesting future experiment would be to test whether the SUMO modification can, as predicted, increase Ntg1/NTHL1 catalytic turnover as seen in TDG[104–106]. Future studies could also investigate *Ntg1/NTHL1* mRNA and ultimately protein expression levels in response DNA damaging agents, including mtDNA damaging agents like antimycin A in combination with H_2O_2.

4.4 BER and repair pathway crosstalk

Despite decades of studies on the major DNA repair pathways, little is known about how these pathways interact within the cell, particularly at the mechanistic level. In recent years, however, information on DNA repair pathway interplay has begun to emerge. Evidence exists of components of BER interacting with components of four of the other five major DNA repair pathways, NER, MMR, DSBR, and damage tolerance[88,94,95,114]. NER and HR proteins (e.g. XPG and BRCA1) increase NTHL1 activity and expression, and recent work suggests that NTHL1 promotes NHEJ while suppressing HR in DSBR[88,94,95]. The YB-1 protein, implicated in nuclear BER and NER, is also involved in mitochondria MMR[55]. Some of the bypassed lesions of the

DNA damage tolerance pathway are repaired by BER, and Ntg1 has the ability to signal for cell cycle arrest when sumoylated[43,53,102]. Previous studies in budding yeast focused on repair of spontaneous DNA damage and discovered that multiple pathways with overlapping specificities are involved in the removal of, or tolerance to, such DNA damage in *S. cerevisiae*[115]. These findings suggest a highly coordinated response to protect cells from potentially deleterious DNA lesions, but the molecular mechanisms that allow and coordinate these interactions are just beginning to be defined. Together, these emerging observations suggests a vast and complex coordination of DNA repair where various systems can serve as backup systems for one another. Further studies need to be conducted to define these interactions and extend the work to explore how post-translational modifications such as sumoylation may play a role in regulating crosstalk between these pathways. Expanding our genetic screen to include more knockouts of DNA damage response genes and translating any data into human cells will help define the BER crosstalk network.

In this dissertation, I describe work done to advance the field of DNA repair pathway regulation. I specifically focus on the BER pathway and the *S. cerevisiae N*-glycosylase, Ntg1. Ultimately, the work here can help streamline studies in human cells as better understanding of the cellular genetic interplay will allow scientists to develop targeted personalized cancer therapeutics.

4.5 References

1. Friedberg, E. C. DNA damage and repair. *Nature* (2003). doi:10.1038/nature01408
2. Hoeijmakers, J. H. J. DNA Damage, Aging, and Cancer. *N. Engl. J. Med.* **361**, 1475–1485 (2009).
3. Lindahl, T. Instability and decay of the primary structure of DNA. *Nature* **362**, 709–715 (1993).
4. Boesch, P. *et al.* DNA repair in organelles: Pathways, organization, regulation, relevance in disease and aging. *Biochim. Biophys. Acta - Mol. Cell Res.* **1813**, 186–200 (2011).
5. Pinto, M. & Moraes, C. T. Mechanisms linking mtDNA damage and aging. *Free Radic. Biol. Med.* **85**, 250–8 (2015).
6. Miquel, J. *An integrated theory of aging as the result of mitochondrial-DNA mutation in differentiated cells *. Arch. Gerontol. Geriatr* **12**, (1991).
7. Bauer, N. C., Corbett, A. H. & Doetsch, P. W. The current state of eukaryotic DNA base damage and repair. *Nucleic Acids Res.* **43**, 10083–101 (2015).
8. Goldman, M. *Ionizing Radiation and Its Risks*. (1982).
9. Wallace, B. M. & Lasker, J. S. Awakenings ... UV light and HIV gene activation. *Science* **257**, 1211–2 (1992).
10. Wakasugi, M. *et al.* Nucleotide excision repair-dependent DNA double-strand break formation and ATM signaling activation in mammalian quiescent cells. *J. Biol. Chem.* **289**, 28730–7 (2014).
11. Tiwari, V. & Wilson, D. M. DNA Damage and Associated DNA Repair Defects in Disease and Premature Aging. *American Journal of Human Genetics* (2019). doi:10.1016/j.ajhg.2019.06.005

12. Jackson, S. P. & Bartek, J. The DNA-damage response in human biology and disease. *Nature* **461**, 1071–1078 (2009).

