Distribution Agreement

In presenting this 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 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 dissertation. I retain all ownership rights to the copyright of the dissertation. I also retain the right to use in future works (such as articles or books) all or part of this dissertation.

Signature:

_____ _____

Nicholas C. Bauer Date

Regulation of Base Excision and Strand Incision Repair by Base Damage–Induced Dynamic Compartmentalization

By

Nicholas Christopher Bauer
Doctor of Philosophy

Graduate Division of Biological and Biomedical Sciences
Biochemistry, Cell and Developmental Biology

_____ _____

Paul W. Doetsch, Ph.D.
Co-Advisor

Anita H. Corbett, Ph.D.
Co-Advisor

_____

Yoke Wah Kow, Ph.D.
Committee Member

_____

Adam I. Marcus, Ph.D.
Committee Member

_____

Eric A. Ortlund, Ph.D.
Committee Member

Accepted:

_____

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

_____

Date

Regulation of Base Excision and Strand Incision Repair by Base Damage–Induced Dynamic Compartmentalization

By

Nicholas Christopher Bauer
B.S., Bates College, 2008

Advisors:
Paul W. Doetsch, Ph.D. and 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 Graduate Division of Biological and Biomedical Sciences
Biochemistry, Cell and Developmental Biology

2014

ABSTRACT

Regulation of Base Excision and Strand Incision Repair by
Base Damage–Induced Dynamic Compartmentalization

By Nicholas Christopher Bauer

Genomes (DNA) provide instructions for the development and function of an organism. Most DNA in eukaryotes resides in the nucleus, though a small but important fraction is located in mitochondria. Genomic fidelity is ensured by DNA repair pathways, many of which are shared by both compartments. While the biochemical mechanisms of repair have been delineated, repair regulation is largely unknown. Recent observations indicated that a base excision and strand incision repair (BESIR) glycosylase from the model eukaryote *Saccharomyces cerevisiae*, Ntg1, undergoes a shift in localization to the compartment enduring oxidative DNA damage. This novel regulatory mechanism was termed dynamic compartmentalization.

The discovery that Ntg1 is regulated by dynamic compartmentalization led to four key questions: 1) Is this mode of regulation general to BESIR? 2) Which lesions are responsible for initiating signaling? 3) How is the signal transduced? 4) How do these signals modulate the protein's distribution? Effectively addressing these questions depended on the availability of methods to measure localization and to introduce compartment-specific base lesions. Solving these methodological challenges and fully investigating dynamic compartmentalization was the main objective of this dissertation research.

Chapter 1 thoroughly reviews the newly intersecting fields of DNA damage/repair and control of protein localization. Chapter 2 describes an analysis of *S. cerevisiae* and human BESIR protein sequences. Chapter 3 contains the published work describing the Quantitative Subcellular Compartmentalization Analysis (Q-SCAn) method developed to robustly, rapidly, and automatically quantify the nucleomitochondrial distribution of BESIR proteins in *S. cerevisiae*. Chapter 4 discusses work characterizing Ung1 and the utility of bisulfite as an in vivo cytosine deamination agent. Chapter 4 also details an experiment to approach the question of which lesions are responsible for generating base damage–dependent reactive oxygen species. Chapter 5 examines a direct comparison between Q-SCAn and the manual scoring technique with Ntg1 dynamic compartmentalization. Finally, Chapter 6 discusses the body of work described here, places it in context, and suggests future work. This dissertation research has developed the methods necessary to vigorously pursue the lines of inquiry surrounding dynamic compartmentalization and has provided important insights into base damage signaling and repair protein localization.

Regulation of Base Excision and Strand Incision Repair by Base Damage–Induced Dynamic Compartmentalization

By

Nicholas Christopher Bauer
B.S., Bates College, 2008

Advisors:
Paul W. Doetsch, Ph.D. and 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 Graduate Division of Biological and Biomedical Sciences
Biochemistry, Cell and Developmental Biology

2014

ACKNOWLEDGEMENTS

This dissertation represents the culmination of over five years of work, but it never would have come together without the advice and support of numerous other individuals. First and foremost I thank my advisors, Paul Doetsch and Anita Corbett. Their advice has helped me to develop both as an experimentalist and as a science writer, and I very much appreciated the degree of freedom I was allowed in pursuing development of the Q-SCAn image analysis algorithm. I also thank my committee members Yoke Wah Kow, Adam Marcus, and Eric Ortlund for helping to keep me on track and for their advice over the years, including their feedback in the preparation of this manuscript. I thank Natasha Degtyareva for her incredible knowledge of yeast biology and helpful experimental advice, former students Dan Swartzlander and Lyra Griffiths who initiated the project, my industrious undergraduate mentee Xi Jiang, and everyone in both the Doetsch and Corbett laboratories who were a source for advice, commiseration, and fun. I thank Stacy Holloway for her crucial support in navigating Emory's institutional systems and Paul's schedule, and Ms. Jannie for washing our glassware and for being a constant friendly presence. I also thank the leaders of the BCDB program (including Anita and Rick Kahn) and the GDBBS for making graduate school at Emory the best choice I could have made.

This dissertation also could not have happened without the support of my friends, both from college and those met in just the past few years, who helped to keep me sane through both game nights and nights out, when even the easiest experiment would seem to not work.

Lastly, but certainly not least, I must thank the National Institutes of Health for their funding of my research and training, through my advisors' grants, through the program's BCMB training grant, and through my individual F31 NRSA fellowship awarded by the National Cancer Institute.

CONTENTS

FIGURES AND TABLES

Figures

Tables

Chapter 1
INTRODUCTION

A heritable set of macromolecules composed of deoxyribonucleic acid (DNA), the genome, provides the instructions for development and function of an organism. Eukaryotic cells (cells with organelles, e.g. human cells) contain two discrete genomes, one located in the nucleus and the other in mitochondria. The nuclear genome encodes the vast majority of the cellular machinery. The mitochondrial genome encodes a limited number of factors crucial for mitochondrial functions including energy metabolism. Though it is much smaller than the nuclear genome, the mitochondrial genome is present in many copies and makes up a significant fraction of the total cellular DNA. A subset of eukaryotes possesses a third compartment, such as the light energy–harvesting plant chloroplast, containing a genome with properties similar to those described for the mitochondrial genome.

The nuclear and organellar genomes are under constant assault from endogenous and exogenous sources. The resulting chemical alterations can cause dysfunction. Thus, these genomes must be protected by efficient, redundant DNA repair pathways. These repair pathways are active in all genome-containing compartments, and many of the pathways are shared, with the same protein pool serving roles at each location. While the biochemical mechanisms of repair have been generally delineated, repair regulation is still relatively unknown. Recent observations indicated that the repair protein Ntg1 in a model eukaryote, budding yeast *Saccharomyces cerevisiae*, undergoes a shift in localization between nuclei and mitochondria dependent on compartment-specific genomic damage. The discovery that Ntg1 is regulated by dynamic compartmentalization led to four key questions: 1) Is dynamic compartmentalization general to base damage repair enzymes, or is it specific to Ntg1? 2) What DNA lesions are responsible for inducing dynamic compartmentalization? 3) How is information about the lesion transmitted to the protein? 4) How is this information

integrated by the protein to modulate its distribution? However, effectively addressing these questions depended on the availability of methods to measure localization and to introduce compartment-specific base lesions. Solving these methodological challenges and fully investigating dynamic compartmentalization was the main objective of this dissertation research.

This chapter provides a broad and thorough review of the newly intersecting fields of DNA damage/repair and control of protein localization. These reviews are followed by a final section discussing current research into the role of localization in regulating DNA repair activity and which is concluded by an overview of the work presented in this dissertation.

DNA Damage and Repair

Living organisms are highly ordered systems. Maintaining this order in the face of the inexorable pull of entropy requires an immense amount of energy; as a first approximation, a median human requires on the order of 230 gigajoules (2000 Calories/day for 74 years) (1). This energy is derived from the complete digestive breakdown of numerous other highly ordered organisms. Assuming that the net order of all life is approximately the same on a per-mass basis, the average human in the United States produces, in total, on the order of 500 times its mass in entropy from food consumption (2,3). Of course, even that is not enough to maintain the organism indefinitely; the system still eventually breaks down, resulting in disease, aging, and death. (Note that theories suggesting that aging and death are programmed have been proposed, but they are currently controversial (4).)

One of the critically important components maintained by this continual intake of energy is the genome, copies of which are contained in each of the approximately 100 trillion cells that compose the human body. The genome stores encoded instructions for making the ribonucleic acid (RNA) and protein molecules that shape the function of cells, tissues, and organs. Due to the central role of the genome in cellular processes, these DNA molecules must be immaculately maintained over the entire lifespan of a cell, which may be minutes for

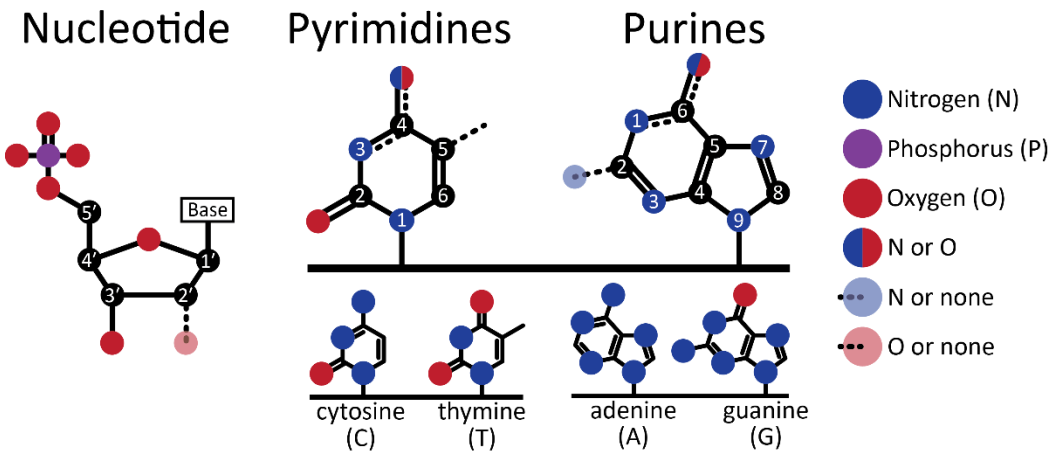

Figure 1-1. Basic nucleic acid chemistry. The basic chemical structure of (2′-deoxy)ribonucleotide monomers are diagramed. The standard ring numbering scheme is also shown. Position 1 of pyrimidines and position 9 of purines are connected to the 1′ position of the (deoxy)ribose. Hydrogens are omitted.

simple bacteria to decades for human neurons. The genome must also be faithfully copied when a cell divides. These goals are undermined by the constant assault from exogenous chemicals and radiation, as well as friendly fire from the metabolic and signaling pathways necessary to sustain life.

DNA is a polymer consisting of a backbone formed by molecules of the 5-carbon sugar 2′-deoxyribose linked together by a phosphate between the 5′ carbon of one sugar and the 3′ carbon of the next (Figure 1-1). Each sugar is linked via an *N*-glycosidic bond to a nitrogenous base at the 1′ carbon of the sugar (Figure 1-1). These nitrogenous bases are the purines adenine (A) and guanine (G), and the pyrimidines cytosine (C) and thymine (T) (Figure 1-1). DNA exists as a double helix of two DNA molecules attached by specific hydrogen bonds formed between paired bases (A:T, G:C) on each strand and stabilized by the hydrophobic effect and π-stacking interactions between adjacent bases.

The functions of DNA depend entirely on the chemical fidelity of the molecule. However, there are many processes that can adulterate the molecule and lead to dysfunction: reactive chemicals can modify the bases; a base can spontaneously hydrolyze from the sugar, leaving an abasic site; reactive chemicals and UV radiation can produce covalent crosslinks between bases on the same or opposing strands; and energetic ionizing radiation can induce

modifications as well as single- or double-strand breaks. Depending on the kind of damage, the biological effect of DNA lesions ranges from silent, to mutagenic, to lethal.

This constant threat of damage and the associated consequences caused life to very quickly evolve a set of defensive mechanisms. These DNA repair and tolerance pathways, many of which are shared by both nuclear and mitochondrial genomes, are as well conserved as any of the other core biological systems, including those responsible for replication of new DNA, transcription of DNA to RNA, and translation of RNA to protein. One of the trade-offs for evolving DNA repair was a reduced mutation rate, which would necessarily lead to a reduced rate of evolutionary innovation even as it stabilized the genome of a thriving organism. Early in the origin of life, relatively high rates of damage and mutagenesis may have allowed the rapid sampling of a wide array of polymer structures and functions, accelerating the chemical evolution of what is now the core biological machinery. Note that the rates of damage may have been lower compared to today, as oxygen, a key agent of DNA damage, was far less abundant at the time (5). As the core biological machinery matured and reached a minimal level of stable function, rapid mutagenesis would have been increasingly likely to act as a destructive force rather than as a creative force, because changes to an optimized macro-molecule or system of macromolecules would be more likely to degrade function than to improve it. Further, high mutation rates would make it less likely for a beneficial allele to propagate widely before being reverted or alternatively hidden by a deleterious change elsewhere. Subsequent evolution of DNA repair systems would have slowed and stabilized mutation rates, allowing single phenotypic changes that could undergo selection against a steady phenotypic background. It is intriguing, therefore, to observe evidence that many bacteria have had alternating evolutionary periods of loss and recovery (by horizontal gene transfer) of DNA repair genes (6). Many bacteria also have the ability to increase mutation rates during stress through the upregulation of low-fidelity polymerases (6). These

observations imply that a balance of repair efficiency and mutation rate is optimal for adaptive evolution.

An important feature of these repair pathways is their partial redundancy. Redundancy ensures that if any one pathway is temporarily overwhelmed, lesions can still be repaired or their effects mitigated. If a lesion-specific pathway is unavailable, a more general pathway can often take its place. If direct, precise repair mechanisms are overwhelmed, lower-fidelity backup pathways can be invoked that prioritize survival over accuracy. The result is a complex, layered defensive strategy to prevent or mitigate the effects of DNA damage. The types and causes of the damage that can be inflicted on DNA are discussed in the next section, followed by a review of the repair pathways brought to bear against these lesions. This topic is concluded by a brief overview of what is currently known about the signaling mechanisms invoked to coordinate the DNA damage response.

Damage types

DNA can undergo multiple types of damage, ranging from common, simple, mildly mutagenic base lesions to severe double-strand breaks that threaten the loss of entire chromosome arms. These lesions are reviewed below.

Base, sugar, and single-strand break lesions

DNA base lesions are the most common type of genomic damage. Bases can be oxidized, deaminated, or alkylated, and they can spontaneously undergo hydrolysis of the *N*-glycosidic bond, leaving behind an abasic (apyrimidinic/apurinic, or AP) site. An estimated 100,000 oxidative lesions, 10,000 deamination events, and 10,000 hydrolysis events occur in the 6.5 Gbp nuclear genome of human liver cells per day (7). These lesions are also potentially mutagenic as they can be misread by DNA polymerases to produce incorrect daughter strands (8) and by RNA polymerases to produce incorrect transcripts (9). Even when not being read by a polymerase, base lesions can also interfere with recognition of regulatory elements by

DNA binding proteins (8).

Reactive oxygen species (ROS; e.g. hydrogen peroxide, hydroxyl radical, superoxide anion) are critical for certain signaling pathways, yet they are also a significant source of base damage (10). As a result, the level of these species is very carefully regulated, deliberately produced by oxidases and removed by scavengers. ROS can also originate from the environment, both directly and as an aftereffect of the reactions of antioxidants and xenobiotics (11), and hydroxyl radicals can be produced by ultraviolet (UV) radiation in the UV-A band (315–400 nm) (12). Radiolysis of water by ionizing radiation can also produce ROS, in addition to reactive free protons and electrons which can produce similar sets of base lesions (13).

Oxidative attack on pyrimidines (C, 5-methylcytosine (5MeC), T) can result in the saturation of the double bond between pyrimidine carbons 5 and 6 to form hydrate (5-hydroxy-6-hydro, 6-hydroxy-5-hydro) and glycol derivatives. 5,6-dihydro derivatives are formed exclusively by the free protons and/or electrons generated by ionizing radiation. Carbon 5/6 hydrogens can be substituted to form 5-hydroxy and 5,6-hydroxy derivatives. C and 5MeC saturated lesions are especially prone to deamination, leading to the uracil (U) and T forms of the lesion, respectively, and uracil glycol rapidly decomposes to 5-hydroxyuracil (14-16). An exception to this instability is 6-hydroxy-5-hydrocytosine (14). Further oxidation of C (and U) can also lead to alloxan and then 5-hydroxyhydantoin, and T to 5-hydroxy-5-methylhydantoin. The 5-methyl group of 5MeC and T can both be oxidized to 5-hydroxy-methyl, 5-formyl, and 5-carboxy C and U derivatives, respectively. Finally, any pyrimidine can be further oxidized to urea. Pyrimidine radical reaction mechanisms and products are reviewed in (13,17) and are illustrated in Figure 1-2.

Oxidation of purines (A, G) can produce the ring-opened formamidopyrimidine derivatives, 8-oxo derivatives, and 2-hydroxyadenine. 8-oxoguanine can undergo extensive

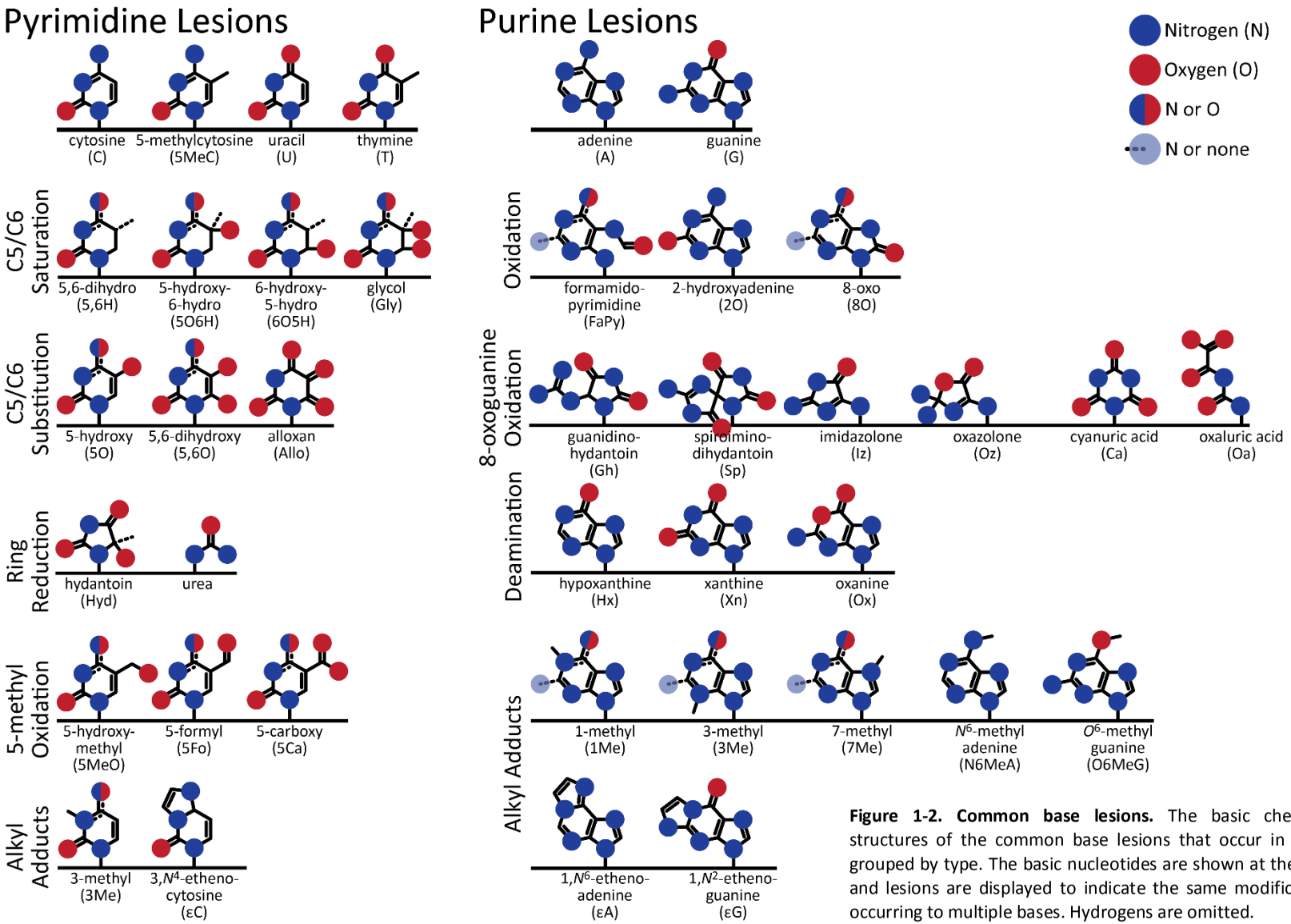

Figure 1-2. Common base lesions. The basic chemical structures of the common base lesions that occur in DNA, grouped by type. The basic nucleotides are shown at the top, and lesions are displayed to indicate the same modification occurring to multiple bases. Hydrogens are omitted.

further oxidation to the more mutagenic lesions guanidinohydantoin, spiroiminohydantoin, imidazolone, oxazolone, cyanuric acid, oxaluric acid, and urea. Purine oxidative reaction mechanisms and products are reviewed in (13,17,18) and are illustrated in Figure 1-2.

Deoxyribose in DNA is also vulnerable to oxidation, which usually results in base hydrolysis and/or a strand break in addition to the modified sugar. Deoxyribose oxidation can also result in a crosslink between the 5′ carbon of the sugar and carbon 8 of an attached purine, forming a cyclic nucleotide (13). There is an older report of oxidative conversion of deoxyribose to ribose in vivo (19), but it has never been followed up. An abasic site may also spontaneously form an *O*-glycosidic bond with an alcohol (20).

Deamination is the replacement of a nitrogen atom with an oxygen atom. Exocyclic amines are the primary target of deamination. C and 5MeC have an exocyclic amine on carbon 4. Deamination of these bases by a basic molecule produces U and T, respectively (21). Purine deamination is primarily mediated by the signaling radical nitric oxide. A has an exocyclic amine at carbon 6, and deamination produces hypoxanthine (inosine). G has an exocyclic amine at carbon 2, and deamination produces xanthine. Uniquely, the internal nitrogen 1 of G can also be replaced by oxygen, producing oxanine (22).

The last major type of base lesion results from alkylation. Methyl groups can be added to any available amine on both pyrimidines and purines, as well as to the carbon 6 oxygen or nitrogen of G and A, respectively (23). Lipid peroxidation products can also react with an exocyclic amine and an adjacent internal amine in C, A, and G to produce, among other lesions, exocyclic etheno adducts (24). Xenobiotic metabolites (resulting from combustion products such as benzo(a)pyrene and other aromatic polyphenols) can also be highly reactive, forming bulky adducts (25,26). The non-bulky adducts are illustrated in Figure 1-2.

Replication of new DNA is another important source of damage. The free nucleotide pool is vulnerable to damage through the same mechanisms described above for bases in

DNA, with U and 8-oxoguanine nucleotides as the major products. These lesions can be incorporated into DNA by DNA polymerase. DNA polymerase may also make coding errors, inserting the wrong nucleotide or slipping backwards or forwards one or more bases to produce a loop in either strand. Common misincorporation of ribonucleotides in DNA was also demonstrated recently, at a frequency of approximately 4 for every 10^4 nucleotides inserted in *S. cerevisiae* (27).

Crosslinks and double-strand breaks

UV light, a component of solar radiation, is a major source of intrastrand crosslinks—covalent bonds between adjacent bases. Neighboring pyrimidines absorb the UV photons and enter an excited state in which they readily react with each other (28) to form cyclobutane pyrimidine dimers (CPDs), 6–4 pyrimidine photoproducts (64PPs), and Dewar valence isomers (reviewed in (12,17)).

Additionally, ROS can produce crosslinks as a minor oxidation product. Intrastrand crosslinks are the result of a covalent bond forming between carbon 8 of a guanine and either the 5-methyl carbon of an adjacent T or 5MeC or the carbon 5 of an adjacent C. Oxidative interstrand crosslinks can also form between the exocyclic amine of A and the 5-methyl group of its paired T. In addition to crosslinks within DNA, ROS can produce crosslinks to amino acid residues, permanently attaching a protein to the DNA. These crosslinks form between pyrimidines and typically either tyrosine or lysine (reviewed in (13,17)).

A unique form of complex lesion occurs when an unusual linkage is formed with a phosphate end group. Topoisomerase 1 (Human: TOP1; *S. cerevisiae*: Top1) is an enzyme responsible for relaxing the helical torsion produced by the unwinding and rewinding of DNA being transcribed or replicated (29). The enzyme accomplishes this task by forming a transient strand break and a covalent bond between a conserved tyrosine and a 3′-phosphate. This strand break releases the helical strain, then TOP1 re-ligates the DNA (29). However, the

presence of a base lesion at the incision site can prevent the sealing of the relaxed DNA, trapping the enzyme on DNA. Similar lesions can occur with topoisomerase 2 (Human: TOP2A, TOP2B; *S. cerevisiae*: Top2), which decatenates intertwined DNA by forming a double-strand break with an overhang, and with topoisomerase 3 (Human: TOP3A, TOP3B; *S. cerevisiae*: Top3), which decatenates intertwined single-stranded DNA, though these linkages are to a 5′-phosphate (29). Topoisomerase-derived and similar lesions are reviewed in (30). Another unusual 5′-5′ linkage between a 5′-phosphate end group and adenosine monophosphate (AMP) is generated during an aborted DNA ligase reaction, such as may occur at a DNA base lesion (31).

Double-strand breaks (DSBs) are the most serious lesions that may affect the genome. A DSB can not only directly disrupt a functional gene, it can also result in the loss of hundreds or thousands of genes by disconnecting an entire chromosome arm from the centromere, the element which directs the segregation of chromosomes to each daughter cell during mitosis. DSBs most often result from an encounter with a single-strand break by the replication fork (32), but they can also be produced during the intense concentrated oxidative assault produced by ionizing radiation. DSBs also occur physiologically in meiotic recombination (33) and recombination of immune receptor genes (34).

Repair pathways

The lesions described above are encountered regularly by most cells. The interconnected set of pathways collectively known as DNA repair is responsible for either reverting the lesion back to its original form or otherwise limiting the potential impact of the lesion on cellular function. As previously noted, these pathways are very highly conserved throughout all domains of life. Underscoring the critical role these pathways play, there are multiple, partially redundant pathways available to repair each lesion (35,36). While the biochemical mechanisms of these pathways are relatively well characterized, comparatively

little is known about how they are regulated; what is known about their regulation is primarily derived from genetic interaction experiments with biological endpoints which include sensitivity to chemical DNA damaging agents and mutagenesis. These pathways are reviewed below, with a particular focus on the critical base excision and strand incision repair pathway.

Direct lesion reversal

Certain lesions can be directly removed from a base while leaving the helix intact. Many alkyl adduct lesions can be directly removed by the AlkB homolog family of dioxygenases (Human: ALKBH2, ALKBH3; *S. cerevisiae*: not present). ALKBH2/3 removes alkyl groups (37-39) and exocyclic adducts (40-42) at the nitrogen 1 position of purines and the nitrogen 3 position of pyrimidines. The most highly mutagenic alkylated base, O^6-methylguanine, is reversed by O^6-methylguanine methyltransferase (Human: MGMT; *S. cerevisiae*: Mgt1), a "suicide protein" which irreversibly transfers the errant methyl group onto itself and is subsequently degraded (43). Photolyases (Human: not present; *S. cerevisiae*: Phr1) are light-dependent enzymes that directly reverse UV-induced pyrimidine dimers (44). Photolyases were lost early in the largely nocturnal mammalian lineage (45).

Base excision and strand incision repair

The base excision repair (BER) pathway efficiently corrects most non-bulky DNA base lesions that are not addressable by direct reversal. Thus, BER is responsible for repairing the vast majority of lesions that occur in DNA, and the pathway is active in both nuclei and mitochondria. BER is initiated by the recognition and hydrolysis of the damaged base by a DNA *N*-glycosylase, leaving an abasic site (46). The glycosylases are reviewed in depth below and their specific substrates are summarized in Figure 1-3. The overall BER pathway is illustrated in Figure 1-4.

Abasic sites can be processed by one of two subpathways. The first subpathway, called long patch BER (LP-BER), is initiated by an AP endonuclease (Human: APEX1, APEX2;

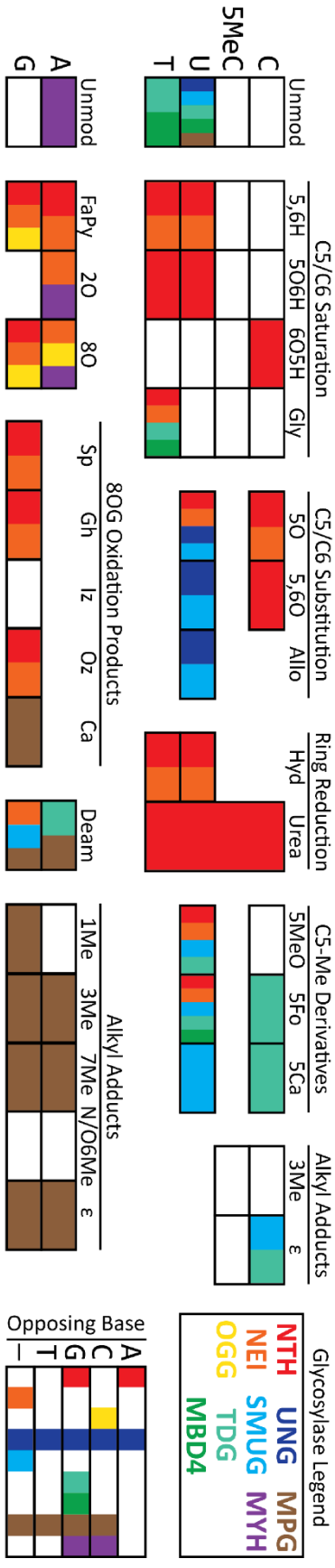

Figure 1-3. Substrate specificities of base excision repair glycosylases. Lesions have been organized into groups with shared features (e.g. same modification on different base), and colored boxes indicate the glycosylase family that recognizes and excises it from DNA, according to the legend at the top-right. The matrix at the bottom-right shows the opposite-base specificity of these glycosylases. This diagram does not account for differences in efficiency between families or within families (paralogs/orthologs). See Figure 1-2 for lesion structures and their assigned abbreviations.

S. cerevisiae: Apn1, Apn2), which cleaves the DNA backbone on the 5′ side of the abasic site (47). This incision leaves a 3′-hydroxyl group that can directly serve as a substrate for DNA polymerase, which fills the removed base and several bases downstream, displacing the strand on the 3′ side of the cut site (48). This displaced strand, which is terminated by the 5′-deoxyribose phosphate (dRP), is removed by the flap endonuclease (Human: FEN1; *S. cerevisiae*: Rad27) at its base and a DNA ligase seals the nick, leaving a fully repaired segment of DNA (46). An AP endonuclease–cleaved abasic site may also be directed into the second pathway. The mechanism controlling this switch is unclear, though it has been proposed that the human BER scaffold protein XRCC1 may play a role (49,50), especially for DNA polymerases possessing 5′-dRPase activity (51).

The other subpathway to process abasic sites, called single nucleotide BER (SN-BER; also termed “short patch”), is initiated by an AP lyase, which cleaves the DNA backbone on the 3′ side of the abasic site by β-elimination (AP β-lyase), which leaves a 3′-deoxyribose phosphate and 5′-phosphate, or by β,δ-elimination (AP β,δ-lyase), which

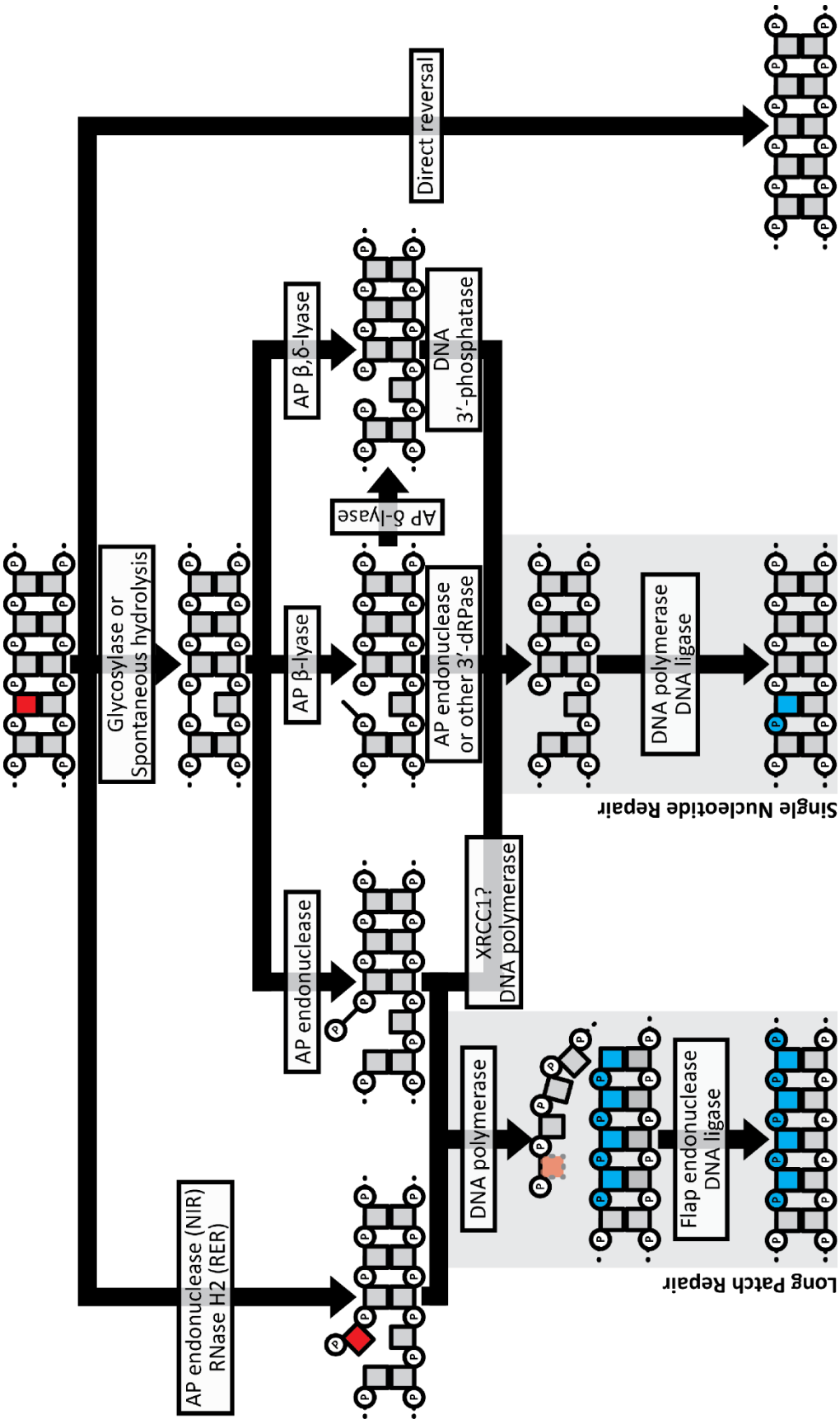

Figure 1-4. Base, sugar, and single-strand lesion repair pathways. Base lesions can be processed by direct reversal (right), base excision repair (middle) or strand incision repair (left). Multiple semi-redundant pathways are available for processing abasic sites. Pathway choice depends on the lesion, and some lesions can be acted on by multiple pathways. Squares (nucleosides) and connected circles (phosphates) represent a section of a DNA molecule, with the base lesion indicated in red. De novo synthesized bases are shown in blue. The enzyme activity responsible for each step is noted.

excises the deoxyribose and leaves 3′- and 5′-phosphates (47). Both of these products block DNA polymerase activity. Removal of the 3′-dRP can be accomplished by as little as a basic tripeptide, but is primarily the responsibility of an AP endonuclease (47). Removal of the 3′-phosphate is catalyzed by the DNA 5′-kinase/3′-phosphatase, polynucleotide kinase/phosphatase (Human: PNKP; *S. cerevisiae*: Tpp1) (52,53). The sequential action of AP δ-lyase and DNA 3′-phosphatase may also process 3′-dRP as an alternative to AP endonuclease (54,55). The result is a clean 1-nucleotide gap, which is then directly filled by a DNA polymerase and sealed by a DNA ligase (56).

The AP lyase activities which promote SN-BER are associated with the subset of glycosylases that recognize oxidative lesions, and this association tends to drive the choice of subpathways (47,57-59). This subpathway may be favored for oxidative lesions which occur in clusters (56). Attempting LP-BER in such a cluster could result in polymerase stalling, mutagenesis, or converting the original simple base lesion into a more serious double-strand break (60,61).

Classical BER has been joined by two strand incision repair pathways that feed into LP-BER: nucleotide incision repair (NIR) and ribonucleotide excision repair (RER). NIR is initiated by AP endonuclease, which can directly cleave the DNA backbone 5′ to pyrimidine lesions (62-68). Since AP endonucleases are so abundant, NIR can constitute the major repair activity against a subset of pyrimidine base lesions (65). RER is initiated by the RNase H2 complex (Human: RNASEH2A-RNASEH2B-RNASEH2C; *S. cerevisiae*: Rnh202-Rnh203), which incises the strand 5′ to the misincorporated ribonucleotide (69).

Other processes that result in lesions that look like BER intermediates can be handled by BER. Bases can spontaneously hydrolyze from the DNA backbone to generate abasic sites, and certain base lesions are more prone to hydrolysis than others. Oxidative attack on deoxyribose can also lead to base hydrolysis and a damaged sugar (13), which are effectively

processed by the coordinated actions of AP endonuclease, AP lyases, and 3′-DNA phosphatases (70,71). Complex single-strand breaks can occur through the action of abortive topoisomerase or DNA ligase activity (via AMP) adjacent to a lesion, which are trapped by covalent linkage to an end-group phosphate. Tyrosyl–DNA phosphodiesterase 1 (Human: TDP1; *S. cerevisiae*: Tdp1) directly catalyzes the release of a crosslinked TOP1/Top1 from the end-group phosphate. PNKP/Tpp1 then removes the 3′-phosphate and adds a 5′-phosphate to the break site. TOP2 lesions are removed by TDP2 (*S. cerevisiae*: not present), leaving a 5′-phosphate (30). Aprataxin deadenylase (Human: APTX; *S. cerevisiae*: Hnt3) directly removes the 5′-5′ AMP left behind by an aborted ligation (72). Resolution of these structures allows BER to complete repair of the instigating lesions.

As discussed above, since the initial discovery and characterization of the BER pathway, its roles have been expanded beyond base excision to include entry of non-base lesions (e.g. spontaneous abasic sites, deoxyribose lesions, single strand breaks, and protein–DNA crosslinks). Furthermore, the LP-BER subpathway may be initiated by direct strand incision without base excision, and this strand incision activity is orchestrated by one of the core BER enzymes, AP endonuclease, as well as by the highly conserved RNase H2. As a result, the terminology has become inaccurate. Perhaps a more appropriate term for this inter-related set of pathways would be “base excision and strand incision repair” (BESIR). This term has been adopted in the remainder of this dissertation, as it more closely reflects the range of activities of the pathway components. The enzymes critical for the initiation and early processing in BESIR, the glycosylases and AP endonucleases, are reviewed below. As RNase H2 and its involvement in RER is a relatively recent discovery and its activity is still under investigation, RNase H2 is not covered here.

Alkylpurine–DNA glycosylases

The alkylpurine–DNA glycosylases (Human: MPG; *S. cerevisiae*: Mag1) are responsible

for excising methylated purines (especially 3-methyladenine) (73-78), purine exocyclic adducts (e.g. 1,N^6-ethenoadenine) (79,80), deaminated purines (e.g. hypoxanthine, xanthine, oxanine) (81,82), the 8-oxoguanine oxidation product cyanuric acid (83), uracil (78), and *O*-glycosidic additions to deoxyribose (20). Many of these substrates can be excised from both double-stranded and single-stranded DNA (78,82). Mag1 can also remove A mispaired with C (79). Intriguingly, MPG can also catalyze the reverse reaction, forming an *N*-glycosidic bond with a free base, potentially allowing the correctly paired base to be directly swapped (84). The biological relevance of this activity would be dependent on the relative concentrations of free bases in the nucleus, and it could be counterproductive if an incorrect base were inserted which was not excisable by MPG.

These glycosylases do not contain lyase activity, and their resulting abasic sites can be processed both by SN and LP subpathways (59). Mag1 expression is inducible by the DNA damage checkpoint pathway (85-89), while MPG transcript level is cell-cycle regulated (90). MPG has been identified both in the nucleus and mitochondria (91), while Mag1 is restricted to nuclei (92). MPG activity is enhanced by XRCC1 and HR23 (93,94). MPG can bind to PCNA along with APEX1, and APEX1 can stimulate MPG turnover by displacing it from the abasic site products (95-97). MPG can also form a dimer with methylcytosine binding domain protein 1 (MBD1), which sequesters it at methylated CpG promoters. Upon alkylative attack on G, MBD1 dissociates from both MPG and the DNA, upon which MPG redistributes throughout the genome (77). This mechanism may provide an alkylation-responsive reservoir for rapid mobilization to repair alkylation damage.

Endonuclease III–like glycosylases

The endonuclease III–like *N*-glycosylase family (Human: NTHL1; *S. cerevisiae*: Ntg1, Ntg2) is responsible for repairing a wide array of oxidative lesions in double-stranded DNA, primarily oxidized pyrimidines (e.g. 5-hydroxycytosine, cytosine hydrates, thymine glycol)

(57,98-108), ring-fragmented purines (100,109-111), and 8-oxoguanine opposite a purine (112,113). There are several differences between these proteins with respect to substrate specificity. For example, Ntg1 and NTHL1 do not as efficiently process 5-hydroxycytosine, while Ntg2 processes it efficiently (65,98). Ntg1, however, is better at processing cytosine hydrates than Ntg2. Ntg2 and NTHL1 can excise certain 8-oxoguanine oxidation products, while Ntg1 does not (83,111,114). Ntg1 can process dihydrothymine, while Ntg2 cannot (106). This family of enzymes possesses AP β-lyase activity which directs lesions into the SN subpathway described above through a coordinated reaction mechanism, but this family can also contribute to the processing of abasic sites generated spontaneously or by monofunctional glycosylases (48,58,115,116).

NTHL1 activity on its own is over 100 times slower than its bacterial counterpart (117), and the rate-limiting step of its biochemical mechanism is release from the lyase-cleaved abasic site (118). APEX1 has been found to enhance release (104). This mechanism may protect the toxic strand-break intermediate in a "passing the baton" or "handoff" substrate channeling mechanism (119). Other binding partners (e.g. XPG, YB-1, XRCC1) enhance the activity of NTHL1, though the precise mechanism behind the enhancement is unknown (94,104,117,120).

Ntg1 and Ntg2 are the result of a genome duplication in the evolutionary history of *S. cerevisiae* (121,122), and they have segregated certain characteristics that the single original enzyme possessed. NTHL1 and Ntg1 localize to both the nucleus and mitochondria and are involved in repairing both genomes, while Ntg2 is strictly nuclear (98,106,123-128). On the other hand, NTHL1 and Ntg2 both contain a conserved iron-sulfur center, but this feature is missing from Ntg1 (57,100,129). This iron-sulfur center is redox-active in vivo and has been hypothesized to be involved DNA damage sensing. In brief, electrons can be transported over long distances through the π-orbitals of the DNA base pair stack between

bound redox-active proteins containing iron-sulfur centers, a process called DNA-mediated charge transport. Reduction of the iron-sulfur center allows dissociation of the repair enzyme from the DNA. DNA lesions which disrupt the base stack retard charge transport, causing the repair protein to remain bound to the DNA and slide along the helix to find the lesion (130-133). Little is known about the transcriptional regulation of this family of proteins, but NTHL1 is upregulated in S phase (110), and the Ntg1 promoter has a conserved (within yeast) promoter element necessary for oxidative stress induction (134).

Endonuclease VIII–like glycosylases

The endonuclease VIII–like *N*-glycosylase family (Human: NEIL1, NEIL2, NEIL3; *S. cerevisiae*: not present) is responsible for repairing a wide array of oxidative lesions primarily in single-stranded DNA. Most of the substrates of this enzyme family overlap with those of NTHL1 (65,107,109,111,135-148). NEIL1 and NEIL3 can also process 8-oxoguanine oxidation products in telomere-associated quadruplex DNA (149), NEIL3 can process thymine glycol in the same structures, and they are additionally able to process lesions near strand breaks which are refractory to NTHL1 and OGG1 (150,151). NEIL1 can excise the internally deaminated G, oxanine (152). NEILs are also involved in repair of some interstrand crosslinks (153).

This family of enzymes possesses AP lyase activity. In contrast to NTHL1, NEIL1 and NEIL2 are AP β,δ-lyases (138,139), while NEIL3 employs the typical β-elimination mechanism (111). The β,δ mechanism directs NEIL1 or NEIL2-excised lesions into the SN subpathway, but there is also evidence that these lesions can undergo LP repair (154). NEIL1 or NEIL2, with PNK, can provide backup dRPase activity to clean up after other AP β-lyases (55).

NEILs provide specialized repair activities in the cell. NEIL1 participates in pre-replication repair, in which it recognizes lesions in the single-stranded region just before being read by the DNA polymerase. NEIL1 excises the base and creates a strand break, forcing

the polymerase to stall and backtrack so that the lesion can be repaired (155,156). NEIL2 associates with RNA polymerase II and CSB, and can repair lesions encountered as transcription is progressing (135,149,157). There is some evidence that NEIL1 and NEIL2 can compensate for one another, but this ability appears to be limited (155,156,158,159).

NEIL1 is strongly upregulated during S phase (138) and under oxidative stress (160). In contrast, expression levels of NEIL2 are constant throughout the cell cycle (139) while still responsive to oxidative stress (161), perhaps reflecting their differing roles in replication and transcription. Both NEIL1 and NEIL2 have been found in mitochondria as well as in the nucleus (162,163). NEIL2 has been associated with microtubules, but the relevance of this interaction is unknown (164). As with NTHL1, NEIL2 activity can be stimulated by the scaffold XRCC1 (94). Intriguingly, NEIL1 is also subject to RNA editing by the adenosine deaminase ADAR1, which results in a K→R change in the lesion recognition site. The edited form is less efficient at removing thymine glycol, but more efficient at removing 8-oxoguanine oxidation products (165). However, the role of this editing is not yet clear.

8-oxoguanine–DNA glycosylases

The 8-oxoguanine–DNA glycosylase family (Human: OGG1; *S. cerevisiae*: Ogg1) is responsible for excising G oxidation products with intact ring systems, including the extremely common 8-oxoguanine and further modifications, specifically across from C (166-179), and G-derived formamidopyrimidine (167,172,177,179). This family of enzymes possesses an AP β-lyase activity specifically across from a C (59,166). Some weak δ-elimination has also been detected (171) and Ogg1 has a minor dRPase activity, perhaps due to δ-elimination coupled with the Tpp1 3′-phosphatase (54). OGG1 and Ogg1 lyase activity is fairly inefficient compared to their glycosylase activities (173), but OGG1 can be stimulated 5-fold in the presence of APEX1 (174,180,181). OGG1 lyase activity can also be substituted by NEIL1 /PNK, which binds abasic sites more strongly (182).

OGG1 is expressed as multiple isoforms resulting from alternative splicing. All isoforms contain a mitochondrial matrix targeting signal (MTS), but only isoform 1a contains a strong nuclear localization signal (NLS) (183). Isoform 1a is primarily localized to nuclei and to the nuclear matrix, but has also been detected in small amounts in mitochondria (184,185). All the other isoforms, which differ in the C-terminus, are primarily localized to mitochondria, with isoform 2a forming the majority of the mitochondrial pool, associating with the inner membrane (184,186). As with NEIL2, OGG1 associates with microtubules, but the function of this interaction has not been elucidated (187).

OGG1 expression can be upregulated by MMS treatment and antioxidants, but not by ROS (181,188,189). OGG1 activity may be stimulated by ribosomal protein S3, which may act to bring OGG1 and APEX1 to lesions; however, it can also bind 8-oxoguanine lesions strongly enough to prevent excision (190,191), which might suggest a role in the nucleolus. OGG1 can also be post-translationally modified. PKC can phosphorylate OGG1, though the function of this modification is unknown (185), while ROS-inducible p300-mediated acetylation weakens abasic site binding, thus enhancing APEX1-induced OGG1 activity (192).

Uracil–DNA glycosylases

The uracil–DNA glycosylase superfamily is one of the most highly conserved and diverse families of BESIR enzymes (193). While *S. cerevisiae* only has one (Ung1), mammals have four (UNG, SMUG1, TDG, MBD4). Mammalian glyceraldehyde-3-phosphate dehydrogenase (GAPDH) and cyclin O (CCNO) have also been reported to have uracil excision activity (194-196). These activities have not been characterized beyond these initial reports. Note that there was some initial confusion in the literature: CCNO was originally named UDG2, so some reports about the nuclear UNG isoform, UNG2, were errantly attributed to CCNO.

Uracil–DNA glycosylase (Human: UNG; *S. cerevisiae*: Ung1) is the major enzyme responsible for removing U (resulting from C deamination or misincorporation) (197-205)

and oxidized derivatives (e.g. 5-hydroxyuracil, alloxan) (206,207), acting on both double-stranded and single-stranded DNA (197,198). UNG localizes to both the nucleus and mitochondria as separate isoforms from alternate promoters (UNG1: mitochondrial; UNG2: nuclear) (208-212), while Ung1 is a single isoform that localizes to both compartments (213). UNG1 is the only known human mitochondrial uracil–DNA glycosylase. UNG2 and Ung1 are upregulated in S phase (208,214-217). UNG2 is controlled by both cyclin-dependent kinases (218,219) and TP53-dependent phosphatase (220,221), and associates with replication complexes (199,222-224). UNG1 is induced with oxidative stress (225). Both human forms are regulated by several miRNAs (226). With a relatively high K_M compared to the other uracil–DNA glycosylases (227), UNG may be specialized for rapidly detecting and excising lesions encountered by the replication fork and to quickly correct U misincorporated by DNA polymerase.

SMUG1 is nuclear and recognizes the same substrates as UNG1 and UNG2 (206,228-231), in addition to U and T oxidation products (e.g. 5-hydroxymethyluracil, 5-formyluracil, alloxan) (206,229-232), with some weak activity for exocyclic adducts of C (e.g. 3,N^4-ethenocytosine) (233) and deaminated purines (e.g. oxanine, xanthine) (152,234). SMUG1 overall is less efficient than UNG, but prefers single-stranded DNA 100-fold more than double-stranded DNA (228). SMUG1 also has a lower K_M than UNG, which gives it an advantage at lower substrate concentrations (227). As a result, SMUG1 is more efficient at repairing rare lesions in nonreplicating chromatin (235). Additionally, SMUG1 has novel activity removing 5-hydroxymethyluracil from rRNA in the nucleolus (236), presumably to perform quality control on rRNAs while they are unfolded and most vulnerable to damage.

TDG is a nuclear mismatch-specific glycosylase which excises U and T (237-239), highly oxidized 5MeC derivatives (e.g. 5-formylcytosine, 5-carboxylcytosine) (240-242), oxidized T derivatives (e.g. 5-hydroxymethyluracil, 5-formyluracil, thymine glycol) (240,243,

244), deaminated A (hypoxanthine) (240), and exocyclic C derivatives (245-248), all when across from G. TDG has a strong preference for a 5′-adjacent G, as is found in CpG islands (249), and has been strongly associated with both DNA methyltransferases (250,251) and transcription activation (252-256). Recently, a set of 5MeC oxidases, the TET dioxygenases, have been discovered which specifically oxidize the 5-methyl group, converting the methylated base into a substrate for TDG (257,258). Thus, TDG has been implicated in active DNA demethylation as well as reducing the mutation rate of these critical genetic control regions (259). TDG is expressed inversely with UNG, degraded on entering S phase and upregulated in G2 (260,261), and is translationally regulated by the miRNA miR-29 (262,263). TDG activity is controlled by p300/CBP-SIRT1 acetylation sites which reduce glycosylase activity, adjacent PKCα phosphorylation sites that block acetylation (261,264,265), and transient sumoylation which stimulates abasic site release (266,267).

MBD4 is a recently discovered nuclear mismatch-specific uracil–DNA glycosylase which is a fusion of a 5MeC binding domain and a uracil–DNA glycosylase domain (268). It associates with methylated CpG islands (269) and excises U, T and oxidized T derivatives (e.g. 5-formyluracil, thymine glycol) (243,244,270,271) when opposite G. Many other functions have been associated with MBD4 (272). The glycosylase activity of MBD4 has not been fully characterized, but it may be involved in active demethylation similar to TDG (251,273).

MutY family glycosylases

The MutY family is a G mismatch–specific adenine–DNA glycosylase (Human: MUTYH; *S. cerevisiae*: not present). A:G mispairs are the result of a misinsertion of A across from 8-oxoguanine. MUTYH can excise both A and oxidized A (e.g. 2-hydroxyadenine) across from either G or 8-oxoguanine (274-276). MUTYH is present in both nuclei and mitochondria (184,277) and is upregulated in S phase (278). In the nucleus, MUTYH is associated with replication complexes (278,279) so that mispairs can be corrected shortly after replication.

There are multiple known isoforms of MUTYH, but their distinct roles are not known (275,277). MUTYH activity is enhanced by the mismatch repair complex MSH2/6 (280) and APEX1 (281), but it also inhibits OGG1 excision across from, and APEX1 excision of, its resulting abasic site (282). As with NTHL1 and Ntg2, MUTYH also contains an iron-sulfur center, and may be regulated by a similar charge transport mechanism (132).

AP endonucleases

AP endonucleases are responsible for cleaving the DNA backbone 5′ to an abasic site or 3′-dRP left behind by AP β-lyase (283-291). AP endonucleases are also capable of recognizing oxidized abasic sites (70,71,292). APEX1 and Apn1 have an additional functionality in directly incising 5′ to oxidized and alkylated pyrimidines (62-68), and they can remove a 3′-terminal lesion with weak exonuclease activity (291,293-296). APEX1 can also cleave at abasic sites within RNA, playing a role in RNA quality control (297-299). But APEX1 has many other known functions as a result of its redox-sensitive transcription factor domain (reviewed in (300)). Interestingly, while APEX1, APEX2, and Apn2 belong to the exonuclease III family (301), Apn1 belongs to the endonuclease IV family (302).

APEX1 and APEX2 are both upregulated in S phase (303,304) and during DNA damage (305). APEX1 activity is reduced by SIRT1-removed acetylation (306) and CK2-added phosphorylation (307). APEX1 activity is also stimulated by HSP70 binding (308,309), the RAD9-RAD1-HUS1 checkpoint clamp (310), and XRCC1 (311).

APEX1, APEX2, and Apn1 all localize to both the nucleus and mitochondria (286,312-317) while Apn2 is exclusively nuclear (287,288), although APEX2 is distributed more to mitochondria (225). APEX1 and Apn1 make up the majority of AP endonuclease activity in both nuclei and mitochondria, with the weaker activities of APEX2 and Apn2 playing minor, backup, or specialized roles (225,286,289,318).

Endonuclease V–mediated excision repair

Recently, a unique repair pathway initiated by endonuclease V (Human: ENDOV; *S. cerevisiae*: not present) was discovered (319). ENDOV recognizes exocyclic-deaminated purines and cuts the DNA backbone 3′ to the nucleotide immediately 3′ to the lesion, preferentially in single-stranded regions. Through an unknown combination of nucleases, a short single-strand gap is created in the region of the nick, which is filled by DNA polymerase and sealed by DNA ligase. The discovery and current knowledge of this enzyme family is reviewed in (320).

Nucleotide excision repair

The nucleotide excision repair (NER) pathway corrects bulky chemical adducts and intrastrand crosslinks such as those produced by UV irradiation which induce helical distortion. NER can partially compensate for BESIR loss (321), and proteins involved in NER are involved in repair of certain BESIR substrates (322-326). NER is active in the nucleus, and there is no evidence of mitochondrial NER activity. Since there are dozens of copies of the small mitochondrial genome, it may be more efficient to degrade severely damaged mitochondrial DNA molecules and resynthesize new ones than to repair them (reviewed in (327)).

The general pathway, global genome repair (GG-NER), in eukaryotes involves recognition of helix distortion caused by the lesion either by xeroderma pigmentosa complementation group C (XPC) (*S. cerevisiae*: Rad4) and RAD23B (*S. cerevisiae*: Rad23), which destabilizes the duplex in the damaged region, or by XPA (*S. cerevisiae*: Rad14) and the single-stranded DNA binding complex RPA (Human: RPA1-RPA2-RPA3; *S. cerevisiae*: Rfa1-Rfa2-Rfa3), which locally unwinds the DNA. Each of these events introduces greater helical distortion and recruits the other complex (328). Once this initiating complex is assembled, the RNA polymerase II transcription factor H (TFIIH) helicase complex, which includes the proteins XPB (*S. cerevisiae*: Ssl2) and XPD (*S. cerevisiae*: Rad3), binds and further unwinds the region

around the damaged DNA (329). The open helix complex recruits XPG (*S. cerevisiae*: Rad2) and XPF-ERCC1 (*S. cerevisiae*: Rad1) to excise the damaged DNA strand on the 3′ and 5′ sides, respectively, leaving a 24–32 nucleotide gap (330). Finally, the gap is filled by a DNA polymerase and sealed with a ligase (331,332).

A second pathway, transcription-coupled repair (TC-NER), is initiated when RNA polymerase stalls at a blocking lesion. Cockayne syndrome group A and B (CSA and CSB; *S. cerevisiae*: Rad28 and Rad26) proteins recognize the stalled polymerase and recruit XPA, RPA, and the TFIIH to the lesion, which completes repair as in the general pathway (333).

Mismatch repair

The mismatch repair (MMR) pathway is responsible for correcting mispaired bases and unpaired loops, rather than modified bases, by recognizing the resulting helical distortion. MSH2 (*S. cerevisiae*: Msh2) paired with either MSH3 (large loops; *S. cerevisiae*: Msh3) or MSH6 (substitutions and small loops; *S. cerevisiae*: Msh6) recognizes the mispaired regions (334,335). These proteins recruit MLH to make an incision into the new strand 5′ to the lesion (336), with direction provided by the replication clamp PCNA (*S. cerevisiae*: Pol30) (337). Alternatively, MMR can make use of pre-existing nicks such as the end of an Okazaki fragment, or nicks resulting from excision of misincorporated uracil or ribonucleotides by UNG/APEX or RNase H2, respectively (27,116,338). A 5′–3′ exonuclease (Human: EXO1; *S. cerevisiae*: Exo1) removes the DNA 3′ to the nick and through the mismatch, leaving a gap (339) which is then filled by DNA polymerase and sealed by DNA ligase. Some alternate mechanisms have been proposed, such as polymerase extending from the nick to create a flap which is later excised, firmly connecting MMR to LP-BESIR, or that the polymerase can be induced to backtrack with 3′–5′ exonuclease activity and then resynthesize the mismatched DNA (340).

Nucleotide pool cleansing

As described above, free nucleotides can become damaged and may be subsequently

misincorporated into DNA. Specialized enzymes recognize these modified nucleotides and cleave off the terminal diphosphate, rendering them incapable of being incorporated into nascent DNA. These enzymes prominently include oxopurine nucleotide triphosphatase (Human: NUDT1; *S. cerevisiae*: Pcd1) and deoxyuridine triphosphatase (Human: DUT; *S. cerevisiae*: Dut1), which remove oxidized purines and U from the nucleotide pool, respectively. These nucleotide pool cleansing/sanitizing enzymes are reviewed in (341).

Lesion bypass by post-replication repair

While not technically repair, lesion bypass pathways allow a cell to mitigate some of the more deleterious effects of small unrepaired lesions and thus to have a greater chance of producing progeny with a correct base. When a processive polymerase reaches a stall-inducing lesion, it can be exchanged with a more permissive polymerase that can read through the lesion at the cost of a higher error rate when reading undamaged DNA. Some of these polymerases do better at inserting the correct base across from a lesion, such as is the case with DNA polymerase η (Human: POLH; *S. cerevisiae*: Rad30) which can properly insert two As across from a T dimer (342,343). Others simply insert a random base, especially when the lesion is an abasic site, just to allow replication to continue, such as with DNA polymerase ζ (Human: REV3L; *S. cerevisiae*: Rev3) (344,345).

Lesion bypass can also occur through recombination-related mechanisms to produce an error-free product. A stalled polymerase may cause regression of the replication fork, allowing synthesis using the other strand's daughter strand as a template, which can then reanneal across the lesion and allow polymerization to continue. Fork regression may also provide a chance for the DNA repair machinery to recognize and repair the lesion before replication continues. Alternatively, a lesion may result in a gap in the lagging strand, which can be repaired by the gapped lagging daughter strand invading the leading daughter strand and using it as a template to fill the gap. These pathways have been reviewed in (340).