13. Caldecott, K. W. Mammalian DNA single-strand break repair: an X-ra(y)ted affair. *BioEssays* **23**, 447–455 (2001).

14. Téoule, R. Radiation-induced DNA Damage and Its Repair. *Int. J. Radiat. Biol. Relat. Stud. Phys.* **51**, 573–589 (1987).

15. Riley, P. A. Free Radicals in Biology: Oxidative Stress and the Effects of Ionizing Radiation. *Int. J. Radiat. Biol.* **65**, 27–33 (1994).

16. Maynard, S., Fang, E. F., Scheibye-Knudsen, M., Croteau, D. L. & Bohr, V. A. DNA damage, DNA repair, aging, and neurodegeneration. *Cold Spring Harb. Perspect. Med.* (2015). doi:10.1101/cshperspect.a025130

17. Turrens, J. F. Mitochondrial formation of reactive oxygen species. *J. Physiol.* **552**, 335–344 (2003).

18. Krokan, H. E. & Bjørås, M. Base excision repair. *Cold Spring Harb. Perspect. Biol.* **5**, a012583 (2013).

19. Chatterjee, N. & Walker, G. C. Mechanisms of DNA damage, repair, and mutagenesis. *Environ. Mol. Mutagen.* **58**, 235–263 (2017).

20. Bohr, V. A. Repair of oxidative DNA damage in nuclear and mitochondrial DNA, and some changes with aging in mammalian cells,. *Free Radic. Biol. Med.* **32**, 804–812 (2002).

21. Jin You, H. *et al.* Saccharomyces cereVisiae Ntg1p and Ntg2p: Broad Specificity N-Glycosylases for the Repair of Oxidative DNA Damage in the Nucleus and Mitochondria †. *Biochemistry* **38**, 11298–11306 (1999).

22. Lindahl, T. *Instability and decay of the primary structure of DNA*. (1993).

23. Faure, V., Saparbaev, M., Dumy, P. & Constant, J.-F. Action of multiple base excision repair enzymes on the 2 0-deoxyribonolactone q. (2005). doi:10.1016/j.bbrc.2005.01.082

24. Ranjha, L., Howard, S. M. & Cejka, P. Main steps in DNA double-strand break repair: an introduction to homologous recombination and related processes. doi:10.1007/s00412-017-0658-1

25. Fraga, C. G., Shigenaga, M. K., Park, J.-W., Degant, P. & Amest, B. N. *Oxidative damage to DNA during aging: 8-Hydroxy-2'-deoxyguanosine in rat organ DNA and urine (cancer/mutation/endogenous DNA adducts/8-hydroxyguanine/oxygen radicals)*. *Proc. Nadl. Acad. Sci. USA* **87**, (1990).

26. Boiteux, S. & Guillet, M. Mini review Abasic sites in DNA: repair and biological consequences in Saccharomyces cerevisiae. *DNA Repair (Amst).* **3**, 1–12 (2004).

27. Limón-Pacheco, J. & Gonsebatt, M. E. The role of antioxidants and antioxidant-related enzymes in protective responses to environmentally induced oxidative stress. *Mutat. Res. Toxicol. Environ. Mutagen.* **674**, 137–147 (2009).

28. Michael, F. & Van Houten, B. *Mitochondrial DNA damage is more extensive and persists longer than nuclear DNA damage in human cells following oxidative stress*. *Cell Biology* **94**, (1997).

29. Michikawa, Y., Mazzucchelli, F., Bresolin, N., Scarlato, G. & Attardi, G. *Aging-Dependent Large Accumulation of Point Mutations in the Human mtDNA Control Region for Replication*. *Source: Science, New Series* **286**, (1999).

30. Pakendorf, B. & Stoneking, M. MITOCHONDRIAL DNA AND HUMAN EVOLUTION. *Annu. Rev. Genomics Hum. Genet.* **6**, 165–183 (2005).

31. Kennedy, S. R., Salk, J. J., Schmitt, M. W. & Loeb, L. A. Ultra-Sensitive Sequencing Reveals an Age-Related Increase in Somatic Mitochondrial Mutations That Are Inconsistent with Oxidative Damage. *PLoS Genet.* **9**, (2013).

32. Corral-Debrinski, M. *et al.* Marked changes in mitochondrial DNA deletion levels in Alzheimer brains.

33. Goto, Y.-I., Nonaka, I. & Horai, S. *A mutation in the tRNA Leu (UUR) gene associated with the MELAS subgroup of mitochondrial encephalomyopathies*.