Homologous recombination and non-homologous end joining

Homologous recombination (HR) is one of two major pathways to repair lethal double-strand breaks in eukaryotic cells. HR is most active in S phase and G_2, when an identical sequence is immediately available nearby: the recently replicated sister chromatid. Repair is initiated when MRN (MRE11-RAD50-NBS1; *S. cerevisiae*: Mre11-Rad50-Xrs2) complexes binds to each end of the break and to each other. Ataxia telangiectasia mutated (ATM) kinase (*S. cerevisiae*: Tel1) binds to the complex and phosphorylates the alternate histone H2AX and other proteins adjacent to the break, which results in the recruitment of 5′–3′ exonucleases (including EXO1/Exo1) to resect the DNA away from the breaks. The single-stranded region becomes coated with RPA, which in turn recruits and activates the ataxia telangiectasia and Rad3 related (ATR) kinase (*S. cerevisiae*: Mec1). BRCA2 (*S. cerevisiae*: not present) and RAD52 (*S. cerevisiae*: Rad52) replace RPA with RAD51 (*S. cerevisiae*: Rad51). RAD51 mediates the invasion of the resected strands onto the homologous duplex DNA. RAD54 (*S. cerevisiae*: Rad54) unwinds the homologous sequence and allows the "branch point," where the invading strand enters, to migrate along the DNA. Finally, DNA polymerase extends the invaded strands, the ends are sealed by ligases, and the Holiday junctions are cleaved by a resolvase (Human: GEN1; *S. cerevisiae*: Yen1).

The second pathway is called non-homologous end-joining (NHEJ), which predominates in G_1 and in postmitotic cells, or when the strand termini are damaged. The Ku70/80 heterodimers (*S. cerevisiae*: Yku70-Yku80) bind to each end of the break and keep them together. The DNA-dependent protein kinase (Human: PRKDC; *S. cerevisiae*: not present) binds to the Ku-sealed DNA ends and recruits factors to process the ends and re-ligate them together. This pathway can result in the loss of several base pairs at the break site, especially if the termini were damaged, but re-sealing the chromosome is preferable to not repairing it at all and losing an entire chromosome arm.

Several NHEJ backup pathways have recently been discovered and are currently being elucidated. If NHEJ fails due to extensive damage to the break ends or if the repair pathway is overwhelmed, the HR machinery can be recruited to perform minor resection. The resected regions can anneal with each other at sites of microhomology, and any remaining gaps are filled by DNA polymerase and sealed by DNA ligase, resulting in a short deletion of sequence at the break site. If NHEJ is completely overwhelmed or absent, MRN can bind to the damaged termini and recruit and activate poly(ADP-ribose) polymerase (Human: PARP1, PARP2; *S. cerevisiae*: not present). PARP creates large, branching chains of ADP-ribose polymer at the lesion (later removed by PAR glycohydrolase (PARG)), which recruits resectioning 5′–3′ exonucleases. The resected regions can then anneal with each other at an extensive exposed region of homology, following which the nonhomologous regions are cut back by a 3′ flap endonuclease, and the remaining single-stranded regions are filled by DNA polymerase and ligated together. This MRN/PARP process would result in a large deletion of sequence at the strand break. These pathways have been thoroughly reviewed in (346).

Interstrand crosslink repair

Interstrand crosslinks are highly lethal lesions which block polymerase activity. Their unique nature requires a complex process involving the cooperation of multiple repair pathways. Crosslinks encountered by the replication fork activate the Fanconi anemia pathway via the ATR/Mec1 kinase. The FANC complex recruits a series of nucleases to cleave one strand on either side of the crosslink, a process called unhooking. Lesion bypass can then synthesize the opposite strand, and a BER glycosylase (e.g. NEIL1) or the NER pathway can excise the bulky unhooked lesion (153,347). Homologous recombination can complete repair of the remaining double strand break. This pathway has been reviewed in (348). Crosslinks encountered by RNA polymerase invoke TC-NER, but proceeds similarly from there, sans the need to repair a double-strand break (349,350).

DNA damage signaling

Several cellular signaling pathways have evolved to allow global responses to DNA damage. These signaling pathways are broadly responsible for: 1) marshalling DNA repair proteins, both in general and to the lesion itself; 2) halting the cell cycle to prevent replication until the lesions are repaired; and 3) initiating senescence or apoptosis if the damage is too severe. There are two well-known signaling pathways that initiate the DNA damage response: the ATM/Tel1 and ATR/Mec1 kinases. As described above, ATM is activated in the presence of a DSB and coordinates homologous recombination. Meanwhile, ATR is activated in the presence of extensive RPA-bound single-stranded regions, which can occur as part of homologous recombination or otherwise. In addition to recruiting and coordinating repair factors, these kinases also activate the TP53 transcription factor, which coordinates the DNA damage checkpoint to halt replication until the damage is repaired (351). In the event of severe and sustained damage, TP53 activation initiates apoptosis (351).

While signaling in response to severe lesions has been known for some time, signaling resulting from base lesions has not been as clear. Recently, ATR was found to be activated by UV and bulky adducts (352). In the case of these NER-repaired lesions, the gap produced by the NER endonucleases can be extended by exonuclease 1. This extended gap recruits RPA and activates ATR/Mec1 signaling (353-356). These findings support a model where slowed, inefficient NER (as could occur in cells heavily damaged by bulky adducts or UV crosslinks and exceeding the NER repair capacity) can lead to increased ATR signaling and thus DNA damage checkpoint activation. Exonuclease 1 is actively involved in MMR (339) and it can also act at RNase H2 incision sites (69), indicating that EXO1 could potentially provide connections between any strand breaks, including those of BESIR intermediates, and ATR signaling. No studies have demonstrated such a link at this time. However, since NER is able to repair certain BESIR lesions (322-326), this pathway could also provide a connection

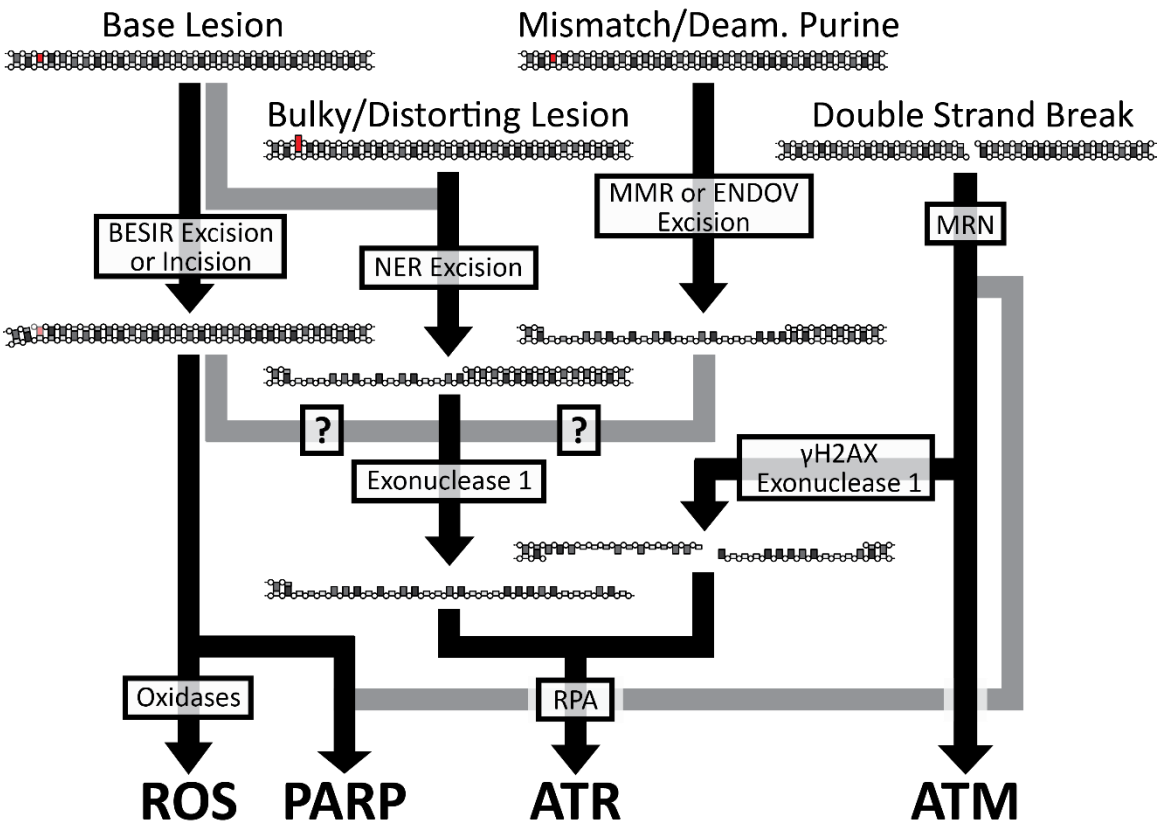

Figure 1-5. DNA damage signaling pathways. This diagram illustrates the connections between DNA lesions and the ROS, PARP, ATR, and ATM signaling pathways. Gray arrows indicate backup signaling pathways that may activate only when lesion levels have overwhelmed primary repair capacity. Proteins responsible for signal transduction is indicated. Question marks indicate a potential pathway, but which has not yet been observed.

between BESIR-repaired lesions and ATR signaling.

In mammalian cells, PARP is recruited to single-strand breaks, including those formed during BESIR. Not much is known about the role of this pathway, though it can act as a DNA damage signal and can, in the extreme, induce cell death. It has also been postulated to protect single-strand breaks under high genotoxic load until repair enzymes can be recruited (reviewed in (357)).

DNA damage, specifically base damage, also leads to the production of ROS as signaling molecules, which can trigger movement of oxidative response transcription factors into the nucleus (358). However, how these lesions lead to ROS generation or what other responses they may trigger are unknown. The relationships between various DNA lesions and DNA damage signaling pathways are illustrated in Figure 1-5.

Protein Localization

Location is a critical component of a protein's function. If a protein is not where it needs to be, it does not matter how active the protein is or how much of it there is. Conversely, if a protein is active in an unusual location, unexpected and undesired consequences could result. Protein function is controlled at numerous levels: rate of transcription, alternative transcription start and termination/polyadenylation sites, alternative splicing, transcript editing, transcript localization, transcript lifetime, rate of translation, alternative translation start sites, protein modifications affecting activity, protein localization, and protein lifetime. Out of all of these levels of regulation, only localization allows a pre-existing pool of proteins to be rapidly mobilized in response to changing cellular conditions without requiring the time or resources required for de novo synthesis.

Eukaryotic subcellular compartments are functionally defined by the proteins present within them. While many proteins are localized to a single compartment, some proteins are found in multiple compartments, either constitutively or in response to cellular signals. For example, many critical transcription factors are localized to the cytosol until a signal triggers nuclear import (e.g. β-catenin). Other proteins simultaneously play roles in multiple compartments, including some DNA repair proteins that are localized to both the nucleus and mitochondria to maintain the integrity of the genomes in each of these locations.

At a basic level, protein localization is determined by binding affinities to other components in the cell. Consider a protein moving through an open compartment. The protein will diffuse freely, undergoing the random Brownian motion resulting from jostling by the surrounding medium; over time, the protein will sample the entire space evenly. Now add a field of anchored binding sites in one half of the compartment. Eventually, the protein will wander into the field of binding sites and bind to one at the edge of the field. Randomly, enough energy will be imparted to the protein from the surrounding medium to allow it to

dissociate from the binding site and resume random diffusion. While the released protein has a 50% chance of wandering back into the open zone, because it was at the edge of the region, it also has a 50% chance of wandering further into the binding zone and binding again for a period. The further into the binding zone the protein wanders, the more likely the protein is to re-bind with the same or another binding site than it is to return to the free zone. The net result of this behavior is that the protein will spend most of the time in the binding zone and relatively little time in the open zone. The differential between these two durations depends directly on the strength of the binding interaction and the concentration of binding sites within the zone. Extended to large numbers of protein molecules and with a large enough differential, the protein is considered to be localized to the binding region. As this exercise demonstrates, it is not the case that a localized protein molecule is locked into a given location. Rather, localization is a statistical phenomenon resulting from the mass action of many protein molecules in a dynamic steady state.

At a small scale, multiple binding sites can be used to produce an increased local concentration of a low-concentration protein, such as is the case with the *Escherichia coli lac* operon. Only the O1 operator site can be bound by LacR to repress transcription of the operon. The adjacent O2 and O3 sites, binding to which does not repress transcription, enhance activity at the O1 site (359). At a large scale, this phenomenon can result in the formation of centers of activity, which are sometimes called foci or puncta when observed microscopically. The most prominent examples of this behavior on a larger scale are nucleoli (360) and centrosomes (361), two entirely self-organizing structures without a membrane barrier. Local concentration is also observed in the case of scaffold proteins that anchor multiple signaling components in close proximity to each other, both allowing efficient transduction from one to the other as well as allowing the same proteins to transduce diverse signaling events despite the ubiquity of their expression (362,363). Such self-organization

may even be involved in the origin of cellular life (364). Locally concentrating related molecules is a critical, central principle of biology.

Localization can be more stringently enforced by a protein-impermeable barrier. Eukaryotic cells take full advantage of this partitioning ability with the endomembrane system. The barrier provides a way to restrict the diffusion of a group of proteins and small molecules within a particular space, and is often used to concentrate two interacting partners together and/or to keep them apart. Physical separation of different biochemical pathways allows a wider variety of processes to co-exist without side-effects on other cellular processes. For instance, processes enhanced by low pH, such as occur in the lysosome, could not be easily accomplished in a prokaryote without negatively impacting proteins that require a more neutral pH. Similarly, the presence of introns can be better tolerated when there is an opportunity to excise them before the translation machinery can start producing protein from them. To take advantage of impermeable boundaries, a cell needs tightly regulated pathways to move proteins across the barriers as necessary. The pathways that accomplish this task are reviewed in the next section, followed by a review of the mechanisms that enable a protein to be differentially distributed to multiple compartments.

Transport pathways

The problem of intracellular protein localization was solved very early in eukaryotic evolution, and the central role of these pathways has ensured that they have been conserved throughout all eukaryotes. These pathways likely co-evolved with the development of the endomembrane system itself (365). The known protein transport pathways are reviewed below.

Nuclear transport

The nucleus contains the majority of the cellular genome and keeps it physically separated from other cellular functions. As noted above, transcription is physically separated

from translation, which provides ample opportunity for RNA processing, including trimming, splicing, folding, and editing, without running the risk of the intermediates inappropriately interfering with normal cellular functions. This separation also enabled an extremely important evolutionary mechanism, exon shuffling, which allowed functional domains to be recombined with other proteins in novel combinations, providing a large amount of raw material for selection to act on (366).

The nucleoplasm is separated from the cytoplasm by the nuclear envelope, a double-bilayer membrane contiguous with the endoplasmic reticulum. Transport through the envelope is controlled by large multiprotein complexes called nuclear pores. The pore complex's role as mediator of transport of large macromolecules makes it a central hub of nuclear regulatory activity, including roles in chromatin organization and transcription (367-369), protein modification (370), DNA repair (371), cell cycle coordination (372), and other functions (373). Components of the nuclear pore, termed nucleoporins, are also some of the longest-lived proteins in postmitotic (non-replicating) cells; for example, lifetimes for the core components in rat brain were measured to be well over 1 year (374). Their long life allows nucleoporins to accumulate oxidative damage over time, which eventually weakens the selectivity barrier of the nuclear pore and may contribute to aging phenotypes (375).

Rather than passing molecules through the lumen of the nuclear envelope, the nuclear pore complex bends the membrane of the nuclear envelope to form a direct connection between the cytoplasm and the nucleoplasm (structure reviewed in (376)). However, this connection is obstructed by the phenylalanine–glycine (FG) repeat–containing unstructured nucleoporin tails that extend into the center of the pore (377). Small molecules (water, ions, nucleotides, amino acids, etc.) and macromolecules up to a molecular weight of approximately 40 kDa can passively diffuse through small (9–10 nm diameter) channels within the

pore (378). Passage of larger macromolecules or macromolecular complexes (up to approximately 39 nm in diameter (379)) requires the assistance of a protein chaperone, a karyopherin. The exact nature of this barrier, and thus the mechanism for facilitated transport through it, is under intensive scrutiny (377,380-396) and a discussion of the merits of the competing models is beyond the scope of this work. The general theme among all of the models is that the interaction of a karyopherin with the FG repeats allows it to hop between binding sites within the pore and thus move through the selectivity barrier with its bound cargo.

Karyopherin movement through the nuclear pore is mediated solely by diffusion, and yet these proteins are able to effectively and efficiently move proteins from one compartment to the other. Directionality is conferred on this process through control of cargo binding and release, which is dependent on the small guanosine triphosphatase (GTPase) RAN (*S. cerevisiae*: Gsp1). Small GTPases comprise a superfamily of proteins that inefficiently cleave the terminal phosphate from GTP to make guanosine diphosphate (GDP) and bind both substrate and product tightly. The function of these proteins depends on whether they are bound to GTP or GDP; GTP is typically associated with the active state. The relative inefficiency of changing states allows fine control over switching between states: GTPase-activating proteins (GAPs) enhance the GTPase activity of the enzyme to convert GTP to GDP; guanine nucleotide–exchange factors (GEFs) enhance the dissociation of GDP from the GTPase and association with a new molecule of GTP. The RANGAPs (Human: RANGAP1 and RANBP1; *S. cerevisiae*: Rna1 and Yrb1) are in the cytosol, while the RANGEF (Human: RCC1; *S. cerevisiae:* Srm1) is in the nucleus, bound to chromatin.

Nuclear transport is illustrated in Figure 1-6. RAN is produced in the cytosol and binds to GDP. It may bind to GTP, but the GTP is quickly hydrolyzed to GDP due to cytosolic RANGAP. RAN-GDP is bound by the non-karyopherin transport factor NTF2 (*S. cerevisiae*:

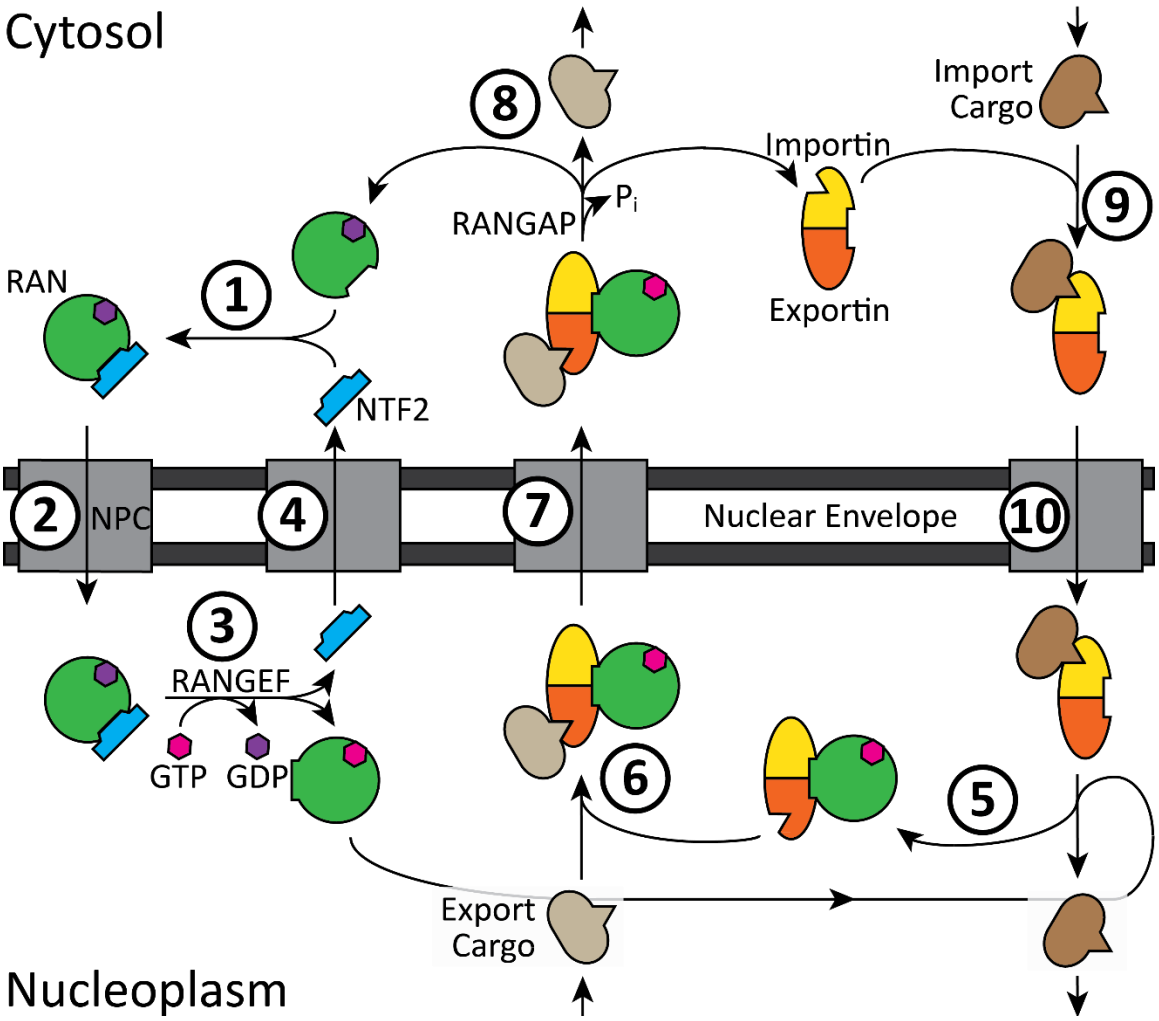

Figure 1-6. The nuclear transport cycle. The components of the nuclear transport cycle are illustrated. 1) RAN-GDP associates with NTF2 and 2) diffuses into the nucleus where 3) RANGEF converts RAN-GDP to RAN-GTP, releasing RAN-GTP from NTF2, which 4) diffuses back out to the cytosol. 5) Karyopherin binds to RAN-GTP, releasing any imported cargo or 6) binding to any exported cargo, then 7) diffuses out to the cytosol where 8) RANGAP converts RAN-GTP to RAN-GDP, releasing the karyopherin and any exported cargo and 9) binding to any imported cargo, then 10) diffuses back into the nucleus. The import adaptor is not shown. It behaves as an import cargo and is exported by another karyopherin.

Ntf2) [1], which escorts RAN into the nucleus [2], where RAN encounters a high concentration of RANGEF. RANGEF causes RAN-GDP to convert to RAN-GTP, inducing a conformational change in RAN releasing NTF2 [3], which diffuses back through the nuclear pore [4], and conferring strong binding affinity to karyopherins [5/6]. The RAN-GTP-karyopherin complex diffuses back through the nuclear pore into the cytosol [7]. Cytosolic RANGAP converts RAN-GTP to RAN-GDP, triggering release of the karyopherin [8]. The karyopherin then diffuses back into the nucleus [9/10]. This process is known as the nuclear transport cycle (reviewed

in (397); Ran and its regulators are reviewed in (398)).

The nuclear transport cycle is used both for import karyopherins (importins) and export karyopherins (exportins). The distinction between the two classes of karyopherins is that importins release their cargo when bound by RAN-GTP [9→10→5], whereas exportins bind their cargo only when bound by RAN-GTP [6→7→8]. The result of this process is that cytosolic proteins bound by an importin are moved into the nucleus, while nuclear proteins bound by an exportin are moved into the cytosol (397).

The importins and exportins which can pass through the nuclear pore and bind to RAN-GTP belong to the karyopherin β superfamily (399). Both human and *S. cerevisiae* genomes encode multiple importins, and most of them bind cargo directly. Importin β (importin β1; Human: KPNB1; *S. cerevisiae*: Kap95), binds most of its cargo through an adaptor karyopherin, importin α (Human: KPNA1 to KPNA7; *S. cerevisiae*: Srp1) (400). Importin α cannot independently diffuse through the nuclear pore and does not directly interact with RAN-GTP, so it relies on an exportin to export it from the nucleus (401,402). The importin α/β system accounts for the nuclear import of most nuclear proteins and is called classical nuclear import (403). Other importins have specialized cargos. For example, transportin 1 (importin β2; Human: TNPO1; *S. cerevisiae*: Kap104) and transportin 2 (Human: TNPO2; *S. cerevisiae*: not present) exclusively recognize hnRNP A1, HuR, and other mRNA-binding proteins (reviewed in (404)), while importin 8 (Human: IPO8; *S. cerevisiae*: not present) imports Argonaute-microRNA (miRNA) complexes into the nucleus for gene silencing (405,406). Importins beyond importin β have not received close attention, and only a handful of cargo proteins are known for each of them (reviewed in (407,408)).

The majority of cargo exported from the nucleus is escorted by exportin 1 (Human: XPO1; *S. cerevisiae*: Crm1), known as classical nuclear export (409). As with the importins, the other exportins have specialized roles. For example: exportin 2 (Human: CSE1L;

S. cerevisiae: Cse1) exports importin α from the nucleus (401,402) and exportin t (Human: XPOT; *S. cerevisiae*: Los1) moves tRNAs out to the cytosol (410,411). Relatively little is known about exportins beyond exportin 1, however (reviewed in (407)).

There are also several bidirectional karyopherins, which can act as both importins and exportins for different cargos. This category includes importin 13 (Human: IPO13; *S. cerevisiae*: not present) (412), exportin 4 (Human: XPO4; *S. cerevisiae*: not present) (413), and exportin 5 (Human: XPO5; *S. cerevisiae*: Msn5) (414). As with many of the other karyopherins, there is still much to learn about these proteins.

The karyopherins mediating the import and export of most cellular proteins recognize their cargo by motifs typically as part of unstructured loops or tails. Importin α most strongly recognizes two basic amino acid residue stretches separated by a short linker (403,415), typified by nucleoplasmin (**KR**PAATKKAGQA**KKKK**L) (416). This sequence is known as a bipartite classical nuclear localization signal (cNLS). Importin α can also recognize a monopartite cNLS composed of a single patch of basic residues, as in the SV40 large T antigen (PK**KKRK**V) (417). Transportin 1 and transportin 2 bind to a loosely hydrophobic- or basic-rich stretch and a proline–tyrosine (PY) separated by a short linker and embedded in a positively charged region (418). Importin 7 recognizes a dual-serine-phosphorylated SPS motif (419). Exportin 1 recognizes an 8–15 residue leucine/hydrophobic-rich motif, termed a classical nuclear export signal (cNES), as seen in cAMP-dependent protein kinase inhibitor α (PKIα) (N**E**LALK**LAGLDI**) (420). The remaining karyopherins either recognize particular protein folds or linear sequences which have yet to be fully defined, but they include basic residues arranged on the surface of a domain and, in at least one case, a hydrophobic stretch (reviewed in (421,422)). Localization signals on one protein are also able to carry along any bound protein partners, allowing proteins without a localization signal to “piggyback” with their partners through the nuclear pore.

Nuclear transport is not limited to soluble proteins. Transmembrane proteins in the outer nuclear membrane containing an NLS in the cytosolic domain can be bound by importins and brought through the nuclear pore and into the inner nuclear membrane. The NLS in these transmembrane proteins is located 150–250 residues away from the membrane, which would provide enough 'slack' for the importin-bound cytosolic domain to move through the pore while allowing the transmembrane domain to remain inside the lipid bilayer and without overly disturbing the organization of the nuclear pore rings (reviewed in (423)).

There are several large proteins which can traverse the nuclear pore independently of karyopherins. Among these transport proteins are NTF2, which is responsible for RAN-GDP nuclear import (424,425); TAP:p15 heterodimer, which is responsible for mRNA export (426); β-catenin (427); proteins containing spectrin repeats, such as dystrophin and actinin (427); Ca^{2+}-bound calmodulin (import) (428,429); and Ca^{2+}-bound calreticulin (export) (430,431). The transport mechanism for β-catenin and spectrin repeat proteins relies on amphiphilic motifs which can readily undergo a conformational change to expose more surface hydrophobicity upon interaction with the FG-repeat proteins in the nuclear pore (427). The mechanisms for the other proteins are not well defined (reviewed in (432)).

Mitochondrial transport

Mitochondria are organelles that derived from an ancient endosymbiosis event of a pre-eukaryote with α-proteobacteria (433,434). Mitochondria are the site of cellular respiration and many metabolic pathways (435,436) and are involved in the control of apoptosis (programmed cell death) (437). The interior mitochondrial matrix contains biosynthetic enzymes, the mitochondrial DNA nucleoids, and transcription and translation machinery. The mitochondrial genome encodes a small number of tRNAs, rRNAs, and certain subunits of the respiratory complexes, whereas most mitochondrial proteins are produced from nuclear genes and imported. The evolutionary retention of a distinct mitochondrial

genome may provide a way to locally sense mitochondrial conditions and regulate mitochondrial function (438). The matrix is enclosed by the highly folded and very protein-rich inner membrane (the folds form structures called cristae), which harbors the electron transport chain, ATP synthases, metabolite carriers, and the protein import machinery. The intermembrane space between the inner and outer membranes is utilized for iron metabolism and hosts the protons pumped out by the electron transport chain to create the proton gradient across the inner membrane. The outer membrane hosts many regulatory factors, including those involved in apoptosis (437), and contains proteins that mediate connections between mitochondria, the endoplasmic reticulum (439), and to microtubule-associated transport motors (440). Mitochondrial structure and functions have been reviewed in (435).

Because the mitochondrion has so many functional locations and due to its bacterial origin, mitochondrial protein import is highly complex. All import begins by the mitochondrial protein (in at least two cases co-translationally (441,442)) binding to and moving through the translocase of the outer membrane (TOM) complex in either an α-helical or an unfolded conformation (443,444). Pathways diverge from this point.

Proteins destined for the mitochondrial matrix carry an N-terminal presequence that forms an amphipathic α-helix (one face of the helix is hydrophobic, the other face is hydrophilic) with a net positive charge within the first approximately 50 residues, called the mitochondrial matrix targeting or localization signal (MTS or MLS) (445). Tom20 recognizes the hydrophobic face of the MTS, while Tom22 recognizes the positively charged face (446,447). The Tom40 pore forms a “polar slide” with increasing affinity towards the intermembrane space, where the Tom22 subunit has another presequence binding domain, which may provide directionality of movement through the pore (448). When the MTS emerges into the intermembrane space, it is recognized by the translocase of the inner membrane 23 (TIM23) complex (449). The positively charged precursor is initially drawn through

TIM23 by the negative electric potential across the inner membrane to the matrix (450); TIM23 is associated with the Complex III and Complex IV proton pumps of the electron transport chain, where the membrane potential gradient would be locally higher (reviewed in (451)). Once the MTS has entered the matrix, a complex associated with TIM called presequence translocase–associated motor (PAM), which includes the mitochondrial chaperone HSP70, consumes ATP to pull the protein the rest of the way across the membranes and into the matrix (452,453). Finally, the MTS is cleaved off by mitochondrial matrix peptidase (MMP) (454).