34. Bebenek, A. DNA replication fidelity. *Postepy biochemii* (2008). doi:10.1146/annurev.biochem.69.1.497

35. Yeeles, J. T. P., Poli, J., Marians, K. J. & Pasero, P. Rescuing stalled or damaged replication forks. *Cold Spring Harb. Perspect. Biol.* **5**, a012815 (2013).

36. Tsegay, P. S., Lai, Y. & Liu, Y. Replication Stress and Consequential Instability of the Genome and Epigenome. *Molecules* **24**, (2019).

37. Dehé, P.-M. & Gaillard, P.-H. L. Control of structure-specific endonucleases to maintain genome stability. (2017). doi:10.1038/nrm.2016.177

38. Doudican, N. A., Song, B., Shadel, G. S. & Doetsch, P. W. Oxidative DNA damage causes mitochondrial genomic instability in Saccharomyces cerevisiae. *Mol. Cell. Biol.* **25**, 5196–204 (2005).

39. Tretyakova, N. Y., Iv, A. G., Ji, S. & Tretyakova, N. DNA-Protein Cross-links: Formation, Structural Identities, and Biological Outcomes HHS Public Access. *Acc Chem Res* **48**, 1631–1644 (2015).

40. Helleday, T., Lo, J., van Gent, D. C. & Engelward, B. P. DNA double-strand break repair: From mechanistic understanding to cancer treatment. *DNA Repair (Amst).* (2007).

doi:10.1016/j.dnarep.2007.02.006

41. Dianov, G. L. Base excision repair targets for cancer therapy. *Am. J. Cancer Res.* (2011).

42. Marullo, R. *et al.* Cisplatin induces a mitochondrial-ROS response that contributes to cytotoxicity depending on mitochondrial redox status and bioenergetic functions. *PLoS One* **8**, e81162 (2013).

43. Swanson, R. L., Morey, N. J., Doetsch, P. W. & Jinks-Robertson, S. *Overlapping Specificities of Base Excision Repair, Nucleotide Excision Repair, Recombination, and Translesion Synthesis Pathways for DNA Base Damage in Saccharomyces cerevisiae*. *MOLECULAR AND CELLULAR BIOLOGY* **19**, (1999).

44. Girard, P. M. & Boiteux, S. *Repair of oxidized DNA bases in the yeast Saccharomyces cerevisiae*. *Biochimie* **79**, (1997).

45. Robertson, A. B., Klungland, A., Rognes, T. & Leiros, I. Base excision repair: The long and short of it. *Cellular and Molecular Life Sciences* **66**, 981–993 (2009).

46. Wallace, S. S. Base excision repair: A critical player in many games. *DNA Repair (Amst).* **19**, 14–26 (2014).

47. Marenstein, D. R. *et al.* Substrate Specificity of Human Endonuclease III (hNTH1). *J. Biol. Chem.* **278**, 9005–9012 (2003).

48. Maclean, M. J. *et al.* Base excision repair activities required for yeast to attain a full chronological life span. *Aging Cell* **2**, 93–104 (2003).

49. Dianov, G. L. & Hü Bscher, U. SURVEY AND SUMMARY Mammalian Base Excision Repair: the Forgotten Archangel. doi:10.1093/nar/gkt076

50. Kuiper, R. P., Nielsen, M., De Voer, R. M. & Hoogerbrugge, N. *NTHL1 Tumor Syndrome*. *GeneReviews®* (University of Washington, Seattle, 1993).

51. Kuiper, R. P. & Hoogerbrugge, N. NTHL1 defines novel cancer syndrome. **6**,

52. Fishel, R. Mismatch Repair *. (2015). doi:10.1074/jbc.R115.660142

53. Fleck, O. & Nielsen, O. DNA repair. *J. Cell Sci.* **117**, 515–7 (2004).

54. Lombard, D. B. *et al.* DNA repair, genome stability, and aging. *Cell* (2005). doi:10.1016/j.cell.2005.01.028

55. de Souza-Pinto, N. C. *et al.* Novel DNA mismatch-repair activity involving YB-1 in human mitochondria. *DNA Repair (Amst).* **8**, 704–719 (2009).

56. Morris, L. P., Conley, A. B., Degtyareva, N., Jordan, I. K. & Doetsch, P. W. Genome-wide map of Apn1 binding sites under oxidative stress in *Saccharomyces cerevisiae*. *Yeast* **34**, 447–458 (2017).