Proteins of the intermembrane space can be imported through one of two mechanisms. The primary mechanism starts exactly the same as for matrix-bound proteins, with an MTS that moves through TOM and TIM23. Translocation through TIM23 stops when it reaches a nearby hydrophobic sequence and the protein is released into the inner membrane (455). Following this step, MMP in the matrix cleaves off the MTS and either the inner membrane peptidase (IMP) or the rhomboid protease PCP1 releases the protein from the inner membrane (456). The second mechanism of intermembrane import involves a class of small cysteine-rich proteins. Although it is not known how TOM recognizes these proteins, the mechanism requires the activity of the inner membrane–anchored mitochondrial intermembrane space assembly (MIA) machinery, which forms a disulfide bond with the incoming protein and catalyzes the formation of intramolecular disulfide bonds in the incoming protein (457,458) (extensively reviewed in (459-461)). There are a few proteins destined for the matrix which depend on the MIA pathway, including the mammalian AP endonuclease APEX1 (462) and the DNA damage response coordinator tumor protein p53 (TP53) (463), but the details of this pathway have not yet been elucidated.

Inner membrane proteins may also be imported through two pathways. Proteins with large hydrophilic domains in the matrix use the same mechanism as matrix proteins but with

a hydrophobic stop-transfer sequence as used by intermembrane space proteins. The other class of proteins, which includes the channel proteins responsible for moving metabolites (e.g. ADP, ATP) across the inner membrane, are highly hydrophobic and are maintained in the cytosol in an unfolded state by chaperones (e.g. HSP70) (464). The precursors contain internal sequences which bind along with the chaperones to TOM through TOM70 (465). Following binding, the precursor protein is imported through TOM in a loop structure which is recognized by intermembrane space chaperones (466). These chaperones recruit the precursor protein to the TIM22 complex (distinct from TIM23), which then coordinates insertion into the inner membrane, dependent on the membrane potential (467) (reviewed in (468)).

Import of outer membrane proteins depends on the type of protein. β-barrel proteins are imported similarly to inner membrane proteins, except that a C-terminal signal is recognized in the intermembrane space by the sorting and assembly machinery (SAM) complex, which coordinates the assembly of the β-barrel inside the outer membrane (469). Insertion of α-helical outer membrane proteins may occur through a variety of pathways directly from the cytosol and not through TOM, but they have not been well characterized (470,471).

Mitochondrial import pathways have been extensively reviewed recently (451,472, 473). There are no known specific mitochondrial export pathways. Given that most mitochondrial import pathways involve one-way events (presequence cleavage, membrane insertion), this is not surprising. However, such proteins could conceivably exit and not return. Generally, mitochondrial proteins are only released back into the cytosol during apoptosis due to mitochondrial membrane disruption (474), though transient disruption could allow retrograde movement outside of apoptosis.

Chloroplast transport

Chloroplasts are the most well known member of the plastid family of organelles. Similar to mitochondria, they are the product of an ancient endosymbiosis event between a

eukaryote and a photosynthesizing cyanobacteria (475). Like mitochondria, chloroplasts have outer and inner membranes, an intermembrane space, and a stroma (matrix). Chloroplasts also have a third membrane system inside the stroma, the thylakoid membranes, which are somewhat analogous to the cristae of mitochondria, enclosing the thylakoid lumen. Because they share similar origins, many of the principles of chloroplast protein import are similar to those in mitochondria. In fact, many mitochondrial proteins are dual-targeted to chloroplasts using an ambiguous targeting signal. These pathways are reviewed in detail in (476).

Peroxisomal transport

Peroxisomes, and microbodies more broadly, are small, single membrane–bound bodies that compartmentalize and concentrate specific metabolic processes (477). The functions of peroxisomes include the conversion of toxic hydrogen peroxide to water, very long chain fatty acid β-oxidation, and bile acid and plasmalogen synthesis in mammals (478). Translocation of proteins into the peroxisomal matrix has not been completely delineated, but the broad features are known. Proteins targeted to the peroxisome may have a C-terminal 10-residue peroxisomal targeting signal (PTS1) or an N-terminal 9-residue PTS2. Each signal has a corresponding import receptor (Pex5 and the Pex7 complex, respectively), which binds to the signal and inserts into the peroxisomal membrane, bringing the cargo protein into the matrix. The mechanism of this translocation is not clear, but one well-supported model is that a transient pore forms in the peroxisomal membrane which allows folded proteins to move through (479). Transient monoubiquitination triggers exit of the import receptor from the membrane, releasing the receptor back into the cytosol for another round of import (reviewed in (477)). As with nuclear transport, proteins bound to PTS1/2-containing partners may “piggyback” into peroxisomes. While there are no known export pathways, there are reports of proteins exported back into the cytosol (reviewed in (480)).

Ciliary transport

Primary cilia, which are present on many eukaryotic cells, are used for motility and/or sensing functions. Despite not being separated from the cytosol by a membrane, large proteins are selectively transported into the compartment. Relatively little is known about the movement of proteins into and out of the cilium. However, there is a growing body of evidence that cilia have co-opted the nuclear transport machinery (481). Similar to the nuclear pore, the ciliary barrier allows small molecules and macromolecules less than approximately 40 kDa in size to pass through while excluding larger molecules (482), and a high level of RAN-GTP is maintained inside the cilia (483,484). Nucleoporins have been detected at this barrier (482), although their arrangement has not been characterized, and karyopherins have been identified as necessary for some large proteins to cross through the barrier (483,485). Indeed, cilia-localized proteins contain NLS-like and NES-like sequences which are sufficient to import a protein to the nucleus (483) or export it from the nucleus (485).

However, if ciliary proteins use the same transport mechanisms as the nucleus, how is targeting achieved? It has been speculated that kinesin (+ end directed) motor activity is required to move proteins along microtubules to the base of the cilium (483), which is strengthened by recent observations that some nuclear proteins require dynein (- end directed) motor activity to bring them close to the nucleus (486-489). This remains an outstanding question in ciliary transport.

Endomembrane system transport

Perhaps one of the most well studied protein translocation systems is the one responsible for insertion into the endoplasmic reticulum (ER) and/or embedding into the ER membrane, which serves as the entry point for the entire endomembrane system, the plasma membrane, and extracellular destinations. When mRNA is initially read by a ribosome and the produced N-terminal sequence matches a signal peptide, which typically contains several

basic residues followed by a 10–15 residue hydrophobic stretch (490), the ribonucleoprotein signal recognition particle (SRP) binds to the signal peptide and ribosome. This interaction halts translation and recruits the whole complex to the rough ER where the import translocons are concentrated (491-493). The signal peptide is inserted into the translocon channel and protein synthesis resumes, which provides the energy for translocation across the ER membrane (490). Once the signal peptide has emerged into the ER lumen, a peptidase cuts off the signal peptide and the protein folds (494). Internal hydrophobic start- and stop-transfer sequences are required for insertion into the ER membrane (495). However, some recent research has discovered translation-independent modes of mRNA localization to the ER or parts of the ER (reviewed in (496)), and some proteins can translocate post-translationally (reviewed in (497)).

While far less common than ER import, retrotranslocation back into the cytosol can also occur. Receptor tyrosine kinases, their growth factors, and several other proteins are able to traverse the membrane in signaling functions, as discussed below (498). Damaged and misfolded ER proteins may also be exported for cytosolic degradation (498).

Vesicular transport

While largely beyond the scope of this work, vesicular transport shares many features with other transport pathways. Open-compartment localization is initiated by extensive intermolecular interactions between transmembrane proteins, which may be mediated by signal motifs in their cytosolic or lumenal domains and/or by glycosylation. Lumenal proteins are concentrated through interactions with these membrane proteins. The budding process generates a physical barrier, allowing a distinct mix of proteins to be transported to other locations in the endomembrane system and to the cell surface/exterior (reviewed in (499)). These principles are used throughout the endomembrane system for trafficking to and from all compartments.

Mechanisms governing protein distribution and relocalization

The distribution of a protein among cellular compartments depends on: 1) the types and number of localization signals on the protein; 2) the relative strength of each signal; 3) the concentration of freely diffusing molecules; and 4) the concentration and activity of localization signal receptors. Evolution has harnessed all four of these factors to determine and modulate cellular protein distribution. The first three factors affect the localization of a specific protein, while the fourth factor allows global control of entire classes of proteins.

Recent reviews have considered the issue of proteins with multiple localizations (often called "dual-targeting"), typically focusing on specific sets of proteins, such as mitochondrial, chloroplast, or peroxisomal proteins (e.g. (500-503)). Extensive portions of the proteome have been found in or are predicted to occupy multiple cellular compartments (504). However, most of these articles consider a narrow perspective and largely static distributions. The goal of this review is to consider localization as a general, intrinsically dynamic process and to delve into the mechanisms that evolution has developed to shape and control the distribution of a protein throughout the cell and among various compartments.

It has been proposed that substantially similar proteins localized to different compartments be termed "echoforms" (500). While this terminology has some utility for distinguishing protein pools with nominally terminal localization, such as mitochondrial and chloroplast populations (the frame in which the term was proposed), it would not be generally applicable. The term, like those it was modeled on (e.g. isoform, isoenzyme), suggests that it refers to an intrinsic property of the protein, rather than a property conferred by other processes. Thus, the term obscures the dynamic and contingent nature of localization. It would further be inappropriate to refer to proteins without terminal localization as being one putative echoform or another. The lack of general applicability and ontological inaccuracy recommends against adoption of this term. Therefore, this review continues the standard

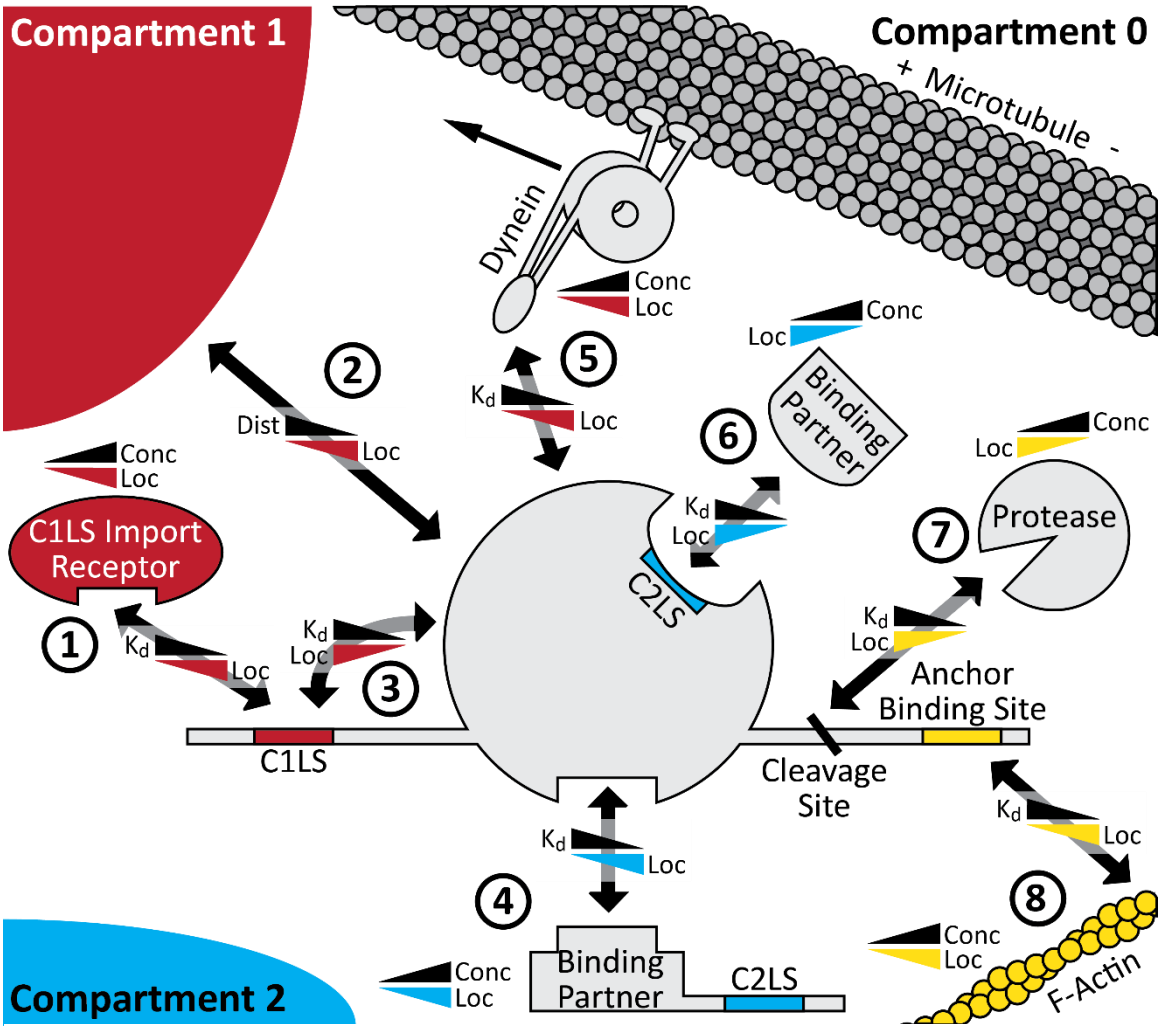

Figure 1-7. Mechanisms governing the regulation of protein localization. This diagram depicts the various ways in which the localization of a protein can be modulated, with a diagram of how the interaction affects localization to the color-coded compartment: compartment 0 is yellow, 1 is red, and 2 is blue. 1) Modulation of the binding strength of signal receptors. 2) Distance to the compartment. 3) Signal masked by conformational change. 4) Contribution of signals by a binding partner. 5) Motor proteins transporting proteins towards the target compartment. 6) Signal masked by a binding partner. 7) Functional domains separated by protease cleavage. 8) Anchoring by binding to a fixed structure.

practice of referring to differently localized proteins as belonging to a localized pool and a protein with minor sequence variations as an "isoform".

This section reviews the common mechanisms invoked to modulate the localization of specific proteins, as gleaned from an extensive review of the literature. This review is concluded by an exploration of several examples demonstrating how these various mechanisms can be integrated to provide fine control over localization.

Modulation of signal strength

One of the most common methods by which cells control protein distribution is to

modify the binding affinity of the localization signal for its matching receptor. This affinity modulation is achieved by post-translationally modifying the protein in or near the localization signal (Figure 1-7[1]). This type of mechanism typically involves serine/threonine phosphorylation, but also occurs through tyrosine phosphorylation, lysine acetylation, or lysine sumoylation. These modifications are primarily employed to interfere with localization receptor binding, but several are used to enhance receptor binding.

The nuclear localization of many cell cycle–related proteins is controlled by importin-blocking phosphorylation in or near the NLS. For example, the paralogous *S. cerevisiae* cell cycle transcription factors Swi5 and Ace2 are phosphorylated in late G_1 by the cyclin-dependent kinase Cdc28 (homolog of mammalian CDK1) on serine residues within and near the NLS (Swi5: ^{636}**KK**YENVVIKRSP**RKRGRPRK**DGTSSVSSS^{674}), which reduces nuclear import receptor binding (505,506). The Cdc14 cell cycle phosphatase reverses this phosphorylation during mitosis, re-enabling nuclear import (507). Importantly, the phosphorylation-dependent regulation of this NLS is completely preserved when it is fused to an exogenous reporter protein (508). Swi5 expressed in mammalian cells exhibits the same regulation, demonstrating the high degree of conservation of this regulatory mechanism (508).

Exportin-blocking phosphorylation in or near an NES is less common than blocking an NLS, but is no less important. For example, this mechanism is involved in regulating the critical tumor suppressor tumor protein p53 (TP53) (351). Threats to genomic integrity like ionizing radiation, ultraviolet light, nitric oxide, and reactive oxygen species (ROS) lead to the activation of various DNA damage response kinases (e.g. ataxia telangiectasia mutated (ATM), ataxia telangiectasia and Rad3 related (ATR), DNA-dependent protein kinase (PRKDC), cyclin-dependent kinase 5 (CDK5)). These kinases phosphorylate serine residues within one of the two TP53 NES motifs (Human: ^{14}**L**S**L**QET**L**S**L**DLWK**L**^{25}), which is sufficient to prevent exportin 1 from binding to the NES, thus preventing nuclear export (509-515). This

effect is stabilized by phospho-S[15]–dependent CK1 phosphorylation of T[18] (516), and the initial phosphorylation requires TP53 to be in the tetrameric form (517).

Phosphorylation may also rarely affect the recognition of the endoplasmic reticulum signal peptide by SRP, as is the case with cytochrome P450 family 1, subfamily A, polypeptide 1 (CYP1A1), a member of the extensive cytochrome P450 superfamily of monooxygenases which are involved in steroid synthesis and metabolize xenobiotic chemicals (518). Protein kinase C phosphorylates CYP1A1 at T[35], reducing the binding affinity of SRP for the signal peptide and allowing CYP1A1 to remain in the cytosol, free to localize elsewhere (519).

There are only a few instances of acetylation interfering with localization signal recognition. One of these instances is RecQ protein–like 4 (RECQL4), a helicase involved in genomic stability. Acetylation by the p300 acetyltransferase within its NLS ([376]**K**QAW**K**Q**K**W **RKK**[386]) efficiently blocks nuclear import (520).

Sumoylation can also rarely interfere with localization signals. There is evidence that retinoic acid receptor (RAR), which translocates to the nucleus to modulate gene transcription upon binding to retinoic acid (vitamin A), is sumoylated near or within its NLS. However, this role for sumoylation of RAR is unclear, as those sumoylation sites are also required for full transcriptional activation (521). Krüppel-like factor 5 (KLF5), a transcription factor involved in proliferation, is sumoylated near its NES. This sumoylation enhances nuclear accumulation of KLF5, but the precise signals and mechanisms involved have not been elucidated (522).

Less common is modification-driven enhancement of localization signals. A prototypical example of a phosphorylation-dependent NLS is found in the mitogen-activated protein kinases 1 and 3 (MAPK1/ERK2/p38MAPK, MAPK3/ERK1/p42MAPK) that mediate growth factor signaling. When an extracellular growth factor is detected by a receptor tyrosine kinase which activates the MAPK pathway, the downstream MAPK kinase 1

(MAP2K1/MEK1) phosphorylates MAPK1/3 on a separate site. This event primes MAPK1/3 for phosphorylation by CK2 on an SPS motif, which is then recognized by importin 7 for nuclear import (419,523). An example of a phosphorylation-dependent NES is the androgen receptor (AR) which translocates to the nucleus on androgen binding to coordinate transcription activation of androgen-responsive genes. AR is phosphorylated near its NES at S^{650} by the stress-inducible JNK or MAPK1/3 which enhances exportin 1 binding (524). This phosphorylation mark can be reversed by PP1 (525).

Signal addition

Another fairly common method of adjusting protein localization is for a binding partner to carry the bound protein between compartments (Figure 1-7[4]). By necessity, this form of regulation requires a pore that can handle multiple folded proteins, such as the nuclear pore, ciliary boundary, and the peroxisomal pore, but not through a system that generally requires an unfolded protein, like the mitochondrial or endoplasmic reticulum translocons. This mechanism allows the localization of a protein to be controlled indirectly by regulating the localization and levels of its binding partner, which plays a very similar role to import receptors.

An excellent example of this mode of regulation is with the minichromosome maintenance proteins (MCM2–7) that form a core part of the replicative helicase and are critical for origin firing in DNA replication (526). MCM2 and MCM3 both have weak NLSs, but when bound together they form a single strong NLS, while MCM3 has an NES. MCM4 through MCM7 do not contain any localization signals, and their localization depends entirely on binding to MCM2 and MCM3 (527,528).

Signal masking

A very efficient method to prevent protein transport is simply to bury the signal. Signals may either be hidden by another protein or ligand (Figure 1-7[6]), or they may be

obscured by an intramolecular interaction (Figure 1-7[3]).

One fairly common method of masking is via a 14-3-3 protein binding to a phosphorylated localization signal. Carbohydrate responsive element binding protein (ChREBP), a glucose-activated transcription factor, is phosphorylated at multiple sites in and around its cNLS by protein kinase A (PKA) (529). 14-3-3β binds at these phosphorylated residues and effectively competes with importin α for binding (530,531). This mark can be removed by the PP2A phosphatase in response to glucose, re-enabling nuclear import (532).

Nuclear receptors, which mediate transcriptional programs based on steroid signals, have an NES in their ligand binding pocket. Without a bound ligand (apo form), nuclear receptors are structurally unstable and are often bound by chaperones, exposing this NES. Binding to the ligand stabilizes the structure and blocks access to the NES by exportin 1, providing an elegant mechanism to induce nuclear accumulation in response to ligand and to clear it from the nucleus when ligand is absent. Nuclear receptors known to employ this mechanism include AR, estrogen receptor α (ERα) and mineralocorticoid receptor (MR) (533).

A well-characterized example of intramolecular masking is the *S. cerevisiae* transcription factor Yap1, which mediates the oxidative stress response (534). Under normal redox conditions, Yap1 shuttles between the nucleus and cytosol due to its NLS and NES. Oxidative stress induces peroxiredoxin to mediate the formation of an intramolecular disulfide bond between a cysteine near the NES and a cysteine on another part of the protein, which folds the NES away, making it inaccessible to exportin 1 (535-538). Because the NES is now masked, Yap1 accumulates in the nucleus and activates the oxidative stress response transcriptional program.

Sequestration and anchoring

Sequestration of a protein within a compartment occurs when the number of binding sites within the compartment significantly reduces the mobility of the protein such that the

rate of export from the compartment is reduced. The term “sequester” is typically applied in the literature to any mechanism which reduces transport to another compartment, but it is productive to distinguish these modes more clearly. However, most proteins considered to be sequestered under this definition may turn out to belong to another mechanistic category upon closer experimental inspection. Anchoring, also called tethering, is a special case of sequestration where the protein is bound to or embedded in an essentially immobile component with respect to other compartments, such as the plasma membrane, cytoskeleton, or chromatin (Figure 1-7[8]). Anchoring reduces the soluble pool available to be acted on by transport receptors.

Calmodulin (CaM), a calcium ion (Ca^{2+})–binding protein and mediator of calcium-sensitive responses (539), is a good example of a sequestered protein. CaM is small enough to passively diffuse through the nuclear pore (540), though all but 5% is bound in the cytosol by many CaM binding proteins, including membrane-bound receptors, at both low and high calcium levels (541). On Ca^{2+} influx, Ca^{2+}-CaM is released from those binding proteins and enters the nucleus through an as-yet-unknown import mechanism. However, high calcium restricts normal nuclear transport (542), and Ca^{2+}-CaM induces nuclear import of certain proteins (432). Once in the nucleus, Ca^{2+}-CaM is retained by nuclear binding proteins (540,543).

Nuclear factor erythroid 2–like 2 (NFE2L2/NRF2), a critical transcription factor coordinating the oxidative stress, antioxidant, and xenobiotic response, is normally anchored in the cytosol by binding to kelch-like ECH-associated protein 1 (KEAP1/INRF2), which in turn is bound to the actin cytoskeleton (544-546). Phosphorylation of NFE2L2 by protein kinase C (PKC), phosphatidylinositol-4,5-bisphosphate 3 kinase (PI3K), MAPK1/3, or oxidative or electrophilic attack on KEAP1 releases NFE2L2 from its anchor and frees it to enter the nucleus (547-553). Once in the nucleus, NFE2L2 dimerizes with other transcription

factors and binds to antioxidant response elements in the genome, which effectively anchors it in the nucleus (553).

Conformational control of localization signals

Some localization signals are not contained within the linear peptide sequence, but are formed by the arrangement of amino acid residues on the surface of the protein. An advantage to such an arrangement is that conformational changes induced by allosteric events can disrupt or reform the localization signal transiently and in response to the state of the protein (Figure 1-8). A well-characterized example of a conformation-sensitive NLS is found in the fatty acid transporter fatty acid binding protein 5 (FABP5). When its fatty acid ligand (e.g. linoleic acid, arachidonic acid) is bound, several basic residues on the surface of the protein align together. This alignment allows an importin to bind to and escort FABP5 and its ligand into the nucleus for delivery to lipid-responsive nuclear receptors. However, when the ligand binding pocket is emptied, those basic residues are shifted out of alignment, rendering the surface unrecognizable to importins and excluding it from the nucleus (554).

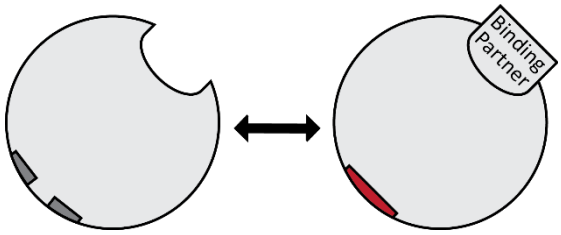

Figure 1-8. Conformational control of localization signals. Events affecting protein conformation may modify the arrangement of amino acids on the protein surface, enabling localization signals to be formed and disrupted in response to cellular conditions.

Motor-driven proximity

The closer a protein molecule is to the compartment of interest, the greater the chance that the protein will be productively bound by a localization signal receptor and translocated into the target compartment. Eukaryotic cells, especially mammalians cells, can reach relatively large sizes. Some cellular signals can be initiated at extreme distances from the target compartment of a signaling protein, such that passive diffusion is a fairly inefficient way to move the protein across the intervening volume. Slower response is only one factor created by large distances. The longer it takes for the activated protein pool to be imported,

the higher the likelihood is that the pool will become deactivated before transport can be completed—bound by an inhibitor, transport receptor dissociated, modification removed, etc. Further, these signals may be transient, which would by necessity only activate a small proportion of the protein. Therefore, it would be beneficial for some proteins to take advantage of microtubule-associated motor proteins to rapidly traverse the cytosol. In fact, the localization of several proteins has been linked to microtubules and motor activity. This mechanism has only been relatively recently discovered and work into the extent of this mode of regulation is ongoing (555-557) (Figure 1-7[2,5]).

Parathyroid hormone–related protein (PTHrP) is generally known for its role as a secreted hormone but it can also be produced in a non-secreted, nuclear-localized form whose functions are still being elucidated (558). PTHrP is also one of the best-characterized proteins that is demonstrated to move to the nucleus in a microtubule- and dynein-dependent manner (559). PTHrP has an NLS which is directly bound by importin β (560) and PTHrP participates in the cell cycle, entering the nucleus in G_1 (561). Intriguingly, the microtubule association sequence (MTAS) directly overlaps with the non-classical NLS, relying on two 4-residue basic clusters, and importin β does not mediate the association with the microtubule (555). When the protein reaches the end of the microtubule at the nuclear periphery, the locally high concentration of importin β allows it to displace the dynein motor complex and import PTHrP into the nucleus (555).

The best-studied protein in terms of ciliary localization is the kinesin-2 motor protein kinesin family member 17 (KIF17), which is involved in ciliogenesis in many cells and axonal transport of proteins and mRNA in neurons (562). KIF17 contains an NLS-like ciliary localization signal which is necessary for import to both the nucleus and the cilium. Efficient localization to the cilium and not the nucleus requires the full length protein, though it has

not yet been demonstrated that the kinesin activity, rather than some other protein interaction, is responsible for specific targeting (483).

Proteolytic disjunction

Most of the mechanisms discussed above are entirely reversible or transient. When a cell needs to change state, a protease may cleave a protein to permanently separate functional domains from each other (Figure 1-7[7]). This mechanism is often employed as a way to release a protein from a strong anchor, such as is the case for membrane-embedded proteins, or it may remove a localization signal.

Many plasma membrane receptors and proteins undergo ligand-induced proteolytic cleavage to release an intracellular domain (ICD) into the cytosol, which often contains an NLS. The most well known protein regulated by this mechanism is the developmental receptor Notch. Upon interaction of Notch with Delta or other ligands expressed on a neighboring cell, γ-secretase and presenilin peptidases cleave the cytosolic side off of the membrane (563-565). The Notch ICD then enters the nucleus, where it coordinates the developmental gene expression program directing cell fate (566,567).

Some soluble proteins undergo proteolytic separation of their signals, which is particularly associated with caspase activity during apoptosis. For example, the mammalian STE20-like kinase 1 (MST1) helps coordinate chromatin condensation and other nuclear apoptotic events (568). Normally the two NESs in its C-terminus keep MST1 from accumulating in the nucleus. During induction of apoptosis, however, caspases can cleave those signals off of the protein, allowing nuclear accumulation (569).

Alternative isoform choice

All of the mechanisms described above allow for rapid mobilization of a pre-existing pool of protein molecules. For longer-term responses, cells can control upstream events to adjust a protein's cellular distribution. One of the major ways this can be accomplished is by

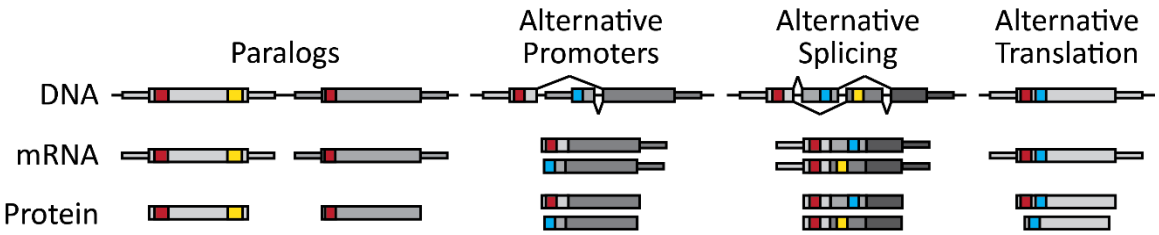

Figure 1-9. Alternative isoform choice produces proteins with different localization signals. Horizontal black line indicates untranscribed regions. Thin rectangle indicates transcribed but untranslated regions. Thick rectangle indicates translated regions. Exons are distinguished by shades of gray and marked with localization signals indicated in color in the translated regions.

generating an alternate isoform of a protein with a different complement of localization signals. Use of paralogous genes, alternative transcription start sites, alternative splicing, and alternative translation initiation can all produce functionally identical proteins but with altered localization (Figure 1-9).

One of the more stark examples of this mode of regulation is the human uracil–DNA glycosylase (UNG). Alternative transcription start sites produce two isoforms with different N-termini: isoform 1 (UNG1) contains an MTS and is localized to mitochondria, while isoform 2 (UNG2) contains an NLS and is localized to nuclei (225). Since each isoform depends on a separate promoter, they are regulated independently, with UNG2 induced more in late G_1 and S phase than UNG1 (208).

The results can also be more subtle, with differences in strength between included localization signals, as is the case with NIMA-related kinase 2 (NEK2) which plays critical roles in mitotic progression (570). NEK2 has three alternate splicing products resulting in protein isoforms differing in the C-terminal residues beyond 370. Isoform A contains a weak bipartite NLS spanning the splice boundary which causes it to distribute evenly to the nucleus and cytosol. Isoform C shares the C-terminus of isoform A but has a shorter spacing between the basic residues, resulting in a stronger NLS and primarily nuclear localization. Isoform B, however, has a completely different C-terminus which lacks the critical second part of the NLS and thus localizes to the cytosol (571). However, it is unclear under which conditions the levels of each isoform are adjusted.

Spastin (SPAST), a microtubule-severing ATPase required for microtubule dynamics in mitosis and neuronal morphogenesis (572), has two alternative translation products from its mRNA, a 60 kDa form and a 68 kDa form. The longer isoform contains 2 NLSs and 2 NESs and shuttles between the nucleus and cytosol but is concentrated in the cytosol, while the shorter isoform has just 1 NLS and no NESs and accumulates in the nucleus (573).

Localized production

One last specific mechanism available to cells to control protein localization is to localize mRNA transcripts near the location where the protein is intended to function (Figure 1-7[2]). As in the case for microtubule-mediated transport, this mode relies on increasing the local concentration of the protein where it is most likely to interact with import receptors, making transport more efficient. One protein that uses this pathway is metallothionein 1A (MT1A), which has its nuclear concentration directly dependent on its mRNA localizing near the nucleus (574). Some mitochondrial (575) and peroxisomal (576) transcripts are directed in this way as well. Localization of protein production is especially important for neurons, which often have extreme distances between the nucleus and sites of protein activity (577). Another form of this mechanism involves co-translational import, as has been most extensively characterized with the endoplasmic reticulum for endomembrane system delivery and/or secretion, as discussed in the "Endomembrane system transport" section. However, several mitochondrial proteins, such as SOD2 and fumerase, are also co-translationally imported (441,442).

Overall concentration

When a protein changes localization, what generally matters for functional regulation is the concentration of the protein within the target compartment, rather than the ratio between compartments. In some cases, especially if the protein's main function is only in one compartment, the only change necessary to increase or decrease the amount of a protein in a

compartment is to increase or decrease the total protein levels. This can be directly controlled through the classical mechanisms of transcription, translation, and degradation of both mRNA and protein, and as such is not discussed here.

A more interesting mechanism is that certain extracellular proteins, either actively secreted or through necrosis, can be actively taken up by neighboring cells. Many of these proteins not only make it into the cytosol, but can localize to other cellular compartments, particularly the nucleus. The exact details of how this may happen is unknown, and as such this is a somewhat contentious area of research. One of the first reports of this phenomenon was fibroblast growth factor 2 (FGF2) (578-581), and has since expanded to include a wide array of growth factors (e.g. growth hormone (582,583)), cytokines (e.g. IL-1 (584)) and other proteins including superoxide dismutase (SOD) (585). Thus, events in one cell can directly lead to an increase in the localization of a protein in another cell.

Combinatorial regulation

The localization of proteins involved in key control points for coordinating cellular responses are frequently regulated through a combination of the above mechanisms, allowing signals to be integrated into a coherent response. Below, several examples are discussed which highlight how cells put these mechanisms to use: protein kinase CDK5, DNA damage checkpoint transcription factor TP53, and the developmental transcription factor GLI.

CDK5: signal affinity, anchoring, proteolytic disjunction

CDK5 is a kinase active in mammalian postmitotic neurons and plays key roles in their development and function, and dysregulation of CDK5 localization is associated with neurodegenerative disease (586). Although CDK5 is a member of the cyclin-dependent kinase family, its two possible regulatory subunits, CDK5R1 (p35) and CDK5R2 (p39), are not cyclins. On its own, CDK5 is localized to the cytosol (587). CDK5R1/2 is cotranslationally, and irreversibly, myristoylated on G^2, which along with an N-terminal positively charged basic

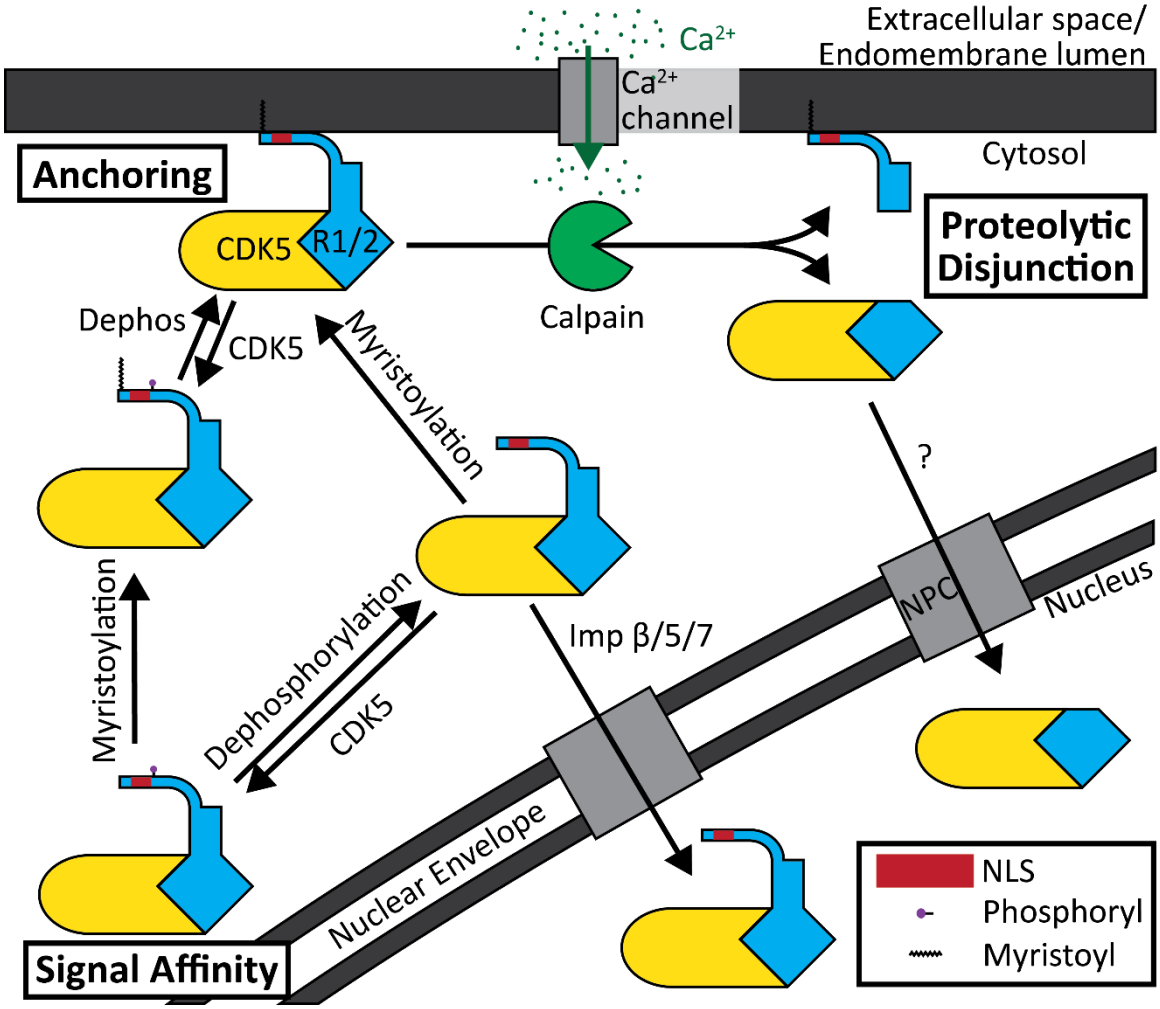

Figure 1-10. Mechanisms regulating CDK5 localization. CDK5 is anchored to cellular membranes through myristoylation of its regulatory subunit, CDK5R1/2. Phosphorylation can release CDK5 complex from the membrane and exclude it from the nucleus by inhibiting importin binding to the NLS. Calcium-induced calpain activity can separate the anchored segment of CDK5R1/2, allowing CDK5 to enter the nucleus through an unknown mechanism.

cluster anchors the protein to the plasma membrane and to perinuclear membranes (including the Golgi apparatus), anchoring CDK5 at those locations (587,588). Without myristoylation, the CDK5 holoenzyme localizes to the nucleus, as the basic clusters are recognized by importins β, 5, and 7 (589), though CDK5R2 is more strongly nuclear than CDK5R1, and there is a small pool of unmodified CDK5 holoenzyme which localizes to the nucleus (590). CDK5-mediated phosphorylation of its regulatory subunits at several N-terminal residues can cause it to dissociate from the membrane and exclude it from the nucleus; specifically, phosphorylation of T^{84} of CDK5R2 within its NLS reduces nuclear localization (590). Calpain protease activity cleaves CDK5R1/2 C-terminal to the basic clusters (591), releasing the

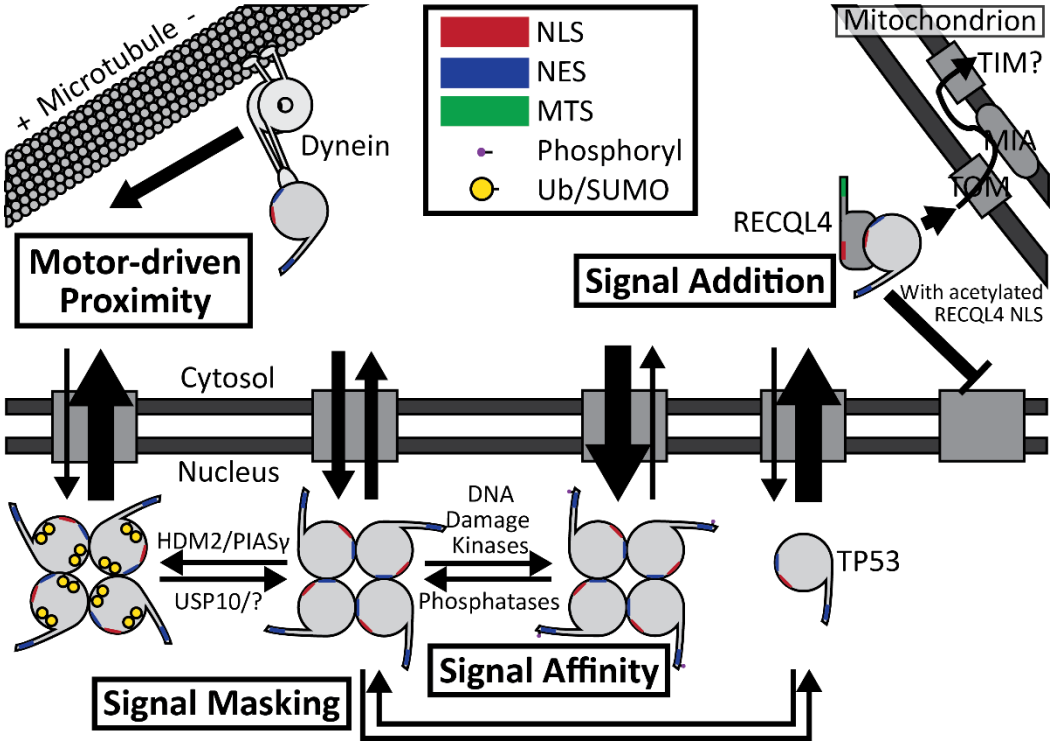

Figure 1-11. Mechanisms regulating TP53 localization. TP53 tetramerization masks an NES which can be exposed by multiple ubiquitin and SUMO modifications. Phosphorylation by DNA damage response kinases can mask a secondary NES. Dynein motor protons bring TP53 closer to the nucleus. RECQL4 recruits TP53 to mitochondria.

holoenzyme from the membrane and allowing the holoenzyme to enter the nucleus through an unknown mechanism (592). The cleaved, released form of CDK5R1/2 is called p25/p29. This proteolytic event occurs en masse when death signals or high calcium stress is present, in which case CDK5 activity triggers neuronal cell death (593). However, normal neuronal activity produces transient calcium pulses, which may induce production of small amounts of CDK5-p25/29 (590). The roles of CDK5 in the nucleus are still being elucidated. These regulatory mechanisms are illustrated in Figure 1-10.

TP53: signal affinity, signal addition, signal masking, motor-driven proximity

The DNA damage response regulator and tumor suppressor TP53 was introduced above as an example of a protein regulated by a phosphorylation-inhibited NES, instigated by the activity of DNA damage response kinases (510). In addition to this NES, a second NES is located within the tetramerization domain, so that this second NES is masked when TP53 tetramerizes (594). TP53 also has an NLS just outside the tetramerization domain (595).

MDM2 (murine; HDM2 in humans) ubiqutinates TP53 at several sites, followed by PIASγ-induced sumoylation, which together disrupt the tetramer enough to expose the masked NES, allowing export (596). USP10 deubiquitinase counters the activity of MDM2/HDM2 (597). TP53 also associates with dynein motors and relies on them for efficient nuclear import (598,599). The RECQL4 helicase, discussed above as an example of a protein with an acetylation-inhibited NLS, can bind to and mask the TP53 NLS. This allows the RECQL4 MTS to dominate, escorting TP53 to mitochondria (600). TP53 is thereby imported into the intermembrane space via the MIA disulfide relay system (463), though it is as yet unclear how it then enters the mitochondrial matrix and gains access to the mitochondrial DNA. These mechanisms are illustrated in Figure 1-11.

GLI: signal affinity, signal masking, anchoring, motor-driven proximity, proteolytic disjunction

The GLI family zinc fingers (GLI1, GLI2, GLI3) and their *Drosophila melanogaster* homolog Cubitus interruptus (Ci) are the key transcription factors in the developmental hedgehog (Human: SHH; *Drosophila*: Hh) signaling pathway (601). These large proteins contain an N-terminal PY-NLS bound by transportin 1, a mid-protein cNLS bound by importin α3, and an NES just C-terminal to the cNLS (602,603). Under non-signaling conditions, the protein is anchored to the microtubule network by a kinesin (Human: KIF7; *Drosophila*: Cos) and the serine/threonine protein kinase fused (Human: STK36; *Drosophila*: Fu) (604,605). Kinesin activity would tend to move the protein away from the nucleus, engaging in the opposite of motor-driven proximity. In mammals, KIF7 is anchored at the base of the primary cilium (606). The anchoring complex recruits PKA, which phosphorylates GLI/Ci immediately adjacent to the cNLS and reduces importin α3 binding (607). Suppressor of fused (Human: SUFU; *Drosophila*: Su(fu)) also binds to and masks the PY-NLS (603,608). Under non-signaling conditions, these mechanisms thoroughly prevent GLI/Ci from localizing to the nucleus. However, the PKA activity recruited by the anchoring complex causes a portion of the protein

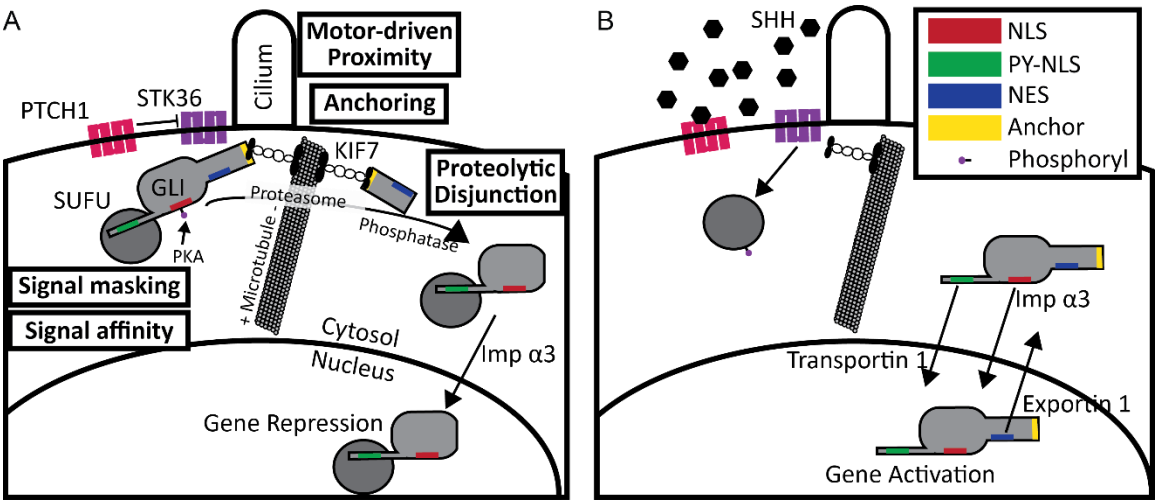

Figure 1-12. Mechanisms regulating GLI localization. A) In absence of SHH signaling, full length GLI is sequestered by masking of its PY-NLS by SUFU, affinity reduction of the GLI cNLS by PKA, and anchoring to KIF7 at the base of the cilium. The full length protein is proteolytically cleaved, separating the NES and anchor from the functional domain, allowing dephosphorylation and nuclear import. B) In the presence of SHH signaling, STK36 phosphorylates SUFU, releasing it from the GLI PY-NLS and allowing the full-length protein to be moved into the nucleus.

to undergo proteasome-dependent proteolysis C-terminal to the cNLS (609). The N-terminal portion is released from the anchoring domain and the NES, allowing the phosphorylation mark on the cNLS to be removed by a cytosolic phosphatase, and the GLI/Ci fragment then enters the nucleus with SUFU/Su(fu) (610), where GLI/Ci engages in transcriptional repression (611).

Under high SHH/Hh signaling, STK36/Fu dimerizes and phosphorylates SUFU/Su(fu) and KIF7/Cos, which exposes the GLI/Ci PY-NLS and primes the cNLS for dephosphorylation (612). The full length protein then enters the nucleus and activates transcription with its retained C-terminal activation domain (602,603,613). NES activity ensures that GLI/Ci shuttles between the nucleus and cytosol. On subsequent loss of SHH/Hh signaling, the shuttling protein is bound by SUFU/Su(fu), which reduces the rate of nuclear re-entry. Then KIF7/Cos recaptures the cytosolic protein and moves it away from the nucleus, meanwhile recruiting PKA to phosphorylate the cNLS and thus fully cutting off nuclear re-entry. This multiplex regulation enables a carefully tuned response dependent on the extracellular SHH /Hh gradient. These mechanisms are illustrated in Figure 1-12.

DNA Repair Protein Localization

While many DNA repair proteins are localized to both the nucleus and mitochondria, the base excision and strand incision repair (BESIR) pathway is unique in that the majority of its components are common between both genome-containing compartments. Despite the extensive biochemical characterization of this critical repair pathway, relatively little is known about the role of localization in its regulation.

Over the past several years, localization of DNA repair proteins has started to be recognized as an important regulatory mechanism. The *S. cerevisiae* DNA *N*-glycosylase responsible for repairing oxidized pyrimidines, Ntg1, localizes to both the nucleus and mitochondria. It was recently discovered that the distribution of Ntg1 between nuclei and mitochondria adjusts depending on the distribution of DNA damage: hydrogen peroxide, which increases whole-cell but not mitochondrial superoxide levels, causes a shift of the protein to nuclei; hydrogen peroxide plus the electron transport chain disruptor antimycin A (614), which increases both mitochondrial and whole-cell superoxide levels, prevents this nuclear shift (125). Additionally, mitochondrial accumulation was reduced in cells lacking mitochondrial DNA (ρ^0), which suggested that DNA damage is the proximal signal leading to this dynamic compartmentalization of Ntg1. This result was followed up by a characterization of the localization sequences of Ntg1 (126). It was found that disrupting the putative nuclear localization signal (NLS) reduced, but did not completely eliminate, nuclear localization, while also eliminating the observed nuclear accumulation upon hydrogen peroxide treatment. These data suggested that the major NLS is responsive to DNA damage or oxidative stress and that a secondary cryptic NLS may be constitutive. This study also found that disrupting both the NLS and mitochondrial matrix targeting signal (MTS) of Ntg1, which did not affect its catalytic activity, reduced survival following hydrogen peroxide or methyl methanesulfonate (MMS) treatment to between 10% and 1% of the wild type protein. Lastly,

the nuclear-only paralog of Ntg1, Ntg2, has largely overlapping functions with Ntg1, and for many purposes can be considered to be a dedicated nuclear pool of the same protein.

A handful of other DNA repair proteins have evidence of being regulated at the level of localization. The *S. cerevisiae* AP endonuclease Apn1 localizes to both the nucleus and mitochondria. A poorly studied cell wall protein, Pir1, masks the Apn1 NLS and thereby increases the proportion of the Apn1 pool imported to mitochondria (317). Meanwhile, the functional mammalian homolog APEX1 localizes to the nucleus and mitochondria from a cytosolic pool in response to oxidative stress (615). APEX1 also contains redox-sensitive cysteine residues which may potentially affect the accessibility of its NES (616) and are involved in mitochondrial import via the MIA pathway (462). Human uracil–DNA glycosylase UNG uses alternate promoters to produce separately regulated nuclear and mitochondrial isoforms; the nuclear promoter is cell-cycle responsive (208,214-217) while the mitochondrial promoter is regulated by oxidative stress (225). Several other mammalians proteins are expressed in multiple isoforms with subtle differences in localization signal strength, including the 8-oxoguanine–DNA glycosylase OGG1, the mispaired adenine–DNA glycosylase MUTYH, and the 8-oxoguanidine triphosphatase MTH1 (617). OGG1 and the single-strand oxopyrimidine–DNA glycosylase NEIL2 have both been associated with microtubules, though whether this association is involved in localization is unknown (164). TDG relocalizes to the nucleus on TP53 activation through an unknown mechanism (618). The methylpurine–DNA glycosylase MPG accumulates in mitochondria on MMS treatment (91). XPF requires binding to RPA to localize to the nucleus (619). Lastly, all DNA repair proteins are potentially subject to anchoring to chromatin, whether through a separate DNA binding domain as with MPG (77), as part of the replication complex as with UNG2 (199,223,224), NEIL1 (155,156), and MUTYH (278,279), with the transcription complex as with NEIL2 (135,149,157), or to the lesions themselves.

There is also evidence of mislocalization of DNA repair proteins in cancer, which may indicate these proteins play a role in oncogenesis. For example, NTHL1 is localized primarily in the nucleus in gastric tissue. In a substantial subset of gastric and colon tumors, however, NTHL1 has been found in the cytoplasm (620,621). NTHL1 has also been found at higher levels in the mitochondria of lung cancer cells compared to normal matched tissue (123). These data hint at a critical role for dysregulation of repair protein localization in cancer.

DNA repair enzymes can act on undamaged DNA at low levels, and overexpression can introduce de novo damage into the genome. Since many repair intermediates are potentially more toxic than the initial lesion, stoichiometrically imbalanced DNA repair pathways can be deleterious. As a result, the levels and activity of these proteins must be tightly controlled. Cells also benefit from being able to repair any DNA damage as rapidly as possible. Thus, dynamic compartmentalization, as observed with Ntg1, would provide a rapid and efficient way to regulate the activity of DNA repair proteins.

Several key questions arose from this research (Figure 1-13): 1) Is dynamic compartmentalization general to BESIR, or is it specific to Ntg1? 2) What DNA lesions are responsible for inducing dynamic compartmentalization? 3) How is information about the lesion transmitted to the protein? 4) How is this information integrated by the target protein to modulate its distribution? However, effectively addressing these questions depended on the availability of methods to measure localization and to introduce compartment-specific base lesions.

Solving these methodological challenges and fully investigating dynamic compartmentalization was the main objective of the work contained herein. Chapter 2 describes an analysis of the *S. cerevisiae* and human BESIR protein sequences for putative localization signals, functional motifs, and possible post-translational modifications. Chapter 3 illustrates

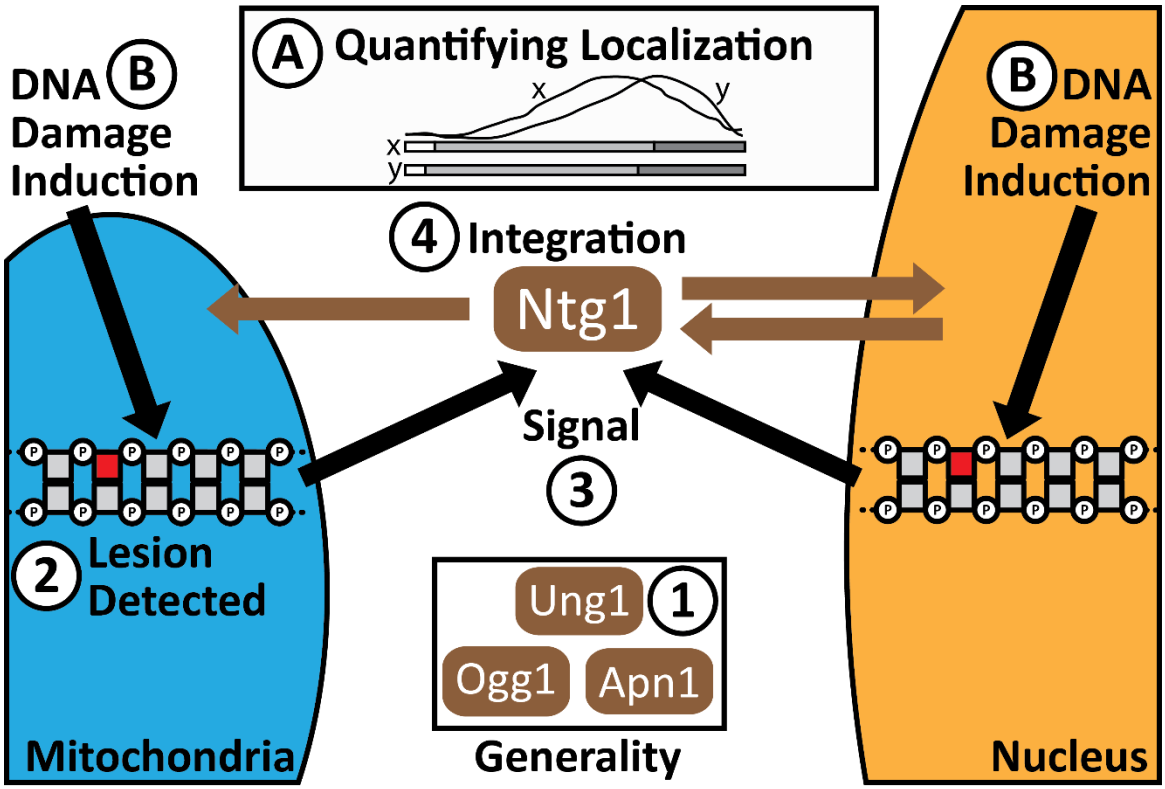

Figure 1-13. Dynamic compartmentalization model, outstanding questions, and challenges. Ntg1 can move in and out of the nucleus and into mitochondria. The distribution of Ntg1 shifts to mitochondria or nuclei in response to a signal transduced in response to compartmental DNA damage. There are four outstanding questions about this model: 1) How general is dynamic compartmentalization? 2) What lesions are detected? 3) What are the signals generated? 4) How are the signals integrated to produce a change in localization of the target protein? There are two major challenges to conducting these studies: A) Quantifying the localization of a target protein. Shown are two hypothetical distributions of a protein between two compartments (above) and how they would be manually scored (below). B) Inducing compartment-specific base damage.

the Quantitative Subcellular Compartmentalization Analysis (Q-SCAn) method developed to robustly, rapidly, and automatically quantify the nucleomitochondrial distribution of BESIR proteins in *S. cerevisiae* to replace the laborious and subjective manual scoring technique previously used for Ntg1. Chapter 4 discusses work characterizing Ung1 and the utility of bisulfite as an agent to induce cytosine deamination in vivo. Chapter 4 also details an experiment to approach the question of which lesions are responsible for generating base damage–dependent ROS. Chapter 5 examines a direct comparison between Q-SCAn and the manual scoring technique in analyzing Ntg1 dynamic compartmentalization. Finally, Chapter 6 discusses the body of work, places it in context, and suggests future research. This dissertation work has developed the methods necessary to vigorously pursue these lines of

inquiry and has provided important insights into base damage signaling and repair protein localization.

References

1. Median Life Expectancy: Wolfram Alpha LLC; [accessed June 11, 2014]. http://www.wolframalpha.com/input/?i=Median+Life+Expectancy.

2. United States Food Consumption: Wolfram Alpha LLC; [accessed June 11, 2014]. http://www.wolframalpha.com/input/?i=United+States+Food+Consumption.

3. Median Human Weight: Wolfram Alpha LLC; [accessed June 11, 2014]. http://www.wolframalpha.com/input/?i=Median+Human+Weight.

4. Kirkwood Thomas BL, Melov S. On the programmed/non-programmed nature of ageing within the life history. Curr Biol. 2011;21(18):R701-R707. doi:10.1016/j.cub.2011.07.020.

5. Sessions AL, Doughty DM, Welander PV, Summons RE, Newman DK. The continuing puzzle of the great oxidation event. Curr Biol. 2009;19(14):R567-574. doi:10.1016/j.cub.2009.05.054.

6. Galhardo RS, Hastings PJ, Rosenberg SM. Mutation as a stress response and the regulation of evolvability. Crit Rev Biochem Mol Biol. 2007;42(5):399-435. doi:10.1080/10409230701648502.

7. Lindahl T. Instability and decay of the primary structure of DNA. Nature. 1993;362(6422):709-715. doi:10.1038/362709a0.

8. Bjelland S, Seeberg E. Mutagenicity, toxicity and repair of DNA base damage induced by oxidation. Mutat Res. 2003;531(1-2):37-80.

9. Morreall JF, Petrova L, Doetsch PW. Transcriptional mutagenesis and its potential roles in the etiology of cancer and bacterial antibiotic resistance. J Cell Physiol. 2013;228(12):2257-2261. doi:10.1002/jcp.24400.

10. Schieber M, Chandel Navdeep S. ROS function in redox signaling and oxidative stress. Curr Biol. 2014;24(10):R453-R462. doi:10.1016/j.cub.2014.03.034.

11. Klaunig JE, Kamendulis LM, Hocevar BA. Oxidative stress and oxidative damage in carcinogenesis. Toxicol Pathol. 2010;38(1):96-109. doi:10.1177/0192623309356453.

12. Cadet J, Mouret S, Ravanat JL, Douki T. Photoinduced damage to cellular DNA: direct and photosensitized reactions. Photochem Photobiol. 2012;88(5):1048-1065. doi:10.1111/j.1751-1097.2012.01200.x.

13. Dizdaroglu M, Jaruga P. Mechanisms of free radical–induced damage to DNA. Free Radical Res. 2012;46(4):382-419. doi:10.3109/10715762.2011.653969.

14. Boorstein RJ, Hilbert TP, Cunningham RP, Teebor GW. Formation and stability of repairable pyrimidine photohydrates in DNA. Biochemistry. 1990;29(46):10455-10460.

15. Labet V, Morell C, Douki T, Cadet J, Eriksson LA, Grand A. Hydrolytic deamination of 5,6-dihydrocytosine in a protic medium: a theoretical study. J Phys Chem A. 2010;114(4):1826-1834. doi:10.1021/jp9049044.

16. Tremblay S, Wagner JR. Dehydration, deamination and enzymatic repair of cytosine glycols from oxidized poly(dG-dC) and poly(dI-dC). Nucleic Acids Res. 2008;36(1):284-293. doi:10.1093/nar/gkm1013.

17. Cadet J, Wagner JR. DNA base damage by reactive oxygen species, oxidizing agents, and UV radiation. Cold Spring Harb Perspect Biol. 2013;5(2). doi:10.1101/cshperspect.a012559.