57. Beljanski, V., Marzilli, L. G. & Doetsch, P. W. DNA Damage-Processing Pathways Involved in the Eukaryotic Cellular Response to Anticancer DNA Cross-Linking Drugs. *Mol. Pharmacol.* **65**, 1496–1506 (2004).

58. Jalal, D., Chalissery, J. & Hassan, A. H. SURVEY AND SUMMARY Genome maintenance in Saccharomyces cerevisiae: the role of SUMO and SUMO-targeted ubiquitin ligases. *Nucleic Acids Res.* **45**, 2242–2261 (2017).

59. Branzei, D. & Foiani, M. Regulation of DNA repair throughout the cell cycle. *Nature Reviews Molecular Cell Biology* **9**, 297–308 (2008).

60. Spivak, G. Nucleotide excision repair in humans. *DNA Repair* (2015). doi:10.1016/j.dnarep.2015.09.003

61. Li, X. & Heyer, W. D. Homologous recombination in DNA repair and DNA damage tolerance. *Cell Research* (2008). doi:10.1038/cr.2008.1

62. Jasin, M. & Rothstein, R. Repair of strand breaks by homologous recombination. *Cold*

Spring Harb. Perspect. Biol. **5**, (2013).

63. Chang, H. H. Y., Pannunzio, N. R., Adachi, N. & Lieber, M. R. Non-homologous DNA end joining and alternative pathways to double-strand break repair HHS Public Access. *Nat Rev Mol Cell Biol* **18**, 495–506 (2017).

64. Malkova, A. & Ira, G. Break-induced replication: functions and molecular mechanism. *Curr. Opin. Genet. Dev.* **23**, 271–9 (2013).

65. Mjelle, R. *et al.* Cell cycle regulation of human DNA repair and chromatin remodeling genes. *DNA Repair (Amst).* (2015). doi:10.1016/j.dnarep.2015.03.007

66. Beljanski, V., Marzilli, L. G. & Doetsch, P. W. DNA damage-processing pathways involved in the eukaryotic cellular response to anticancer DNA cross-linking drugs. *Mol. Pharmacol.* **65**, 1496–506 (2004).

67. Friedberg, E. C. SUFFERING IN SILENCE: THE TOLERANCE OF DNA DAMAGE. (2005). doi:10.1038/nrm1781

68. Ghosal, G. & Chen, J. DNA damage tolerance: A double-edged sword guarding the genome. *Translational Cancer Research* **2**, 107–129 (2013).

69. Alseth, I. *et al. The Saccharomyces cerevisiae Homologues of Endonuclease III from Escherichia coli, Ntg1 and Ntg2, Are Both Required for Efficient Repair of Spontaneous and Induced Oxidative DNA Damage in Yeast*. *MOLECULAR AND CELLULAR BIOLOGY* **19**, (1999).

70. Thayer, M. M., Ahern', H., Xing1, D., Cunningham1l2, R. P. & Tainer2, J. A. *Novel DNA binding motifs in the DNA repair enzyme endonuclease Ill crystal structure*. *The EMBO Journal* **14**, (1995).

71. Imai, K. *et al. Genomic structure and sequence of a human homologue (NTHL1/NTH1) of*

Escherichia coli endonuclease III with those of the adjacent parts of TSC2 and SLC9A3R2 genes. *Gene* **222**, (1998).

72. Aspinwall, R. *et al.* Cloning and characterization of a functional human homolog of Escherichia coli endonuclease III. *Proc. Natl. Acad. Sci. U. S. A.* **94**, 109–14 (1997).
73. Shida, T., Noda, M. & Sekiguchi, J. Cleavage of single- and double-stranded DNAs containing an abasic residue by Escherichia coli exonuclease III (AP endonuclease VI). *Nucleic Acids Res.* **24**, 4572–4576 (1996).
74. Kanchan, S., Mehrotra, R. & Chowdhury, S. In Silico Analysis of the Endonuclease III Protein Family Identifies Key Residues and Processes During Evolution. *J. Mol. Evol.* **81**,
75. Lukianova, O. A. & David, S. S. A role for iron-sulfur clusters in DNA repair. *Current Opinion in Chemical Biology* **9**, 145–151 (2005).
76. Wolfe, K. H. Origin of the Yeast Whole-Genome Duplication. (2015). doi:10.1371/journal.pbio.1002221
77. Kellis, M., Birren, B. W. & Lander, E. S. *Proof and evolutionary analysis of ancient genome duplication in the yeast Saccharomyces cerevisiae*. (2004).
78. Jin You, H., Swanson, R. L. & Doetsch, P. W. *Saccharomyces cereVisiae Possesses Two Functional Homologues of Escherichia coli Endonuclease III †*. (1998).
79. Swartzlander, D. B. *et al.* Regulation of base excision repair: Ntg1 nuclear and mitochondrial dynamic localization in response to genotoxic stress. doi:10.1093/nar/gkq108
80. Sentürker, S. *et al. Substrate specificities of the Ntg1 and Ntg2 proteins of Saccharomyces cerevisiae for oxidized DNA bases are not identical*. *Nucleic Acids Research* **26**, (1998).
81. Hilbert, T. P., Chaung, W., Boorstein, R. J., Cunningham, R. P. & Teebor, G. W. *Cloning*