18. Jena NR, Mishra PC. Formation of ring-opened and rearranged products of guanine: mechanisms and biological significance. Free Radical Biol Med. 2012;53(1):81-94. doi:10.1016/j.freeradbiomed.2012.04.008.

19. Randerath K, Reddy R, Danna TF, Watson WP, Crane AE, Randerath E. Formation of ribonucleotides in DNA modified by oxidative damage in vitro and in vivo. Characterization by ^{32}P-postlabeling. Mutat Res. 1992;275(3-6):355-366. doi:10.1016/0921-8734(92)90038-Q.

20. Admiraal SJ, O'Brien PJ. DNA-*N*-glycosylases process novel *O*-glycosidic sites in DNA. Biochemistry. 2013;52(23):4066-4074. doi:10.1021/bi400218j.

21. Yonekura S, Nakamura N, Yonei S, Zhang-Akiyama QM. Generation, biological consequences and repair mechanisms of cytosine deamination in DNA. J Radiat Res. 2009;50(1):19-26. doi:10.1269/jrr.08080.

22. Doi A, Pack SP, Makino K. Comparison of the molecular influences of NO-induced lesions in DNA strands on the reactivity of polynucleotide kinases, DNA ligases and DNA polymerases. J Biochemistry. 2010;147(5):697-703. doi:10.1093/jb/mvq003.

23. Sedgwick B, Bates PA, Paik J, Jacobs SC, Lindahl T. Repair of alkylated DNA: recent advances. DNA Repair. 2007;6(4):429-442. doi:10.1016/j.dnarep.2006.10.005.

24. Winczura A, Zdzalik D, Tudek B. Damage of DNA and proteins by major lipid peroxidation products in genome stability. Free Radical Res. 2012;46(4):442-459. doi:10.3109/10715762.2012.658516.

25. Billet S, Abbas I, Goff JL, Verdin A, André V, Lafargue P-E, Hachimi A, Cazier F, Sichel F, Shirali P, Garçon G. Genotoxic potential of polycyclic aromatic hydrocarbons-coated onto airborne particulate matter (PM2.5) in human lung epithelial A549 cells. Cancer Lett. 2008;270(1):144-155. doi:10.1016/j.canlet.2008.04.044.

26. Øvrevik J, Arlt VM, Øya E, Nagy E, Mollerup S, Phillips DH, Låg M, Holme JA. Differential effects of nitro-PAHs and amino-PAHs on cytokine and chemokine responses in human bronchial epithelial BEAS-2B cells. Toxicol Appl Pharmacol. 2010;242(3):270-280. doi:10.1016/j.taap.2009.10.017.

27. Nick McElhinny SA, Watts BE, Kumar D, Watt DL, Lundström E-B, Burgers PMJ, Johansson E, Chabes A, Kunkel TA. Abundant ribonucleotide incorporation into DNA by yeast replicative polymerases. Proc Natl Acad Sci USA. 2010;107(11):4949-4954. doi:10.1073/pnas.0914857107.

28. Banyasz A, Douki T, Improta R, Gustavsson T, Onidas D, Vayá I, Perron M, Markovitsi D. Electronic excited states responsible for dimer formation upon UV absorption directly by thymine strands: joint experimental and theoretical study. J Am Chem Soc. 2012;134(36):14834-14845. doi:10.1021/ja304069f.

29. Pommier Y, Leo E, Zhang H, Marchand C. DNA topoisomerases and their poisoning by anticancer and antibacterial drugs. Chem Biol. 2010;17(5):421-433. doi:10.1016/j.chembiol.2010.04.012.

30. Pommier Y, Huang S-yN, Gao R, Das BB, Murai J, Marchand C. Tyrosyl-DNA-phosphodiesterases (TDP1 and TDP2). DNA Repair. 2014;19(0):114-129. doi:10.1016/j.dnarep.2014.03.020.

31. Ellenberger T, Tomkinson AE. Eukaryotic DNA ligases: structural and functional insights. Annu Rev Biochem. 2008;77(1):313-338. doi:10.1146/annurev.biochem.77.061306.123941.

32. Pfeiffer P, Goedecke W, Obe G. Mechanisms of DNA double-strand break repair and their potential to induce chromosomal aberrations. Mutagenesis. 2000;15(4):289-302. doi:10.1093/mutage/15.4.289.

33. Baudat F, Imai Y, de Massy B. Meiotic recombination in mammals: localization and regulation. Nat Rev Genet. 2013;14(11):794-806. doi:10.1038/nrg3573.

34. Lieber MR, Yu K, Raghavan SC. Roles of nonhomologous DNA end joining, V(D)J recombination, and class switch recombination in chromosomal translocations. DNA Repair. 2006;5(9-10):1234-1245. doi:10.1016/j.dnarep.2006.05.013.

35. 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.

36. Doetsch PW, Morey NJ, Swanson RL, Jinks-Robertson S. Yeast base excision repair: interconnections and networks. Prog Nucleic Acid Res Mol Biol. 2001;68:29-39.

37. Aas PA, Otterlei M, Falnes PO, Vagbo CB, Skorpen F, Akbari M, Sundheim O, Bjoras M, Slupphaug G, Seeberg E, Krokan HE. Human and bacterial oxidative demethylases repair alkylation damage in both RNA and DNA. Nature. 2003;421(6925):859-863. doi:10.1038/nature01363.

38. Duncan T, Trewick SC, Koivisto P, Bates PA, Lindahl T, Sedgwick B. Reversal of DNA alkylation damage by two human dioxygenases. Proc Natl Acad Sci USA. 2002;99(26):16660-16665. doi:10.1073/pnas.262589799.

39. Falnes PO. Repair of 3-methylthymine and 1-methylguanine lesions by bacterial and human AlkB proteins. Nucleic Acids Res. 2004;32(21):6260-6267. doi:10.1093/nar/gkh964.

40. Maciejewska AM, Poznanski J, Kaczmarska Z, Krowisz B, Nieminuszczy J, Polkowska-Nowakowska A, Grzesiuk E, Kusmierek JT. AlkB dioxygenase preferentially repairs protonated substrates: specificity against exocyclic adducts and molecular mechanism of action. J Biol Chem. 2013;288(1):432-441. doi:10.1074/jbc.M112.353342.

41. Ringvoll J, Moen MN, Nordstrand LM, Meira LB, Pang B, Bekkelund A, Dedon PC, Bjelland S, Samson LD, Falnes PO, Klungland A. AlkB homologue 2-mediated repair of ethenoadenine lesions in mammalian DNA. Cancer Res. 2008;68(11):4142-4149. doi:10.1158/0008-5472.can-08-0796.

42. Delaney JC, Smeester L, Wong C, Frick LE, Taghizadeh K, Wishnok JS, Drennan CL, Samson LD, Essigmann JM. AlkB reverses etheno DNA lesions caused by lipid oxidation in vitro and in vivo. Nat Struct Mol Biol. 2005;12(10):855-860. doi:10.1038/nsmb996.

43. Daniels DS, Woo TT, Luu KX, Noll DM, Clarke ND, Pegg AE, Tainer JA. DNA binding and nucleotide flipping by the human DNA repair protein AGT. Nat Struct Mol Biol. 2004;11(8):714-720. doi:10.1038/nsmb791.

44. Yang W. Surviving the sun: repair and bypass of DNA UV lesions. Protein Sci. 2011;20(11):1781-1789. doi:10.1002/pro.723.

45. Gerkema MP, Davies WI, Foster RG, Menaker M, Hut RA. The nocturnal bottleneck and the evolution of activity patterns in mammals. Proc R Soc Lond, Ser B: Biol Sci. 2013;280(1765):20130508. doi:10.1098/rspb.2013.0508.

46. 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. doi:10.1016/j.biochi.2003.11.003.

47. Doetsch PW, Cunningham RP. The enzymology of apurinic/apyrimidinic endonucleases. Mutat Res. 1990;236(2-3):173-201. doi:10.1016/0921-8777(90)90004-O.

48. Hanna M, Chow BL, Morey NJ, Jinks-Robertson S, Doetsch PW, Xiao W. Involvement of two endonuclease III homologs in the base excision repair pathway for the processing of DNA alkylation damage in *Saccharomyces cerevisiae*. DNA Repair. 2004;3(1):51-59. doi:10.1016/j.dnarep.2003.09.005.

49. Hanssen-Bauer A, Solvang-Garten K, Akbari M, Otterlei M. X-ray repair cross complementing protein 1 in base excision repair. Int J Mol Sci. 2012;13(12):17210-17229. doi:10.3390/ijms131217210.

50. Sokhansanj BA, Rodrigue GR, Fitch JP, Wilson DM, 3rd. A quantitative model of human DNA base excision repair. I. Mechanistic insights. Nucleic Acids Res. 2002;30(8):1817-1825. doi:10.1093/nar/30.8.1817.

51. Podlutsky AJ, Dianova, II, Wilson SH, Bohr VA, Dianov GL. DNA synthesis and dRPase activities of polymerase β are both essential for single-nucleotide patch base excision repair in mammalian cell extracts. Biochemistry. 2001;40(3):809-813. doi:10.1021/bi002064s.

52. Wiederhold L, Leppard JB, Kedar P, Karimi-Busheri F, Rasouli-Nia A, Weinfeld M, Tomkinson AE, Izumi T, Prasad R, Wilson SH, Mitra S, Hazra TK. AP endonuclease-independent DNA base excision repair in human cells. Mol Cell. 2004;15(2):209-220. doi:10.1016/j.molcel.2004.06.003.

53. Deshpande RA, Wilson TE. Identification of DNA 3′-phosphatase active site residues and their differential role in DNA binding, Mg2+ coordination, and catalysis. Biochemistry. 2004;43(26):8579-8589. doi:10.1021/bi049434n.

54. Sandigursky M, Yacoub A, Kelley MR, Xu Y, Franklin WA, Deutsch WA. The yeast 8-oxoguanine DNA glycosylase (Ogg1) contains a DNA deoxyribophosphodiesterase (dRpase) activity. Nucleic Acids Res. 1997;25(22):4557-4561. doi:10.1093/nar/25.22.4557.

55. Grin IR, Khodyreva SN, Nevinsky GA, Zharkov DO. Deoxyribophosphate lyase activity of mammalian endonuclease VIII-like proteins. FEBS Lett. 2006;580(20):4916-4922. doi:10.1016/j.febslet.2006.08.011.

56. Dianov GL, Souza-Pinto N, Nyaga SG, Thybo T, Stevnsner T, Bohr VA. Base excision repair in nuclear and mitochondrial DNA. Prog Nucleic Acid Res Mol Biol. 2001;68:285-297.

57. 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. doi:10.1093/nar/26.23.5270.

58. Meadows KL, Song B, Doetsch PW. Characterization of AP lyase activities of *Saccharomyces cerevisiae* Ntg1p and Ntg2p: implications for biological function. Nucleic Acids Res. 2003;31(19):5560-5567. doi:10.1093/nar/gkg749.

59. Fortini P, Parlanti E, Sidorkina OM, Laval J, Dogliotti E. The type of DNA glycosylase determines the base excision repair pathway in mammalian cells. J Biol Chem. 1999;274(21):15230-15236. doi:10.1074/jbc.274.21.15230.

60. Yang N, Galick H, Wallace SS. Attempted base excision repair of ionizing radiation damage in human lymphoblastoid cells produces lethal and mutagenic double strand breaks. DNA Repair. 2004;3(10):1323-1334. doi:10.1016/j.dnarep.2004.04.014.

61. Yang N, Chaudhry MA, Wallace SS. Base excision repair by hNTH1 and hOGG1: a two edged sword in the processing of DNA damage in γ-irradiated human cells. DNA Repair. 2006;5(1):43-51. doi:10.1016/j.dnarep.2005.07.003.

62. Gros L, Ishchenko AA, Ide H, Elder RH, Saparbaev MK. The major human AP endonuclease (Ape1) is involved in the nucleotide incision repair pathway. Nucleic Acids Res. 2004;32(1):73-81. doi:10.1093/nar/gkh165.

63. Prorok P, Alili D, Saint-Pierre C, Gasparutto D, Zharkov DO, Ishchenko AA, Tudek B, Saparbaev MK. Uracil in duplex DNA is a substrate for the nucleotide incision repair pathway in human cells. Proc Natl Acad Sci USA. 2013;110(39):E3695-3703. doi:10.1073/pnas.1305624110.

64. Prorok P, Saint-Pierre C, Gasparutto D, Fedorova OS, Ishchenko AA, Leh H, Buckle M, Tudek B, Saparbaev M. Highly mutagenic exocyclic DNA adducts are substrates for the human nucleotide incision repair pathway. PLoS One. 2012;7(12):e51776. doi:10.1371/journal.pone.0051776.

65. Daviet S, Couve-Privat S, Gros L, Shinozuka K, Ide H, Saparbaev M, Ishchenko AA. Major oxidative products of cytosine are substrates for the nucleotide incision repair pathway. DNA Repair. 2007;6(1):8-18. doi:10.1016/j.dnarep.2006.08.001.

66. Ischenko AA, Saparbaev MK. Alternative nucleotide incision repair pathway for oxidative DNA damage. Nature. 2002;415(6868):183-187. doi:10.1038/415183a.

67. Ishchenko AA, Ide H, Ramotar D, Nevinsky G, Saparbaev M. α-anomeric deoxynucleotides, anoxic products of ionizing radiation, are substrates for the endonuclease IV-type AP endonucleases. Biochemistry. 2004;43(48):15210-15216. doi:10.1021/bi049214+.

68. Ishchenko AA, Sanz G, Privezentzev CV, Maksimenko AV, Saparbaev M. Characterisation of new substrate specificities of *Escherichia coli* and *Saccharomyces cerevisiae* AP endonucleases. Nucleic Acids Res. 2003;31(21):6344-6353. doi:10.1093/nar/gkg812.

69. Sparks Justin L, Chon H, Cerritelli Susana M, Kunkel Thomas A, Johansson E, Crouch Robert J, Burgers Peter M. RNase H2-initiated ribonucleotide excision repair. Mol Cell. 2012;47(6):980-986. doi:10.1016/j.molcel.2012.06.035.

70. Xu YJ, DeMott MS, Hwang JT, Greenberg MM, Demple B. Action of human apurinic endonuclease (Ape1) on C1′-oxidized deoxyribose damage in DNA. DNA Repair. 2003;2(2):175-185. doi:10.1016/S1568-7864(02)00194-5.

71. Xu YJ, Kim EY, Demple B. Excision of C-4′-oxidized deoxyribose lesions from double-stranded DNA by human apurinic/apyrimidinic endonuclease (Ape1 protein) and DNA polymerase β. J Biol Chem. 1998;273(44):28837-28844. doi:10.1074/jbc.273.44.28837.

72. Ahel I, Rass U, El-Khamisy SF, Katyal S, Clements PM, McKinnon PJ, Caldecott KW, West SC. The neurodegenerative disease protein aprataxin resolves abortive DNA ligation intermediates. Nature. 2006;443(7112):713-716. doi:10.1038/nature05164.

73. Chakravarti D, Ibeanu GC, Tano K, Mitra S. Cloning and expression in *Escherichia coli* of a human cDNA encoding the DNA repair protein *N*-methylpurine-DNA glycosylase. J Biol Chem. 1991;266(24):15710-15715.

74. Engelward BP, Boosalis MS, Chen BJ, Deng Z, Siciliano MJ, Samson LD. Cloning and characterization of a mouse 3-methyladenine/7-methylguanine/3-methylguanine DNA glycosylase cDNA whose gene maps to chromosome 11. Carcinogenesis. 1993;14(2):175-181. doi:10.1093/carcin/14.2.175.

75. Roy R, Brooks C, Mitra S. Purification and biochemical characterization of recombinant *N*-methylpurine-DNA glycosylase of the mouse. Biochemistry. 1994;33(50):15131-15140.

76. Xiao W, Penugonde V, Rank GH. The *MAG1** 3-methyladenine DNA glycosylase gene is closely linked to the SPT15 TATA-binding *TFIID* gene on chromosome V-R in *Saccharomyces cerevisiae*. Yeast. 1994;10(5):687-691. doi:10.1002/yea.320100513.

77. Watanabe S, Ichimura T, Fujita N, Tsuruzoe S, Ohki I, Shirakawa M, Kawasuji M, Nakao M. Methylated DNA-binding domain 1 and methylpurine-DNA glycosylase link transcriptional repression and DNA repair in chromatin. Proc Natl Acad Sci USA. 2003;100(22):12859-12864. doi:10.1073/pnas.2131819100.

78. Lee CY, Delaney JC, Kartalou M, Lingaraju GM, Maor-Shoshani A, Essigmann JM, Samson LD. Recognition and processing of a new repertoire of DNA substrates by human 3-methyladenine DNA glycosylase (AAG). Biochemistry. 2009;48(9):1850-1861. doi:10.1021/bi8018898.

79. Saparbaev M, Kleibl K, Laval J. *Escherichia coli*, *Saccharomyces cerevisiae*, rat and human 3-methyladenine DNA glycosylases repair 1,N^6-ethenoadenine when present in DNA. Nucleic Acids Res. 1995;23(18):3750-3755. doi:10.1093/nar/23.18.3750.

80. Choudhury S, Adhikari S, Cheema A, Roy R. Evidence of complete cellular repair of 1,N^6-ethenoadenine, a mutagenic and potential damage for human cancer, revealed by a novel method. Mol Cell Biochem. 2008;313(1-2):19-28. doi:10.1007/s11010-008-9737-1.

81. Miao F, Bouziane M, O'Connor TR. Interaction of the recombinant human methylpurine-DNA glycosylase (MPG protein) with oligodeoxyribonucleotides containing either hypoxanthine or abasic sites. Nucleic Acids Res. 1998;26(17):4034-4041. doi:10.1093/nar/26.17.4034.

82. Hitchcock TM, Dong L, Connor EE, Meira LB, Samson LD, Wyatt MD, Cao W. Oxanine DNA glycosylase activity from mammalian alkyladenine glycosylase. J Biol Chem. 2004;279(37):38177-38183. doi:10.1074/jbc.M405882200.

83. Dherin C, Gasparutto D, O'Connor TR, Cadet J, Boiteux S. Excision by the human methylpurine DNA *N*-glycosylase of cyanuric acid, a stable and mutagenic oxidation product of 8-oxo-7,8-dihydroguanine. Int J Radiat Biol. 2004;80(1):21-27. doi:10.1080/09553000310001632976.

84. Admiraal SJ, O'Brien PJ. *N*-glycosyl bond formation catalyzed by human alkyladenine DNA glycosylase. Biochemistry. 2010;49(42):9024-9026. doi:10.1021/bi101380d.

85. Liu Y, Dai H, Xiao W. UAS(MAG1), a yeast cis-acting element that regulates the expression of *MAG1*, is located within the protein coding region of *DDI1*. Mol Gen Genet. 1997;255(5):533-542.

86. Liu Y, Xiao W. Bidirectional regulation of two DNA-damage–inducible genes, *MAG1* and *DDI1*, from *Saccharomyces cerevisiae*. Mol Microbiol. 1997;23(4):777-789. doi:10.1046/j.1365-2958.1997.2701631.x.

87. Zhu Y, Xiao W. Differential regulation of two closely clustered yeast genes, *MAG1* and *DDI1*, by cell-cycle checkpoints. Nucleic Acids Res. 1998;26(23):5402-5408. doi:10.1093/nar/26.23.5402.

88. Zhu Y, Xiao W. Two alternative cell cycle checkpoint pathways differentially control DNA damage–dependent induction of *MAG1* and *DDI1* expression in yeast. Mol Genet Genomics. 2001;266(3):436-444. doi:10.1007/s004380100538.

89. Zhu Y, Xiao W. Pdr3 is required for DNA damage induction of *MAG1* and *DDI1* via a bi-directional promoter element. Nucleic Acids Res. 2004;32(17):5066-5075. doi:10.1093/nar/gkh838.

90. Bouziane M, Miao F, Bates SE, Somsouk L, Sang BC, Denissenko M, O'Connor TR. Promoter structure and cell cycle dependent expression of the human methylpurine-DNA glycosylase gene. Mutat Res. 2000;461(1):15-29. doi:10.1016/S0921-8777(00)00036-7.

91. van Loon B, Samson LD. Alkyladenine DNA glycosylase (AAG) localizes to mitochondria and interacts with mitochondrial single-stranded binding protein (mtSSB). DNA Repair. 2013;12(3):177-187. doi:10.1016/j.dnarep.2012.11.009.

92. Huh W-K, Falvo JV, Gerke LC, Carroll AS, Howson RW, Weissman JS, O'Shea EK. Global analysis of protein localization in budding yeast. Nature. 2003;425(6959):686-691. doi:10.1038/nature02026.

93. Miao F, Bouziane M, Dammann R, Masutani C, Hanaoka F, Pfeifer G, O'Connor TR. 3-Methyladenine-DNA glycosylase (MPG protein) interacts with human RAD23 proteins. J Biol Chem. 2000;275(37):28433-28438. doi:10.1074/jbc.M001064200.

94. Campalans A, Marsin S, Nakabeppu Y, O'Connor T R, Boiteux S, Radicella JP. XRCC1 interactions with multiple DNA glycosylases: a model for its recruitment to base excision repair. DNA Repair. 2005;4(7):826-835. doi:10.1016/j.dnarep.2005.04.014.

95. Xia L, Zheng L, Lee HW, Bates SE, Federico L, Shen B, O'Connor TR. Human 3-methyladenine-DNA glycosylase: effect of sequence context on excision, association with PCNA, and stimulation by AP endonuclease. J Mol Biol. 2005;346(5):1259-1274. doi:10.1016/j.jmb.2005.01.014.

96. Maher RL, Vallur AC, Feller JA, Bloom LB. Slow base excision by human alkyladenine DNA glycosylase limits the rate of formation of AP sites and AP endonuclease 1 does not stimulate base excision. DNA Repair. 2007;6(1):71-81. doi:10.1016/j.dnarep.2006.09.001.

97. Baldwin MR, O'Brien PJ. Human AP endonuclease 1 stimulates multiple-turnover base excision by alkyladenine DNA glycosylase. Biochemistry. 2009;48(25):6022-6033. doi:10.1021/bi900517y.

98. Alseth I, Eide L, Pirovano M, Rognes T, Seeberg E, Bjoras M. 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. Mol Cell Biol. 1999;19(5):3779-3787.

99. Dizdaroglu M, Karahalil B, Senturker S, Buckley TJ, Roldan-Arjona T. Excision of products of oxidative DNA base damage by human NTH1 protein. Biochemistry. 1999;38(1):243-246. doi:10.1021/bi9819071.

100. Eide L, Bjoras M, Pirovano M, Alseth I, Berdal KG, Seeberg E. 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*. Proc Natl Acad Sci USA. 1996;93(20):10735-10740.

101. Eide L, Luna L, Gustad EC, Henderson PT, Essigmann JM, Demple B, Seeberg E. Human endonuclease III acts preferentially on DNA damage opposite guanine residues in DNA. Biochemistry. 2001;40(22):6653-6659. doi:10.1021/bi0028901.

102. Gasparutto D, Ait-Abbas M, Jaquinod M, Boiteux S, Cadet J. Repair and coding properties of 5-hydroxy-5-methylhydantoin nucleosides inserted into DNA oligomers. Chem Res Toxicol. 2000;13(7):575-584. doi:10.1021/tx000005+.

103. Gasparutto D, Muller E, Boiteux S, Cadet J. Excision of the oxidatively formed 5-hydroxyhydantoin and 5-hydroxy-5-methylhydantoin pyrimidine lesions by *Escherichia coli* and *Saccharomyces cerevisiae* DNA *N*-glycosylases. Biochim Biophys Acta. 2009;1790(1):16-24. doi:10.1016/j.bbagen.2008.10.001.

104. Marenstein DR, Chan MK, Altamirano A, Basu AK, Boorstein RJ, Cunningham RP, Teebor GW. Substrate specificity of human endonuclease III (hNTH1). Effect of human APE1 on hNTH1 activity. J Biol Chem. 2003;278(11):9005-9012. doi:10.1074/jbc.M212168200.

105. Miyabe I, Zhang QM, Kino K, Sugiyama H, Takao M, Yasui A, Yonei S. Identification of 5-formyluracil DNA glycosylase activity of human hNTH1 protein. Nucleic Acids Res. 2002;30(15):3443-3448. doi:10.1093/nar/gkf460.

106. You HJ, Swanson RL, Harrington C, Corbett AH, Jinks-Robertson S, Sentürker S, Wallace SS, Boiteux S, Dizdaroglu M, Doetsch PW. *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. doi:10.1021/bi991121i.

107. Redrejo-Rodriguez M, Saint-Pierre C, Couve S, Mazouzi A, Ishchenko AA, Gasparutto D, Saparbaev M. New insights in the removal of the hydantoins, oxidation product of pyrimidines, via the base excision and nucleotide incision repair pathways. PLoS One. 2011;6(7):e21039. doi:10.1371/journal.pone.0021039.

108. Zhang QM, Hashiguchi K, Kino K, Sugiyama H, Yonei S. Ntg1 and Ntg2 proteins as 5-formyluracil-DNA glycosylases/AP lyases in *Saccharomyces cerevisiae*. Int J Radiat Biol. 2003;79(5):341-349.

109. Kino K, Takao M, Miyazawa H, Hanaoka F. A DNA oligomer containing 2,2,4-triamino-5(2H)-oxazolone is incised by human NEIL1 and NTH1. Mutat Res. 2012;734(1-2):73-77. doi:10.1016/j.mrfmmm.2012.03.007.

110. Luna L, Bjoras M, Hoff E, Rognes T, Seeberg E. Cell-cycle regulation, intracellular sorting and induced overexpression of the human NTH1 DNA glycosylase involved in removal of formamidopyrimidine residues from DNA. Mutat Res. 2000;460(2):95-104. doi:10.1016/S0921-8777(00)00015-X.

111. Krokeide SZ, Laerdahl JK, Salah M, Luna L, Cederkvist FH, Fleming AM, Burrows CJ, Dalhus B, Bjoras M. Human NEIL3 is mainly a monofunctional DNA glycosylase removing spiroimindiohydantoin and guanidinohydantoin. DNA Repair. 2013;12(12):1159-1164. doi:10.1016/j.dnarep.2013.04.026.

112. Bruner SD, Nash HM, Lane WS, Verdine GL. Repair of oxidatively damaged guanine in *Saccharomyces cerevisiae* by an alternative pathway. Curr Biol. 1998;8(7):393-403. doi:10.1016/S0960-9822(98)70158-7.

113. Matsumoto Y, Zhang QM, Takao M, Yasui A, Yonei S. *Escherichia coli* Nth and human hNTH1 DNA glycosylases are involved in removal of 8-oxoguanine from 8-oxoguanine/guanine mispairs in DNA. Nucleic Acids Res. 2001;29(9):1975-1981. doi:10.1093/nar/29.9.1975.

114. Kim JE, You HJ, Choi JY, Doetsch PW, Kim JS, Chung MH. Ntg2 of *Saccharomyces cerevisiae* repairs the oxidation products of 8-hydroxyguanine. Biochem Biophys Res Commun. 2001;285(5):1186-1191. doi:10.1006/bbrc.2001.5305.

115. Boiteux S, Guillet M. Abasic sites in DNA: repair and biological consequences in *Saccharomyces cerevisiae*. DNA Repair. 2004;3(1):1-12. doi:10.1016/j.dnarep.2003.10.002.

116. Guillet M, Boiteux S. Origin of endogenous DNA abasic sites in *Saccharomyces cerevisiae*. Mol Cell Biol. 2003;23(22):8386-8394. doi:10.1128/MCB.23.22.8386-8394.2003.

117. Klungland A, Hoss M, Gunz D, Constantinou A, Clarkson SG, Doetsch PW, Bolton PH, Wood RD, Lindahl T. Base excision repair of oxidative DNA damage activated by XPG protein. Mol Cell. 1999;3(1):33-42. doi:10.1016/S1097-2765(00)80172-0.

118. Liu X, Roy R. Truncation of amino-terminal tail stimulates activity of human endonuclease III (hNTH1). J Mol Biol. 2002;321(2):265-276. doi:10.1016/S0022-2836(02)00623-X.

119. Prasad R, Shock DD, Beard WA, Wilson SH. Substrate channeling in mammalian base excision repair pathways: passing the baton. J Biol Chem. 2010;285(52):40479-40488. doi:10.1074/jbc.M110.155267.

120. Oyama M, Wakasugi M, Hama T, Hashidume H, Iwakami Y, Imai R, Hoshino S, Morioka H, Ishigaki Y, Nikaido O, Matsunaga T. Human NTH1 physically interacts with p53 and proliferating cell nuclear antigen. Biochem Biophys Res Commun. 2004;321(1):183-191. doi:10.1016/j.bbrc.2004.06.136.

121. Wolfe KH, Shields DC. Molecular evidence for an ancient duplication of the entire yeast genome. Nature. 1997;387(6634):708-713. doi:10.1038/42711.

122. Kellis M, Birren BW, Lander ES. Proof and evolutionary analysis of ancient genome duplication in the yeast *Saccharomyces cerevisiae*. Nature. 2004;428(6983):617-624. doi:10.1038/nature02424.

123. Karahalil B, Bohr VA, De Souza-Pinto NC. Base excision repair activities differ in human lung cancer cells and corresponding normal controls. Anticancer Res. 2010;30(12):4963-4971.

124. Doudican NA, Song B, Shadel GS, Doetsch PW. Oxidative DNA damage causes mitochondrial genomic instability in *Saccharomyces cerevisiae*. Mol Cell Biol. 2005;25(12):5196-5204. doi:10.1128/mcb.25.12.5196-5204.2005.

125. 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. doi:10.1128/mcb.01357-08.

126. 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. doi:10.1093/nar/gkq108.

127. 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. 2002;1(10):847-854. doi:10.1016/S1568-7864(02)00145-3.

128. O'Rourke TW, Doudican NA, Mackereth MD, Doetsch PW, Shadel GS. Mitochondrial dysfunction due to oxidative mitochondrial DNA damage is reduced through cooperative actions of diverse proteins. Mol Cell Biol. 2002;22(12):4086-4093. doi:10.1128/MCB.22.12.4086-4093.2002.

129. Ikeda S, Biswas T, Roy R, Izumi T, Boldogh I, Kurosky A, Sarker AH, Seki S, Mitra S. 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. 1998;273(34):21585-21593. doi:10.1074/jbc.273.34.21585.

130. Grodick MA, Segal HM, Zwang TJ, Barton JK. DNA-mediated signaling by proteins with 4Fe-4S clusters is necessary for genomic integrity. J Am Chem Soc. 2014;136(17):6470-6478. doi:10.1021/ja501973c.

131. Boal AK, Yavin E, Barton JK. DNA repair glycosylases with a [4Fe-4S] cluster: a redox cofactor for DNA-mediated charge transport? J Inorg Biochem. 2007;101(11-12):1913-1921. doi:10.1016/j.jinorgbio.2007.05.001.

132. Lin JC, Singh RR, Cox DL. Theoretical study of DNA damage recognition via electron transfer from the [4Fe-4S] complex of MutY. Biophys J. 2008;95(7):3259-3268. doi:10.1529/biophysj.108.132183.

133. Romano CA, Sontz PA, Barton JK. Mutants of the base excision repair glycosylase, endonuclease III: DNA charge transport as a first step in lesion detection. Biochemistry. 2011;50(27):6133-6145. doi:10.1021/bi2003179.

134. Yoshitani A, Yoshida M, Ling F. A novel cis-acting element required for DNA damage-inducible expression of yeast *DIN7*. Biochem Biophys Res Commun. 2008;365(1):183-188. doi:10.1016/j.bbrc.2007.10.177.

135. Dou H, Mitra S, Hazra TK. Repair of oxidized bases in DNA bubble structures by human DNA glycosylases NEIL1 and NEIL2. J Biol Chem. 2003;278(50):49679-49684. doi:10.1074/jbc.M308658200.

136. Grin IR, Dianov GL, Zharkov DO. The role of mammalian NEIL1 protein in the repair of 8-oxo-7,8-dihydroadenine in DNA. FEBS Lett. 2010;584(8):1553-1557. doi:10.1016/j.febslet.2010.03.009.

137. Hailer MK, Slade PG, Martin BD, Rosenquist TA, Sugden KD. Recognition of the oxidized lesions spiroiminodihydantoin and guanidinohydantoin in DNA by the mammalian base excision repair glycosylases NEIL1 and NEIL2. DNA Repair. 2005;4(1):41-50. doi:10.1016/j.dnarep.2004.07.006.

138. Hazra TK, Izumi T, Boldogh I, Imhoff B, Kow YW, Jaruga P, Dizdaroglu M, Mitra S. Identification and characterization of a human DNA glycosylase for repair of modified bases in oxidatively damaged DNA. Proc Natl Acad Sci USA. 2002;99(6):3523-3528. doi:10.1073/pnas.062053799.

139. Hazra TK, Kow YW, Hatahet Z, Imhoff B, Boldogh I, Mokkapati SK, Mitra S, Izumi T. Identification and characterization of a novel human DNA glycosylase for repair of cytosine-derived lesions. J Biol Chem. 2002;277(34):30417-30420. doi:10.1074/jbc.C200355200.

140. Jaruga P, Birincioglu M, Rosenquist TA, Dizdaroglu M. Mouse NEIL1 protein is specific for excision of 2,6-diamino-4-hydroxy-5-formamidopyrimidine and 4,6-diamino-5-formamidopyrimidine from oxidatively damaged DNA. Biochemistry. 2004;43(50):15909-15914. doi:10.1021/bi048162l.

141. Jia L, Shafirovich V, Geacintov NE, Broyde S. Lesion specificity in the base excision repair enzyme hNeil1: modeling and dynamics studies. Biochemistry. 2007;46(18):5305-5314. doi:10.1021/bi062269m.

142. Katafuchi A, Nakano T, Masaoka A, Terato H, Iwai S, Hanaoka F, Ide H. Differential specificity of human and *Escherichia coli* endonuclease III and VIII homologues for oxidative base lesions. J Biol Chem. 2004;279(14):14464-14471. doi:10.1074/jbc.M400393200.

143. Krishnamurthy N, Zhao X, Burrows CJ, David SS. Superior removal of hydantoin lesions relative to other oxidized bases by the human DNA glycosylase hNEIL1. Biochemistry. 2008;47(27):7137-7146. doi:10.1021/bi800160s.

144. McKibbin PL, Fleming AM, Towheed MA, Van Houten B, Burrows CJ, David SS. Repair of hydantoin lesions and their amine adducts in DNA by base and nucleotide excision repair. J Am Chem Soc. 2013;135(37):13851-13861. doi:10.1021/ja4059469.

145. Takao M, Kanno S, Kobayashi K, Zhang QM, Yonei S, van der Horst GT, Yasui A. A back-up glycosylase in Nth1 knock-out mice is a functional Nei (endonuclease VIII) homologue. J Biol Chem. 2002;277(44):42205-42213. doi:10.1074/jbc.M206884200.

146. Zhang QM, Yonekura S, Takao M, Yasui A, Sugiyama H, Yonei S. DNA glycosylase activities for thymine residues oxidized in the methyl group are functions of the hNEIL1 and hNTH1 enzymes in human cells. DNA Repair. 2005;4(1):71-79. doi:10.1016/j.dnarep.2004.08.002.

147. Zhao X, Krishnamurthy N, Burrows CJ, David SS. Mutation versus repair: NEIL1 removal of hydantoin lesions in single-stranded, bulge, bubble, and duplex DNA contexts. Biochemistry. 2010;49(8):1658-1666. doi:10.1021/bi901852q.

148. Morland I, Rolseth V, Luna L, Rognes T, Bjoras M, Seeberg E. Human DNA glycosylases of the bacterial Fpg/MutM superfamily: an alternative pathway for the repair of 8-oxoguanine and other oxidation products in DNA. Nucleic Acids Res. 2002;30(22):4926-4936. doi:10.1093/nar/gkf618.

149. Zhou J, Liu M, Fleming AM, Burrows CJ, Wallace SS. Neil3 and NEIL1 DNA glycosylases remove oxidative damages from quadruplex DNA and exhibit preferences for lesions in the telomeric sequence context. J Biol Chem. 2013;288(38):27263-27272. doi:10.1074/jbc.M113.479055.

150. Dobbs TA, Palmer P, Maniou Z, Lomax ME, O'Neill P. Interplay of two major repair pathways in the processing of complex double-strand DNA breaks. DNA Repair. 2008;7(8):1372-1383. doi:10.1016/j.dnarep.2008.05.001.

151. Parsons JL, Kavli B, Slupphaug G, Dianov GL. NEIL1 is the major DNA glycosylase that processes 5-hydroxyuracil in the proximity of a DNA single-strand break. Biochemistry. 2007;46(13):4158-4163. doi:10.1021/bi0622569.

152. Dong L, Meira LB, Hazra TK, Samson LD, Cao W. Oxanine DNA glycosylase activities in mammalian systems. DNA Repair. 2008;7(1):128-134. doi:10.1016/j.dnarep.2007.09.004.

153. Couve-Privat S, Mace G, Rosselli F, Saparbaev MK. Psoralen-induced DNA adducts are substrates for the base excision repair pathway in human cells. Nucleic Acids Res. 2007;35(17):5672-5682. doi:10.1093/nar/gkm592.

154. Hegde ML, Theriot CA, Das A, Hegde PM, Guo Z, Gary RK, Hazra TK, Shen B, Mitra S. Physical and functional interaction between human oxidized base-specific DNA glycosylase NEIL1 and flap endonuclease 1. J Biol Chem. 2008;283(40):27028-27037. doi:10.1074/jbc.M802712200.

155. Dou H, Theriot CA, Das A, Hegde ML, Matsumoto Y, Boldogh I, Hazra TK, Bhakat KK, Mitra S. Interaction of the human DNA glycosylase NEIL1 with proliferating cell nuclear antigen. The potential for replication-associated repair of oxidized bases in mammalian genomes. J Biol Chem. 2008;283(6):3130-3140. doi:10.1074/jbc.M709186200.

156. Hegde ML, Hegde PM, Bellot LJ, Mandal SM, Hazra TK, Li GM, Boldogh I, Tomkinson AE, Mitra S. Prereplicative repair of oxidized bases in the human genome is mediated by NEIL1 DNA glycosylase together with replication proteins. Proc Natl Acad Sci USA. 2013;110(33):E3090-3099. doi:10.1073/pnas.1304231110.

157. Banerjee D, Mandal SM, Das A, Hegde ML, Das S, Bhakat KK, Boldogh I, Sarkar PS, Mitra S, Hazra TK. Preferential repair of oxidized base damage in the transcribed genes of mammalian cells. J Biol Chem. 2011;286(8):6006-6016. doi:10.1074/jbc.M110.198796.

158. Hegde ML, Banerjee S, Hegde PM, Bellot LJ, Hazra TK, Boldogh I, Mitra S. Enhancement of NEIL1 protein-initiated oxidized DNA base excision repair by heterogeneous nuclear ribonucleoprotein U (hnRNP-U) via direct interaction. J Biol Chem. 2012;287(41):34202-34211. doi:10.1074/jbc.M112.384032.

159. Muftuoglu M, de Souza-Pinto NC, Dogan A, Aamann M, Stevnsner T, Rybanska I, Kirkali G, Dizdaroglu M, Bohr VA. Cockayne syndrome group B protein stimulates repair of formamidopyrimidines by NEIL1 DNA glycosylase. J Biol Chem. 2009;284(14):9270-9279. doi:10.1074/jbc.M807006200.

160. Das A, Hazra TK, Boldogh I, Mitra S, Bhakat KK. Induction of the human oxidized base-specific DNA glycosylase NEIL1 by reactive oxygen species. J Biol Chem. 2005;280(42):35272-35280. doi:10.1074/jbc.M505526200.

161. Kinslow CJ, El-Zein RA, Rondelli CM, Hill CE, Wickliffe JK, Abdel-Rahman SZ. Regulatory regions responsive to oxidative stress in the promoter of the human DNA glycosylase gene NEIL2. Mutagenesis. 2010;25(2):171-177. doi:10.1093/mutage/gep058.

162. Hu J, de Souza-Pinto NC, Haraguchi K, Hogue BA, Jaruga P, Greenberg MM, Dizdaroglu M, Bohr VA. Repair of formamidopyrimidines in DNA involves different glycosylases: role of the OGG1, NTH1, and NEIL1 enzymes. J Biol Chem. 2005;280(49):40544-40551. doi:10.1074/jbc.M508772200.

163. Mandal SM, Hegde ML, Chatterjee A, Hegde PM, Szczesny B, Banerjee D, Boldogh I, Gao R, Falkenberg M, Gustafsson CM, Sarkar PS, Hazra TK. Role of human DNA glycosylase Nei-like 2 (NEIL2) and single strand break repair protein polynucleotide kinase 3′-phosphatase in maintenance of mitochondrial genome. J Biol Chem. 2012;287(4):2819-2829. doi:10.1074/jbc.M111.272179.

164. Conlon KA, Miller H, Rosenquist TA, Zharkov DO, Berrios M. The murine DNA glycosylase NEIL2 (mNEIL2) and human DNA polymerase β bind microtubules in situ and in vitro. DNA Repair. 2005;4(4):419-431. doi:10.1016/j.dnarep.2004.10.010.

165. Yeo J, Goodman RA, Schirle NT, David SS, Beal PA. RNA editing changes the lesion specificity for the DNA repair enzyme NEIL1. Proc Natl Acad Sci USA. 2010;107(48):20715-20719. doi:10.1073/pnas.1009231107.

166. Aburatani H, Hippo Y, Ishida T, Takashima R, Matsuba C, Kodama T, Takao M, Yasui A, Yamamoto K, Asano M. Cloning and characterization of mammalian 8-hydroxyguanine-specific DNA glycosylase/apurinic, apyrimidinic lyase, a functional mutM homologue. Cancer Res. 1997;57(11):2151-2156.

167. Bjoras M, Luna L, Johnsen B, Hoff E, Haug T, Rognes T, Seeberg E. Opposite base-dependent reactions of a human base excision repair enzyme on DNA containing 7,8-dihydro-8-oxoguanine and abasic sites. EMBO J. 1997;16(20):6314-6322. doi:10.1093/emboj/16.20.6314.

168. Lu R, Nash HM, Verdine GL. A mammalian DNA repair enzyme that excises oxidatively damaged guanines maps to a locus frequently lost in lung cancer. Curr Biol. 1997;7(6):397-407. doi:10.1016/S0960-9822(06)00187-4.

169. Radicella JP, Dherin C, Desmaze C, Fox MS, Boiteux S. Cloning and characterization of *hOGG1*, a human homolog of the *OGG1* gene of *Saccharomyces cerevisiae*. Proc Natl Acad Sci USA. 1997;94(15):8010-8015.

170. Rosenquist TA, Zharkov DO, Grollman AP. Cloning and characterization of a mammalian 8-oxoguanine DNA glycosylase. Proc Natl Acad Sci USA. 1997;94(14):7429-7434.

171. Shinmura K, Kasai H, Sasaki A, Sugimura H, Yokota J. 8-hydroxyguanine (7,8-dihydro-8-oxoguanine) DNA glycosylase and AP lyase activities of hOGG1 protein and their substrate specificity. Mutat Res. 1997;385(1):75-82. doi:10.1016/S0921-8777(97)00041-4.

172. Dherin C, Radicella JP, Dizdaroglu M, Boiteux S. Excision of oxidatively damaged DNA bases by the human α-hOgg1 protein and the polymorphic α-hOgg1(Ser326Cys) protein which is frequently found in human populations. Nucleic Acids Res. 1999;27(20):4001-4007. doi:10.1093/nar/27.20.4001.

173. Zharkov DO, Rosenquist TA, Gerchman SE, Grollman AP. Substrate specificity and reaction mechanism of murine 8-oxoguanine-DNA glycosylase. J Biol Chem. 2000;275(37):28607-28617. doi:10.1074/jbc.M002441200.

174. Sidorenko VS, Nevinsky GA, Zharkov DO. Specificity of stimulation of human 8-oxoguanine-DNA glycosylase by AP endonuclease. Biochem Biophys Res Commun. 2008;368(1):175-179. doi:10.1016/j.bbrc.2008.01.076.

175. Boldogh I, Hajas G, Aguilera-Aguirre L, Hegde ML, Radak Z, Bacsi A, Sur S, Hazra TK, Mitra S. Activation of ras signaling pathway by 8-oxoguanine DNA glycosylase bound to its excision product, 8-oxoguanine. J Biol Chem. 2012;287(25):20769-20773. doi:10.1074/jbc.C112.364620.

176. Nash HM, Bruner SD, Scharer OD, Kawate T, Addona TA, Spooner E, Lane WS, Verdine GL. Cloning of a yeast 8-oxoguanine DNA glycosylase reveals the existence of a base-excision DNA-repair protein superfamily. Curr Biol. 1996;6(8):968-980. doi:10.1016/S0960-9822(02)00641-3.

177. van der Kemp PA, Thomas D, Barbey R, de Oliveira R, Boiteux S. Cloning and expression in *Escherichia coli* of the *OGG1* gene of *Saccharomyces cerevisiae*, which codes for a DNA glycosylase that excises 7,8-dihydro-8-oxoguanine and 2,6-diamino-4-hydroxy-5-*N*-methylformamidopyrimidine. Proc Natl Acad Sci USA. 1996;93(11):5197-5202.

178. Girard PM, Guibourt N, Boiteux S. The Ogg1 protein of *Saccharomyces cerevisiae*: a 7,8-dihydro-8-oxoguanine DNA glycosylase/AP lyase whose lysine 241 is a critical residue for catalytic activity. Nucleic Acids Res. 1997;25(16):3204-3211. doi:10.1093/nar/25.16.3204.

179. Karahalil B, Girard PM, Boiteux S, Dizdaroglu M. Substrate specificity of the Ogg1 protein of *Saccharomyces cerevisiae*: excision of guanine lesions produced in DNA by ionizing radiation- or hydrogen peroxide/metal ion-generated free radicals. Nucleic Acids Res. 1998;26(5):1228-1233. doi:10.1093/nar/26.5.1228.

180. Hill JW, Hazra TK, Izumi T, Mitra S. Stimulation of human 8-oxoguanine-DNA glycosylase by AP-endonuclease: potential coordination of the initial steps in base excision repair. Nucleic Acids Res. 2001;29(2):430-438. doi:10.1093/nar/29.2.430.

181. Saitoh T, Shinmura K, Yamaguchi S, Tani M, Seki S, Murakami H, Nojima Y, Yokota J. Enhancement of OGG1 protein AP lyase activity by increase of APEX protein. Mutat Res. 2001;486(1):31-40. doi:10.1016/S0921-8777(01)00078-7.

182. Mokkapati SK, Wiederhold L, Hazra TK, Mitra S. Stimulation of DNA glycosylase activity of OGG1 by NEIL1: functional collaboration between two human DNA glycosylases. Biochemistry. 2004;43(36):11596-11604. doi:10.1021/bi049097i.

183. Monden Y, Arai T, Asano M, Ohtsuka E, Aburatani H, Nishimura S. Human MMH (OGG1) type 1a protein is a major enzyme for repair of 8-hydroxyguanine lesions in human cells. Biochem Biophys Res Commun. 1999;258(3):605-610. doi:10.1006/bbrc.1999.0649.

184. Takao M, Aburatani H, Kobayashi K, Yasui A. Mitochondrial targeting of human DNA glycosylases for repair of oxidative DNA damage. Nucleic Acids Res. 1998;26(12):2917-2922. doi:10.1093/nar/26.12.2917.

185. Dantzer F, Luna L, Bjoras M, Seeberg E. Human OGG1 undergoes serine phosphorylation and associates with the nuclear matrix and mitotic chromatin in vivo. Nucleic Acids Res. 2002;30(11):2349-2357. doi:10.1093/nar/30.11.2349.

186. Nishioka K, Ohtsubo T, Oda H, Fujiwara T, Kang D, Sugimachi K, Nakabeppu Y. Expression and differential intracellular localization of two major forms of human 8-oxoguanine DNA glycosylase encoded by alternatively spliced *OGG1* mRNAs. Mol Biol Cell. 1999;10(5):1637-1652. doi:10.1091/mbc.10.5.1637.

187. Conlon KA, Zharkov DO, Berrios M. Cell cycle regulation of the murine 8-oxoguanine DNA glycosylase (mOGG1): mOGG1 associates with microtubules during interphase and mitosis. DNA Repair. 2004;3(12):1601-1615. doi:10.1016/j.dnarep.2004.06.011.

188. Lee MR, Kim SH, Cho HJ, Lee KY, Moon AR, Jeong HG, Lee JS, Hyun JW, Chung MH, You HJ. Transcription factors NF-YA regulate the induction of human OGG1 following DNA-alkylating agent methylmethane sulfonate (MMS) treatment. J Biol Chem. 2004;279(11):9857-9866. doi:10.1074/jbc.M311132200.

189. Singh B, Chatterjee A, Ronghe AM, Bhat NK, Bhat HK. Antioxidant-mediated up-regulation of *OGG1* via NRF2 induction is associated with inhibition of oxidative DNA damage in estrogen-induced breast cancer. BMC Cancer. 2013;13:253. doi:10.1186/1471-2407-13-253.

190. Hegde V, Wang M, Deutsch WA. Human ribosomal protein S3 interacts with DNA base excision repair proteins hAPE/Ref-1 and hOGG1. Biochemistry. 2004;43(44):14211-14217. doi:10.1021/bi049234b.

191. Hegde V, Wang M, Deutsch WA. Characterization of human ribosomal protein S3 binding to 7,8-dihydro-8-oxoguanine and abasic sites by surface plasmon resonance. DNA Repair. 2004;3(2):121-126. doi:10.1016/j.dnarep.2003.10.004.

192. Bhakat KK, Mokkapati SK, Boldogh I, Hazra TK, Mitra S. Acetylation of human 8-oxoguanine-DNA glycosylase by p300 and its role in 8-oxoguanine repair in vivo. Mol Cell Biol. 2006;26(5):1654-1665. doi:10.1128/MCB.26.5.1654-1665.2006.

193. Lucas-Lledo JI, Maddamsetti R, Lynch M. Phylogenomic analysis of the uracil-DNA glycosylase superfamily. Mol Biol Evol. 2011;28(3):1307-1317. doi:10.1093/molbev/msq318.

194. Meyer-Siegler K, Mauro DJ, Seal G, Wurzer J, deRiel JK, Sirover MA. A human nuclear uracil DNA glycosylase is the 37-kDa subunit of glyceraldehyde-3-phosphate dehydrogenase. Proc Natl Acad Sci USA. 1991;88(19):8460-8464.

195. Muller SJ, Caradonna S. Cell cycle regulation of a human cyclin-like gene encoding uracil-DNA glycosylase. J Biol Chem. 1993;268(2):1310-1319.

196. Muller SJ, Caradonna S. Isolation and characterization of a human cDNA encoding uracil-DNA glycosylase. Biochim Biophys Acta. 1991;1088(2):197-207. doi:10.1016/0167-4781(91)90055-Q.

197. Domena JD, Mosbaugh DW. Purification of nuclear and mitochondrial uracil-DNA glycosylase from rat liver. Identification of two distinct subcellular forms. Biochemistry. 1985;24(25):7320-7328.

198. Domena JD, Timmer RT, Dicharry SA, Mosbaugh DW. Purification and properties of mitochondrial uracil-DNA glycosylase from rat liver. Biochemistry. 1988;27(18):6742-6751.

199. Krokan H. Preferential association of uracil-DNA glycosylase activity with replicating SV40 minichromosomes. FEBS Lett. 1981;133(1):89-91. doi:10.1016/0014-5793(81)80477-2.

200. Otterlei M, Warbrick E, Nagelhus TA, Haug T, Slupphaug G, Akbari M, Aas PA, Steinsbekk K, Bakke O, Krokan HE. Post-replicative base excision repair in replication foci. EMBO J. 1999;18(13):3834-3844. doi:10.1093/emboj/18.13.3834.

201. Slupphaug G, Eftedal I, Kavli B, Bharati S, Helle NM, Haug T, Levine DW, Krokan HE. Properties of a recombinant human uracil-DNA glycosylase from the UNG gene and evidence that UNG encodes the major uracil-DNA glycosylase. Biochemistry. 1995;34(1):128-138.

202. Impellizzeri KJ, Anderson B, Burgers PM. The spectrum of spontaneous mutations in a *Saccharomyces cerevisiae* uracil-DNA-glycosylase mutant limits the function of this enzyme to cytosine deamination repair. J Bacteriol. 1991;173(21):6807-6810.

203. Percival KJ, Klein MB, Burgers PM. Molecular cloning and primary structure of the uracil-DNA-glycosylase gene from *Saccharomyces cerevisiae*. J Biol Chem. 1989;264(5):2593-2598.

204. Burgers PM, Klein MB. Selection by genetic transformation of a *Saccharomyces cerevisiae* mutant defective for the nuclear uracil-DNA-glycosylase. J Bacteriol. 1986;166(3):905-913.

205. Crosby B, Prakash L, Davis H, Hinkle DC. Purification and characterization of a uracil-DNA glycosylase from the yeast, *Saccharomyces cerevisiae*. Nucleic Acids Res. 1981;9(21):5797-5809. doi:10.1093/nar/9.21.5797.

206. An Q, Robins P, Lindahl T, Barnes DE. C --> T mutagenesis and γ-radiation sensitivity due to deficiency in the Smug1 and Ung DNA glycosylases. EMBO J. 2005;24(12):2205-2213. doi:10.1038/sj.emboj.7600689.

207. Dizdaroglu M, Karakaya A, Jaruga P, Slupphaug G, Krokan HE. Novel activities of human uracil DNA N-glycosylase for cytosine-derived products of oxidative DNA damage. Nucleic Acids Res. 1996;24(3):418-422. doi:10.1093/nar/24.3.418.

208. Haug T, Skorpen F, Aas PA, Malm V, Skjelbred C, Krokan HE. Regulation of expression of nuclear and mitochondrial forms of human uracil-DNA glycosylase. Nucleic Acids Res. 1998;26(6):1449-1457. doi:10.1093/nar/26.6.1449.

209. Nilsen H, Otterlei M, Haug T, Solum K, Nagelhus TA, Skorpen F, Krokan HE. Nuclear and mitochondrial uracil-DNA glycosylases are generated by alternative splicing and transcription from different positions in the *UNG* gene. Nucleic Acids Res. 1997;25(4):750-755. doi:10.1093/nar/25.4.750.

210. Otterlei M, Haug T, Nagelhus TA, Slupphaug G, Lindmo T, Krokan HE. Nuclear and mitochondrial splice forms of human uracil-DNA glycosylase contain a complex nuclear localisation signal and a strong classical mitochondrial localisation signal, respectively. Nucleic Acids Res. 1998;26(20):4611-4617. doi:10.1093/nar/26.20.4611.

211. Slupphaug G, Markussen FH, Olsen LC, Aasland R, Aarsaether N, Bakke O, Krokan HE, Helland DE. Nuclear and mitochondrial forms of human uracil-DNA glycosylase are encoded by the same gene. Nucleic Acids Res. 1993;21(11):2579-2584. doi:10.1093/nar/21.11.2579.

212. Wittwer CU, Krokan H. Uracil-DNA glycosylase in HeLa S3 cells: interconvertibility of 50 and 20 kDa forms and similarity of the nuclear and mitochondrial form of the enzyme. Biochim Biophys Acta. 1985;832(3):308-318.

213. Chatterjee A, Singh KK. Uracil-DNA glycosylase–deficient yeast exhibit a mitochondrial mutator phenotype. Nucleic Acids Res. 2001;29(24):4935-4940. doi:10.1093/nar/29.24.4935.

214. Gupta PK, Sirover MA. Stimulation of the nuclear uracil DNA glycosylase in proliferating human fibroblasts. Cancer Res. 1981;41(8):3133-3136.

215. Sirover MA. Induction of the DNA repair enzyme uracil-DNA glycosylase in stimulated human lymphocytes. Cancer Res. 1979;39(6 1):2090-2095.

216. Slupphaug G, Olsen LC, Helland D, Aasland R, Krokan HE. Cell cycle regulation and in vitro hybrid arrest analysis of the major human uracil-DNA glycosylase. Nucleic Acids Res. 1991;19(19):5131-5137. doi:10.1093/nar/19.19.5131.

217. Johnston LH, Johnson AL. The DNA repair genes *RAD54* and *UNG1* are cell cycle regulated in budding yeast but MCB promoter elements have no essential role in the DNA damage response. Nucleic Acids Res. 1995;23(12):2147-2152. doi:10.1093/nar/23.12.2147.

218. Hagen L, Kavli B, Sousa MM, Torseth K, Liabakk NB, Sundheim O, Pena-Diaz J, Otterlei M, Horning O, Jensen ON, Krokan HE, Slupphaug G. Cell cycle-specific UNG2 phosphorylations regulate protein turnover, activity and association with RPA. EMBO J. 2008;27(1):51-61. doi:10.1038/sj.emboj.7601958.

219. Muller-Weeks S, Mastran B, Caradonna S. The nuclear isoform of the highly conserved human uracil-DNA glycosylase is an M_r 36,000 phosphoprotein. J Biol Chem. 1998;273(34):21909-21917.

220. Lu X, Nguyen TA, Appella E, Donehower LA. Homeostatic regulation of base excision repair by a p53-induced phosphatase: linking stress response pathways with DNA repair proteins. Cell Cycle. 2004;3(11):1363-1366. doi:10.4161/cc.3.11.1241.

221. Yamaguchi H, Minopoli G, Demidov ON, Chatterjee DK, Anderson CW, Durell SR, Appella E. Substrate specificity of the human protein phosphatase 2Cδ, Wip1. Biochemistry. 2005;44(14):5285-5294. doi:10.1021/bi0476634.

222. Akbari M, Otterlei M, Pena-Diaz J, Aas PA, Kavli B, Liabakk NB, Hagen L, Imai K, Durandy A, Slupphaug G, Krokan HE. Repair of U/G and U/A in DNA by UNG2-associated repair complexes takes place predominantly by short-patch repair both in proliferating and growth-arrested cells. Nucleic Acids Res. 2004;32(18):5486-5498. doi:10.1093/nar/gkh872.

223. Ko R, Bennett SE. Physical and functional interaction of human nuclear uracil-DNA glycosylase with proliferating cell nuclear antigen. DNA Repair. 2005;4(12):1421-1431. doi:10.1016/j.dnarep.2005.08.006.

224. Parlanti E, Locatelli G, Maga G, Dogliotti E. Human base excision repair complex is physically associated to DNA replication and cell cycle regulatory proteins. Nucleic Acids Res. 2007;35(5):1569-1577. doi:10.1093/nar/gkl1159.

225. Akbari M, Otterlei M, Pena-Diaz J, Krokan HE. Different organization of base excision repair of uracil in DNA in nuclei and mitochondria and selective upregulation of mitochondrial uracil-DNA glycosylase after oxidative stress. Neuroscience. 2007;145(4):1201-1212. doi:10.1016/j.neuroscience.2006.10.010.

226. Hegre SA, Saetrom P, Aas PA, Pettersen HS, Otterlei M, Krokan HE. Multiple microRNAs may regulate the DNA repair enzyme uracil-DNA glycosylase. DNA Repair. 2013;12(1):80-86. doi:10.1016/j.dnarep.2012.10.007.

227. Nilsen H, Haushalter KA, Robins P, Barnes DE, Verdine GL, Lindahl T. Excision of deaminated cytosine from the vertebrate genome: role of the SMUG1 uracil-DNA glycosylase. EMBO J. 2001;20(15):4278-4286. doi:10.1093/emboj/20.15.4278.

228. Haushalter KA, Todd Stukenberg MW, Kirschner MW, Verdine GL. Identification of a new uracil-DNA glycosylase family by expression cloning using synthetic inhibitors. Curr Biol. 1999;9(4):174-185.

229. Masaoka A, Matsubara M, Hasegawa R, Tanaka T, Kurisu S, Terato H, Ohyama Y, Karino N, Matsuda A, Ide H. Mammalian 5-formyluracil-DNA glycosylase. 2. Role of SMUG1 uracil-DNA glycosylase in repair of 5-formyluracil and other oxidized and deaminated base lesions. Biochemistry. 2003;42(17):5003-5012. doi:10.1021/bi0273213.

230. Matsubara M, Masaoka A, Tanaka T, Miyano T, Kato N, Terato H, Ohyama Y, Iwai S, Ide H. Mammalian 5-formyluracil-DNA glycosylase. 1. Identification and characterization of a novel activity that releases 5-formyluracil from DNA. Biochemistry. 2003;42(17):4993-5002. doi:10.1021/bi027322v.

231. Matsubara M, Masaoka A, Tanaka T, Terato H, Ohyama Y, Ide H. Identification and characterization of mammalian 5-formyluracil-DNA glycosylase. Nucleic Acids Res Suppl. 2003;(3):233-234. doi:10.1093/nass/3.1.233.

232. Boorstein RJ, Cummings A, Jr., Marenstein DR, Chan MK, Ma Y, Neubert TA, Brown SM, Teebor GW. Definitive identification of mammalian 5-hydroxymethyluracil DNA *N*-glycosylase activity as SMUG1. J Biol Chem. 2001;276(45):41991-41997. doi:10.1074/jbc.M106953200.

233. Krokan HE, Drablos F, Slupphaug G. Uracil in DNA – occurrence, consequences and repair. Oncogene. 2002;21(58):8935-8948. doi:10.1038/sj.onc.1205996.