*and Expression of the cDNA Encoding the Human Homologue of the DNA Repair Enzyme, Escherichia coli Endonuclease III**. (1997).

82. Yoshitani, A., Yoshida, M. & Ling, F. A novel cis-acting element required for DNA damage-inducible expression of yeast DIN7. *Biochem. Biophys. Res. Commun.* **365**, 183–188 (2008).

83. Luna, L., Bjøras, M., Hoff, E., Rognes, T. & Seeberg˚, E. S. *Cell-cycle regulation, intracellular sorting and induced overexpression of the human NTH1 DNA glycosylase involved in removal of formamidopyrimidine residues from DNA*. *Mutation Research* **460**, (2000).

84. Bouziane, M. *et al. Promoter structure and cell cycle dependent expression of the human methylpurine-DNA glycosylase gene*. *Mutation Research* **461**, (2000).

85. Liu, X., Choudhury, S. & Roy, R. In Vitro and in Vivo Dimerization of Human Endonuclease III Stimulates Its Activity. *J. Biol. Chem.* **278**, 50061–50069 (2003).

86. Marenstein, D. R. *et al.* Substrate specificity of human endonuclease III (hNTH1): Effect of human ape1 on hNTH1 activity. *J. Biol. Chem.* **278**, 9005–9012 (2003).

87. Eide, L. *et al. Base excision of oxidative purine and pyrimidine DNA damage in Saccharomyces cerevisiae by a DNA glycosylase with sequence similarity to endonuclease III from Escherichia coli (formamidopyrimidine/thymine glycol/NTGI/DNA damage inducibility)*. *Biochemistry* **93**, (1996).

88. Klungland, A. *et al.* Base Excision Repair of Oxidative DNA Damage Activated by XPG Protein. *Mol. Cell* **3**, 33–42 (1999).

89. Limpose, K. L., Corbett, A. H. & Doetsch, P. W. BERing the burden of damage: Pathway crosstalk and posttranslational modification of base excision repair proteins regulate DNA

damage management. *DNA Repair* **56**, 51–64 (2017).

90. Griffiths, L. M. *et al.* Dynamic Compartmentalization of Base Excision Repair Proteins in Response to Nuclear and Mitochondrial Oxidative Stress. *Mol. Cell. Biol.* **29**, 794–807 (2009).

91. A Weren, R. D. *et al.* A germline homozygous mutation in the base-excision repair gene NTHL1 causes adenomatous polyposis and colorectal cancer. *Nat. Publ. Gr.* (2015). doi:10.1038/ng.3287

92. Goto, M. *et al.* Altered expression of the human base excision repair gene NTH1 in gastric cancer. *Carcinogenesis* **30**, 1345–1352 (2009).

93. Koketsu, S., Watanabe, T. & Nagawa, H. Expression of DNA repair protein: MYH, NTH1, and MTH1 in colorectal cancer. *Hepatogastroenterology.* **51**, 638–42 (2004).

94. Limpose, K. L. *et al.* Overexpression of the base excision repair NTHL1 glycosylase causes genomic instability and early cellular hallmarks of cancer. *Nucleic Acids Res.* **46**, 4515–4532 (2018).

95. Saha, T., Rih, J. K., Roy, R., Ballal, R. & Rosen, E. M. Transcriptional Regulation of the Base Excision Repair Pathway by BRCA1. *J. Biol. Chem.* **285**, 19092–19105 (2010).

96. Doetsch, P. W., Morey, N. J., Swanson, R. L. & Jinks-Robertson, S. Yeast base excision repair: interconnections and networks. *Prog. Nucleic Acid Res. Mol. Biol.* **68**, 29–39 (2001).

97. Knudsen, N. Ø., Andersen, S. D., Lützen, A., Nielsen, F. C. & Rasmussen, L. J. Nuclear translocation contributes to regulation of DNA excision repair activities. *DNA Repair (Amst).* **8**, 682–689 (2009).

98. Hieter, P., Mann, C., Snyder, M. & Davis, R. W. *Mitotic Stability of Yeast Chromosomes:*

A Colony Color Assay That Measures Nondisjunction and Chromosome Loss. *Cell* **40**, (1985).

99. Au, W.-C., Crisp, M. J., Deluca, S. Z., Rando, O. J. & Basrai, M. A. Altered Dosage and Mislocalization of Histone H3 and Cse4p Lead to Chromosome Loss in Saccharomyces cerevisiae. *Genetics* **179**, 263–75 (2008).

100. Johnson, E. S. Protein Modification by SUMO. *Annu. Rev. Biochem.* **73**, 355–382 (2004).

101. McPherson-Davie, A. J., Jowhar, Z. M., Doetsch, P. W. & Corbett, A. H. *A Saccharomyces cerevisiae Model for Overexpression of Ntg1 a Base Excision DNA Repair Protein Reveals Novel Genetic Interactions*.

102. Swartzlander, D. B. *et al.* Identification of SUMO modification sites in the base excision repair protein, Ntg1. *DNA Repair (Amst).* **48**, 51–62 (2016).

103. Borgermann, N. *et al.* SUMOylation promotes protective responses to DNA-protein crosslinks. *EMBO J.* **38**, (2019).

104. Hardeland, U., Steinacher, R., Jiricny, J. & Schär, P. Modification of the human thymine-DNA glycosylase by ubiquitin-like proteins facilitates enzymatic turnover. *EMBO J.* (2002). doi:10.1093/emboj/21.6.1456

105. Smet-Nocca, C., Wieruszeski, J. M., Léger, H., Eilebrecht, S. & Benecke, A. SUMO-1 regulates the conformational dynamics of Thymine-DNA Glycosylase regulatory domain and competes with its DNA binding activity. *BMC Biochem.* (2011). doi:10.1186/1471-2091-12-4

106. Steinacher, R. & Schär, P. Functionality of human thymine DNA glycosylase requires SUMO-regulated changes in protein conformation. *Curr. Biol.* (2005). doi:10.1016/j.cub.2005.02.054

107. Ouspenski, I. I., Elledge, S. J. & Brinkley, B. R. New yeast genes important for chromosome integrity and segregation identified by dosage effects on genome stability. *Nucleic Acids Res.* (1999). doi:10.1093/nar/27.15.3001

108. Barnes, D. E., Lindahl, T. & Sedgwick, B. DNA repair. *Curr. Opin. Cell Biol.* **5**, 424–433 (1993).

109. Biallelic NTHL1 Mutations in a Woman with Multiple Primary Tumors. (2015). doi:10.1056/NEJMc1506878

110. Belhadj, S. *et al.* NTHL1 biallelic mutations seldom cause colorectal cancer, serrated polyposis or a multi-tumor phenotype, in absence of colorectal adenomas. **IDIBELL**, 8908

111. Weren, R. DA *et al. NTHL1* and *MUTYH* polyposis syndromes: two sides of the same coin? *J. Pathol.* **244**, 135–142 (2018).

112. Altaraihi, M., Gerdes, A.-M. & Wadt, K. A new family with a homozygous nonsense variant in NTHL1 further delineated the clinical phenotype of NTHL1-associated polyposis. doi:10.1038/s41439-019-0077-3

113. Ikeda, S., Kohmoto, T., Tabata, R. & Seki, Y. Differential intracellular localization of the human and mouse endonuclease III homologs and analysis of the sorting signals. *DNA Repair (Amst).* (2002). doi:10.1016/S1568-7864(02)00145-3

114. Davies, A. A., Friedberg, E. C., Tomkinson, A. E., Wood, R. D. & West, S. C. Role of the Rad1 and Rad10 proteins in nucleotide excision repair and recombination. *J. Biol. Chem.* **270**, 24638–41 (1995).

115. Swanson, R. L., Morey, N. J., Doetsch, P. W. & Jinks-Robertson, S. *Overlapping Specificities of Base Excision Repair, Nucleotide Excision Repair, Recombination, and*

Translesion Synthesis Pathways for DNA Base Damage in Saccharomyces cerevisiae. *MOLECULAR AND CELLULAR BIOLOGY* **19**, (1999).