234. Mi R, Dong L, Kaulgud T, Hackett KW, Dominy BN, Cao W. Insights from xanthine and uracil DNA glycosylase activities of bacterial and human SMUG1: switching SMUG1 to UDG. J Mol Biol. 2009;385(3):761-778. doi:10.1016/j.jmb.2008.09.038.

235. Pettersen HS, Sundheim O, Gilljam KM, Slupphaug G, Krokan HE, Kavli B. Uracil-DNA glycosylases SMUG1 and UNG2 coordinate the initial steps of base excision repair by distinct mechanisms. Nucleic Acids Res. 2007;35(12):3879-3892. doi:10.1093/nar/gkm372.

236. Jobert L, Skjeldam HK, Dalhus B, Galashevskaya A, Vagbo CB, Bjoras M, Nilsen H. The human base excision repair enzyme SMUG1 directly interacts with DKC1 and contributes to RNA quality control. Mol Cell. 2013;49(2):339-345. doi:10.1016/j.molcel.2012.11.010.

237. Gallinari P, Jiricny J. A new class of uracil-DNA glycosylases related to human thymine-DNA glycosylase. Nature. 1996;383(6602):735-738. doi:10.1038/383735a0.

238. Neddermann P, Jiricny J. The purification of a mismatch-specific thymine-DNA glycosylase from HeLa cells. J Biol Chem. 1993;268(28):21218-21224.

239. Wiebauer K, Jiricny J. Mismatch-specific thymine DNA glycosylase and DNA polymerase β mediate the correction of G.T mispairs in nuclear extracts from human cells. Proc Natl Acad Sci USA. 1990;87(15):5842-5845.

240. Hardeland U, Bentele M, Jiricny J, Schar P. The versatile thymine DNA-glycosylase: a comparative characterization of the human, *Drosophila* and fission yeast orthologs. Nucleic Acids Res. 2003;31(9):2261-2271. doi:10.1093/nar/gkg344.

241. Maiti A, Drohat AC. Thymine DNA glycosylase can rapidly excise 5-formylcytosine and 5-carboxylcytosine: potential implications for active demethylation of CpG sites. J Biol Chem. 2011;286(41):35334-35338. doi:10.1074/jbc.C111.284620.

242. Zhang L, Lu X, Lu J, Liang H, Dai Q, Xu GL, Luo C, Jiang H, He C. Thymine DNA glycosylase specifically recognizes 5-carboxylcytosine-modified DNA. Nat Chem Biol. 2012;8(4):328-330. doi:10.1038/nchembio.914.

243. Liu P, Burdzy A, Sowers LC. Repair of the mutagenic DNA oxidation product, 5-formyluracil. DNA Repair. 2003;2(2):199-210.

244. Yoon JH, Iwai S, O'Connor TR, Pfeifer GP. Human thymine DNA glycosylase (TDG) and methyl-CpG-binding protein 4 (MBD4) excise thymine glycol (Tg) from a Tg:G mispair. Nucleic Acids Res. 2003;31(18):5399-5404.

245. Borys-Brzywczy E, Arczewska KD, Saparbaev M, Hardeland U, Schar P, Kusmierek JT. Mismatch dependent uracil/thymine-DNA glycosylases excise exocyclic hydroxyethano and hydroxypropano cytosine adducts. Acta Biochim Pol. 2005;52(1):149-165. doi:055201149.

246. Hang B, Guliaev AB. Substrate specificity of human thymine-DNA glycosylase on exocyclic cytosine adducts. Chem-Biol Interact. 2007;165(3):230-238. doi:10.1016/j.cbi.2006.12.013.

247. Jurado J, Maciejewska A, Krwawicz J, Laval J, Saparbaev MK. Role of mismatch-specific uracil-DNA glycosylase in repair of 3,N^4-ethenocytosine in vivo. DNA Repair. 2004;3(12):1579-1590. doi:10.1016/j.dnarep.2004.06.012.

248. Sagi J, Perry A, Hang B, Singer B. Differential destabilization of the DNA oligonucleotide double helix by a T.G mismatch, 3,N^4-ethenocytosine, 3,N^4-ethanocytosine, or an 8-(hydroxymethyl)-3,N^4-ethenocytosine adduct incorporated into the same sequence contexts. Chem Res Toxicol. 2000;13(9):839-845. doi:10.1021/tx000040g.

249. Lari SU, Al-Khodairy F, Paterson MC. Substrate specificity and sequence preference of G:T mismatch repair: incision at G:T, O^6-methylguanine:T, and G:U mispairs in DNA by human cell extracts. Biochemistry. 2002;41(29):9248-9255. doi:10.1021/bi020239n.

250. Li YQ, Zhou PZ, Zheng XD, Walsh CP, Xu GL. Association of Dnmt3a and thymine DNA glycosylase links DNA methylation with base-excision repair. Nucleic Acids Res. 2007;35(2):390-400. doi:10.1093/nar/gkl1052.

251. Boland MJ, Christman JK. Characterization of Dnmt3b:thymine-DNA glycosylase interaction and stimulation of thymine glycosylase-mediated repair by DNA methyltransferase(s) and RNA. J Mol Biol. 2008;379(3):492-504. doi:10.1016/j.jmb.2008.02.049.

252. Chen D, Lucey MJ, Phoenix F, Lopez-Garcia J, Hart SM, Losson R, Buluwela L, Coombes RC, Chambon P, Schar P, Ali S. T:G mismatch-specific thymine-DNA glycosylase potentiates transcription of estrogen-regulated genes through direct interaction with estrogen receptor α. J Biol Chem. 2003;278(40):38586-38592. doi:10.1074/jbc.M304286200.

253. Chiang S, Burch T, Van Domselaar G, Dick K, Radziwon A, Brusnyk C, Edwards MR, Piper J, Cutts T, Cao J, Li X, He R. The interaction between thymine DNA glycosylase and nuclear receptor coactivator 3 is required for the transcriptional activation of nuclear hormone receptors. Molecular Cell Biochem. 2010;333(1-2):221-232. doi:10.1007/s11010-009-0223-1.

254. Leger H, Smet-Nocca C, Attmane-Elakeb A, Morley-Fletcher S, Benecke AG, Eilebrecht S. A TDG/CBP/RARα ternary complex mediates the retinoic acid-dependent expression of DNA methylation-sensitive genes. Genomics Proteomics Bioinformatics. 2014;12(1):8-18. doi:10.1016/j.gpb.2013.11.001.

255. Lucey MJ, Chen D, Lopez-Garcia J, Hart SM, Phoenix F, Al-Jehani R, Alao JP, White R, Kindle KB, Losson R, Chambon P, Parker MG, Schar P, Heery DM, Buluwela L, Ali S. T:G mismatch-specific thymine-DNA glycosylase (TDG) as a coregulator of transcription interacts with SRC1 family members through a novel tyrosine repeat motif. Nucleic Acids Res. 2005;33(19):6393-6404. doi:10.1093/nar/gki940.

256. Um S, Harbers M, Benecke A, Pierrat B, Losson R, Chambon P. Retinoic acid receptors interact physically and functionally with the T:G mismatch-specific thymine-DNA glycosylase. J Biol Chem. 1998;273(33):20728-20736. doi:10.1074/jbc.273.33.20728.

257. Cortellino S, Xu J, Sannai M, Moore R, Caretti E, Cigliano A, Le Coz M, Devarajan K, Wessels A, Soprano D, Abramowitz LK, Bartolomei MS, Rambow F, Bassi MR, Bruno T, Fanciulli M, Renner C, Klein-Szanto AJ, Matsumoto Y, Kobi D, Davidson I, Alberti C, Larue L, Bellacosa A. Thymine DNA glycosylase is essential for active DNA demethylation by linked deamination-base excision repair. Cell. 2011;146(1):67-79. doi:10.1016/j.cell.2011.06.020.

258. He YF, Li BZ, Li Z, Liu P, Wang Y, Tang Q, Ding J, Jia Y, Chen Z, Li L, Sun Y, Li X, Dai Q, Song CX, Zhang K, He C, Xu GL. Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA. Science. 2011;333(6047):1303-1307. doi:10.1126/science.1210944.

259. Kohli RM, Zhang Y. TET enzymes, TDG and the dynamics of DNA demethylation. Nature. 2013;502(7472):472-479. doi:10.1038/nature12750.

260. Hardeland U, Kunz C, Focke F, Szadkowski M, Schar P. Cell cycle regulation as a mechanism for functional separation of the apparently redundant uracil DNA glycosylases TDG and UNG2. Nucleic Acids Res. 2007;35(11):3859-3867. doi:10.1093/nar/gkm337.

261. Madabushi A, Hwang BJ, Jin J, Lu AL. Histone deacetylase SIRT1 modulates and deacetylates DNA base excision repair enzyme thymine DNA glycosylase. Biochem J. 2013;456(1):89-98. doi:10.1042/BJ20130670.

262. Morita S, Horii T, Kimura M, Ochiya T, Tajima S, Hatada I. miR-29 represses the activities of DNA methyltransferases and DNA demethylases. Int J Mol Sci. 2013;14(7):14647-14658. doi:10.3390/ijms140714647.

263. Zhang P, Huang B, Xu X, Sessa WC. Ten-eleven translocation (Tet) and thymine DNA glycosylase (TDG), components of the demethylation pathway, are direct targets of miRNA-29a. Biochem Biophys Res Commun. 2013;437(3):368-373. doi:10.1016/j.bbrc.2013.06.082.

264. Mohan RD, Litchfield DW, Torchia J, Tini M. Opposing regulatory roles of phosphorylation and acetylation in DNA mispair processing by thymine DNA glycosylase. Nucleic Acids Res. 2010;38(4):1135-1148. doi:10.1093/nar/gkp1097.

265. Tini M, Benecke A, Um SJ, Torchia J, Evans RM, Chambon P. Association of CBP/p300 acetylase and thymine DNA glycosylase links DNA repair and transcription. Mol Cell. 2002;9(2):265-277. doi:10.1016/S1097-2765(02)00453-7.

266. 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. doi:10.1093/emboj/21.6.1456.

267. 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. doi:10.1186/1471-2091-12-4.

268. Owen RM, Baker RD, Bader S, Dunlop MG, Nicholl ID. The identification of a novel alternatively spliced form of the MBD4 DNA glycosylase. Oncol Rep. 2007;17(1):111-116. doi:10.3892/or.17.1.111.

269. Kondo E, Gu Z, Horii A, Fukushige S. The thymine DNA glycosylase MBD4 represses transcription and is associated with methylated *p16(INK4a)* and *hMLH1* genes. Mol Cell Biol. 2005;25(11):4388-4396. doi:10.1128/MCB.25.11.4388-4396.2005.

270. Bellacosa A, Cicchillitti L, Schepis F, Riccio A, Yeung AT, Matsumoto Y, Golemis EA, Genuardi M, Neri G. MED1, a novel human methyl-CpG-binding endonuclease, interacts with DNA mismatch repair protein MLH1. Proc Natl Acad Sci USA. 1999;96(7):3969-3974. doi:10.1073/pnas.96.7.3969.

271. Wu P, Qiu C, Sohail A, Zhang X, Bhagwat AS, Cheng X. Mismatch repair in methylated DNA. Structure and activity of the mismatch-specific thymine glycosylase domain of methyl-CpG-binding protein MBD4. J Biol Chem. 2003;278(7):5285-5291. doi:10.1074/jbc.M210884200.

272. Sjolund AB, Senejani AG, Sweasy JB. MBD4 and TDG: multifaceted DNA glycosylases with ever expanding biological roles. Mutat Res. 2013;743-744:12-25. doi:10.1016/j.mrfmmm.2012.11.001.

273. Laget S, Miotto B, Chin HG, Esteve PO, Roberts RJ, Pradhan S, Defossez PA. MBD4 cooperates with DNMT1 to mediate methyl-DNA repression and protects mammalian cells from oxidative stress. Epigenetics. 2014;9(4):546-556. doi:10.4161/epi.27695.

274. McGoldrick JP, Yeh YC, Solomon M, Essigmann JM, Lu AL. Characterization of a mammalian homolog of the *Escherichia coli* MutY mismatch repair protein. Mol Cell Biol. 1995;15(2):989-996.

275. Ohtsubo T, Nishioka K, Imaiso Y, Iwai S, Shimokawa H, Oda H, Fujiwara T, Nakabeppu Y. Identification of human MutY homolog (hMYH) as a repair enzyme for 2-hydroxyadenine in DNA and detection of multiple forms of hMYH located in nuclei and mitochondria. Nucleic Acids Res. 2000;28(6):1355-1364. doi:10.1093/nar/28.6.1355.

276. Parker A, Gu Y, Lu AL. Purification and characterization of a mammalian homolog of *Escherichia coli* MutY mismatch repair protein from calf liver mitochondria. Nucleic Acids Res. 2000;28(17):3206-3215. doi:10.1093/nar/28.17.3206.

277. Ichinoe A, Behmanesh M, Tominaga Y, Ushijima Y, Hirano S, Sakai Y, Tsuchimoto D, Sakumi K, Wake N, Nakabeppu Y. Identification and characterization of two forms of mouse MUTYH proteins encoded by alternatively spliced transcripts. Nucleic Acids Res. 2004;32(2):477-487. doi:10.1093/nar/gkh214.

278. Boldogh I, Milligan D, Lee MS, Bassett H, Lloyd RS, McCullough AK. hMYH cell cycle-dependent expression, subcellular localization and association with replication foci: evidence suggesting replication-coupled repair of adenine:8-oxoguanine mispairs. Nucleic Acids Res. 2001;29(13):2802-2809. doi:10.1093/nar/29.13.2802.

279. Hayashi H, Tominaga Y, Hirano S, McKenna AE, Nakabeppu Y, Matsumoto Y. Replication-associated repair of adenine:8-oxoguanine mispairs by MYH. Curr Biol. 2002;12(4):335-339. doi:10.1016/S0960-9822(02)00686-3.

280. Gu Y, Parker A, Wilson TM, Bai H, Chang DY, Lu AL. Human MutY homolog, a DNA glycosylase involved in base excision repair, physically and functionally interacts with mismatch repair proteins human MutS

homolog 2/human MutS homolog 6. J Biol Chem. 2002;277(13):11135-11142. doi:10.1074/jbc.M108618200.

281. Yang H, Clendenin WM, Wong D, Demple B, Slupska MM, Chiang JH, Miller JH. Enhanced activity of adenine-DNA glycosylase (Myh) by apurinic/apyrimidinic endonuclease (Ape1) in mammalian base excision repair of an A/GO mismatch. Nucleic Acids Res. 2001;29(3):743-752. doi:10.1093/nar/29.3.743.

282. Tominaga Y, Ushijima Y, Tsuchimoto D, Mishima M, Shirakawa M, Hirano S, Sakumi K, Nakabeppu Y. MUTYH prevents OGG1 or APEX1 from inappropriately processing its substrate or reaction product with its C-terminal domain. Nucleic Acids Res. 2004;32(10):3198-3211. doi:10.1093/nar/gkh642.

283. Izumi T, Hazra TK, Boldogh I, Tomkinson AE, Park MS, Ikeda S, Mitra S. Requirement for human AP endonuclease 1 for repair of 3′-blocking damage at DNA single-strand breaks induced by reactive oxygen species. Carcinogenesis. 2000;21(7):1329-1334. doi:10.1093/carcin/21.7.1329.

284. Seki S, Hatsushika M, Watanabe S, Akiyama K, Nagao K, Tsutsui K. cDNA cloning, sequencing, expression and possible domain structure of human APEX nuclease homologous to *Escherichia coli* exonuclease III. Biochim Biophys Acta. 1992;1131(3):287-299.

285. Seki S, Ikeda S, Watanabe S, Hatsushika M, Tsutsui K, Akiyama K, Zhang B. A mouse DNA repair enzyme (APEX nuclease) having exonuclease and apurinic/apyrimidinic endonuclease activities: purification and characterization. Biochim Biophys Acta. 1991;1079(1):57-64.

286. Hadi MZ, Wilson DM, 3rd. Second human protein with homology to the *Escherichia coli* abasic endonuclease exonuclease III. Environ Mol Mutag. 2000;36(4):312-324. doi:10.1002/1098-2280(2000)36:4<312::AID-EM7>3.0.CO;2-K.

287. Johnson AW, Demple B. Yeast DNA diesterase for 3′-fragments of deoxyribose: purification and physical properties of a repair enzyme for oxidative DNA damage. J Biol Chem. 1988;263(34):18009-18016.

288. Johnson AW, Demple B. Yeast DNA 3′-repair diesterase is the major cellular apurinic/apyrimidinic endonuclease: substrate specificity and kinetics. J Biol Chem. 1988;263(34):18017-18022.

289. Bennett RA. The *Saccharomyces cerevisiae ETH1* gene, an inducible homolog of exonuclease III that provides resistance to DNA-damaging agents and limits spontaneous mutagenesis. Mol Cell Biol. 1999;19(3):1800-1809.

290. Unk I, Haracska L, Johnson RE, Prakash S, Prakash L. Apurinic endonuclease activity of yeast Apn2 protein. J Biol Chem. 2000;275(29):22427-22434. doi:10.1074/jbc.M002845200.

291. Unk I, Haracska L, Prakash S, Prakash L. 3'-phosphodiesterase and 3'-->5' exonuclease activities of yeast Apn2 protein and requirement of these activities for repair of oxidative DNA damage. Mol Cell Biol. 2001;21(5):1656-1661. doi:10.1128/MCB.21.5.1656-1661.2001.

292. Winters TA, Henner WD, Russell PS, McCullough A, Jorgensen TJ. Removal of 3′-phosphoglycolate from DNA strand-break damage in an oligonucleotide substrate by recombinant human apurinic/apyrimidinic endonuclease 1. Nucleic Acids Res. 1994;22(10):1866-1873. doi:10.1093/nar/22.10.1866.

293. Parsons JL, Dianova, II, Dianov GL. APE1-dependent repair of DNA single-strand breaks containing 3′-end 8-oxoguanine. Nucleic Acids Res. 2005;33(7):2204-2209. doi:10.1093/nar/gki518.

294. Burkovics P, Szukacsov V, Unk I, Haracska L. Human Ape2 protein has a 3'-5' exonuclease activity that acts preferentially on mismatched base pairs. Nucleic Acids Res. 2006;34(9):2508-2515. doi:10.1093/nar/gkl259.

295. Ishchenko AA, Yang X, Ramotar D, Saparbaev M. The 3′->5′ exonuclease of Apn1 provides an alternative pathway to repair 7,8-dihydro-8-oxodeoxyguanosine in *Saccharomyces cerevisiae*. Mol Cell Biol. 2005;25(15):6380-6390. doi:10.1128/MCB.25.15.6380-6390.2005.

296. Mazouzi A, Vigouroux A, Aikeshev B, Brooks PJ, Saparbaev MK, Morera S, Ishchenko AA. Insight into mechanisms of 3′-5′ exonuclease activity and removal of bulky 8,5′-cyclopurine adducts by

apurinic/apyrimidinic endonucleases. Proc Natl Acad Sci USA. 2013;110(33):E3071-3080. doi:10.1073/pnas.1305281110.

297. Barnes T, Kim WC, Mantha AK, Kim SE, Izumi T, Mitra S, Lee CH. Identification of Apurinic/apyrimidinic endonuclease 1 (APE1) as the endoribonuclease that cleaves c-myc mRNA. Nucleic Acids Res. 2009;37(12):3946-3958. doi:10.1093/nar/gkp275.

298. Kim SE, Gorrell A, Rader SD, Lee CH. Endoribonuclease activity of human apurinic/apyrimidinic endonuclease 1 revealed by a real-time fluorometric assay. Anal Biochem. 2010;398(1):69-75. doi:10.1016/j.ab.2009.11.024.

299. Vascotto C, Fantini D, Romanello M, Cesaratto L, Deganuto M, Leonardi A, Radicella JP, Kelley MR, D'Ambrosio C, Scaloni A, Quadrifoglio F, Tell G. APE1/Ref-1 interacts with NPM1 within nucleoli and plays a role in the rRNA quality control process. Mol Cell Biol. 2009;29(7):1834-1854. doi:10.1128/mcb.01337-08.

300. Tell G, Quadrifoglio F, Tiribelli C, Kelley MR. The many functions of APE1/Ref-1: not only a DNA repair enzyme. Antioxidants & Redox Signaling. 2009;11(3):601-620. doi:10.1089/ars.2008.2194.

301. Johnson RE, Torres-Ramos CA, Izumi T, Mitra S, Prakash S, Prakash L. Identification of *APN2*, the *Saccharomyces cerevisiae* homolog of the major human AP endonuclease *HAP1*, and its role in the repair of abasic sites. Genes Dev. 1998;12(19):3137-3143. doi:10.1101/gad.12.19.3137.

302. Popoff SC, Spira AI, Johnson AW, Demple B. Yeast structural gene (*APN1*) for the major apurinic endonuclease: homology to *Escherichia coli* endonuclease IV. Proc Natl Acad Sci USA. 1990;87(11):4193-4197.

303. Fung H, Bennett RA, Demple B. Key role of a downstream specificity protein 1 site in cell cycle-regulated transcription of the AP endonuclease gene *APE1/APEX* in NIH3T3 cells. J Biol Chem. 2001;276(45):42011-42017. doi:10.1074/jbc.M106423200.

304. Ide Y, Tsuchimoto D, Tominaga Y, Iwamoto Y, Nakabeppu Y. Characterization of the genomic structure and expression of the mouse *Apex2* gene. Genomics. 2003;81(1):47-57. doi:10.1016/S0888-7543(02)00009-5.

305. Zaky A, Busso C, Izumi T, Chattopadhyay R, Bassiouny A, Mitra S, Bhakat KK. Regulation of the human AP-endonuclease (*APE1/Ref-1*) expression by the tumor suppressor p53 in response to DNA damage. Nucleic Acids Res. 2008;36(5):1555-1566. doi:10.1093/nar/gkm1173.

306. Yamamori T, DeRicco J, Naqvi A, Hoffman TA, Mattagajasingh I, Kasuno K, Jung SB, Kim CS, Irani K. SIRT1 deacetylates APE1 and regulates cellular base excision repair. Nucleic Acids Res. 2010;38(3):832-845. doi:10.1093/nar/gkp1039.

307. Yacoub A, Kelley MR, Deutsch WA. The DNA repair activity of human redox/repair protein APE/Ref-1 is inactivated by phosphorylation. Cancer Res. 1997;57(24):5457-5459.

308. Kenny MK, Mendez F, Sandigursky M, Kureekattil RP, Goldman JD, Franklin WA, Bases R. Heat shock protein 70 binds to human apurinic/apyrimidinic endonuclease and stimulates endonuclease activity at abasic sites. J Biol Chem. 2001;276(12):9532-9536. doi:10.1074/jbc.M009297200.

309. Mendez F, Sandigursky M, Kureekattil RP, Kenny MK, Franklin WA, Bases R. Specific stimulation of human apurinic/apyrimidinic endonuclease by heat shock protein 70. DNA Repair. 2003;2(3):259-271. doi:10.1016/S1568-7864(02)00215-X.

310. Gembka A, Toueille M, Smirnova E, Poltz R, Ferrari E, Villani G, Hubscher U. The checkpoint clamp, Rad9-Rad1-Hus1 complex, preferentially stimulates the activity of apurinic/apyrimidinic endonuclease 1 and DNA polymerase β in long patch base excision repair. Nucleic Acids Res. 2007;35(8):2596-2608. doi:10.1093/nar/gkl1139.

311. Vidal AE, Boiteux S, Hickson ID, Radicella JP. XRCC1 coordinates the initial and late stages of DNA abasic site repair through protein-protein interactions. EMBO J. 2001;20(22):6530-6539. doi:10.1093/emboj/20.22.6530.

312. Chattopadhyay R, Wiederhold L, Szczesny B, Boldogh I, Hazra TK, Izumi T, Mitra S. Identification and characterization of mitochondrial abasic (AP)-endonuclease in mammalian cells. Nucleic Acids Res. 2006;34(7):2067-2076. doi:10.1093/nar/gkl177.

313. Jackson EB, Theriot CA, Chattopadhyay R, Mitra S, Izumi T. Analysis of nuclear transport signals in the human apurinic/apyrimidinic endonuclease (APE1/Ref1). Nucleic Acids Res. 2005;33(10):3303-3312. doi:10.1093/nar/gki641.

314. Kakolyris S, Kaklamanis L, Giatromanolaki A, Koukourakis M, Hickson ID, Barzilay G, Turley H, Leek RD, Kanavaros P, Georgoulias V, Gatter KC, Harris AL. Expression and subcellular localization of human AP endonuclease 1 (HAP1/Ref-1) protein: a basis for its role in human disease. Histopathology. 1998;33(6):561-569. doi:10.1046/j.1365-2559.1998.00541.x.

315. Tsuchimoto D, Sakai Y, Sakumi K, Nishioka K, Sasaki M, Fujiwara T, Nakabeppu Y. Human APE2 protein is mostly localized in the nuclei and to some extent in the mitochondria, while nuclear APE2 is partly associated with proliferating cell nuclear antigen. Nucleic Acids Res. 2001;29(11):2349-2360. doi:10.1093/nar/29.11.2349.

316. Ramotar D, Kim C, Lillis R, Demple B. Intracellular localization of the Apn1 DNA repair enzyme of *Saccharomyces cerevisiae*. Nuclear transport signals and biological role. J Biol Chem. 1993;268(27):20533-20539.

317. Vongsamphanh R, Fortier PK, Ramotar D. Pir1p mediates translocation of the yeast Apn1p endonuclease into the mitochondria to maintain genomic stability. Mol Cell Biol. 2001;21(5):1647-1655. doi:10.1128/MCB.21.5.1647-1655.2001.

318. Parsons JL, Dianova, II, Dianov GL. APE1 is the major 3′-phosphoglycolate activity in human cell extracts. Nucleic Acids Res. 2004;32(12):3531-3536. doi:10.1093/nar/gkh676.

319. Moe A, Ringvoll J, Nordstrand LM, Eide L, Bjoras M, Seeberg E, Rognes T, Klungland A. Incision at hypoxanthine residues in DNA by a mammalian homologue of the *Escherichia coli* antimutator enzyme endonuclease V. Nucleic Acids Res. 2003;31(14):3893-3900. doi:10.1093/nar/gkg472.

320. Cao W. Endonuclease V: an unusual enzyme for repair of DNA deamination. Cell Mol Life Sci. 2013;70(17):3145-3156. doi:10.1007/s00018-012-1222-z.

321. 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. doi:10.1074/jbc.M400468200.

322. D'Errico M, Parlanti E, Teson M, de Jesus BM, Degan P, Calcagnile A, Jaruga P, Bjoras M, Crescenzi M, Pedrini AM, Egly JM, Zambruno G, Stefanini M, Dizdaroglu M, Dogliotti E. New functions of XPC in the protection of human skin cells from oxidative damage. EMBO J. 2006;25(18):4305-4315. doi:10.1038/sj.emboj.7601277.

323. Pascucci B, D'Errico M, Parlanti E, Giovannini S, Dogliotti E. Role of nucleotide excision repair proteins in oxidative DNA damage repair: an updating. Biochemistry (Mosc). 2011;76(1):4-15. doi:10.1134/S0006297911010032.

324. Melis JP, van Steeg H, Luijten M. Oxidative DNA damage and nucleotide excision repair. Antioxid Redox Signal. 2013;18(18):2409-2419. doi:10.1089/ars.2012.5036.

325. Plosky B, Samson L, Engelward BP, Gold B, Schlaen B, Millas T, Magnotti M, Schor J, Scicchitano DA. Base excision repair and nucleotide excision repair contribute to the removal of *N*-methylpurines from active genes. DNA Repair. 2002;1(8):683-696. doi:10.1016/S1568-7864(02)00075-7.

326. Torres-Ramos CA, Johnson RE, Prakash L, Prakash S. Evidence for the involvement of nucleotide excision repair in the removal of abasic sites in yeast. Mol Cell Biol. 2000;20(10):3522-3528. doi:10.1128/MCB.20.10.3522-3528.2000.

327. Alexeyev M, Shokolenko I, Wilson G, LeDoux S. The maintenance of mitochondrial DNA integrity—critical analysis and update. Cold Spring Harb Perspect Biol. 2013;5(5). doi:10.1101/cshperspect.a012641.

328. Shuck SC, Short EA, Turchi JJ. Eukaryotic nucleotide excision repair: from understanding mechanisms to influencing biology. Cell Res. 2008;18(1):64-72. doi:10.1038/cr.2008.2.

329. Araujo SJ, Tirode F, Coin F, Pospiech H, Syvaoja JE, Stucki M, Hubscher U, Egly JM, Wood RD. Nucleotide excision repair of DNA with recombinant human proteins: definition of the minimal set of factors, active forms of TFIIH, and modulation by CAK. Genes Dev. 2000;14(3):349-359. doi:10.1101/gad.14.3.349.

330. Evans E, Moggs JG, Hwang JR, Egly JM, Wood RD. Mechanism of open complex and dual incision formation by human nucleotide excision repair factors. EMBO J. 1997;16(21):6559-6573. doi:10.1093/emboj/16.21.6559.

331. Shivji MK, Podust VN, Hubscher U, Wood RD. Nucleotide excision repair DNA synthesis by DNA polymerase ε in the presence of PCNA, RFC, and RPA. Biochemistry. 1995;34(15):5011-5017.

332. Moser J, Kool H, Giakzidis I, Caldecott K, Mullenders LH, Fousteri MI. Sealing of chromosomal DNA nicks during nucleotide excision repair requires XRCC1 and DNA ligase IIIα in a cell-cycle-specific manner. Mol Cell. 2007;27(2):311-323. doi:10.1016/j.molcel.2007.06.014.

333. Fousteri M, Mullenders LHF. Transcription-coupled nucleotide excision repair in mammalian cells: molecular mechanisms and biological effects. Cell Res. 2008;18(1):73-84. doi:10.1038/cr.2008.6.

334. Johnson RE, Kovvali GK, Prakash L, Prakash S. Requirement of the yeast *MSH3* and *MSH6 g*enes for *MSH2*-dependent genomic stability. J Biol Chem. 1996;271(13):7285-7288. doi:10.1074/jbc.271.13.7285.

335. Marsischky GT, Filosi N, Kane MF, Kolodner R. Redundancy of *Saccharomyces cerevisiae* MSH3 and MSH6 in MSH2-dependent mismatch repair. Genes Dev. 1996;10(4):407-420. doi:10.1101/gad.10.4.407.

336. Kadyrov FA, Holmes SF, Arana ME, Lukianova OA, O'Donnell M, Kunkel TA, Modrich P. *Saccharomyces cerevisiae* MutLα is a mismatch repair endonuclease. J Biol Chem. 2007;282(51):37181-37190. doi:10.1074/jbc.M707617200.

337. Pluciennik A, Dzantiev L, Iyer RR, Constantin N, Kadyrov FA, Modrich P. PCNA function in the activation and strand direction of MutLα endonuclease in mismatch repair. Proc Natl Acad Sci USA. 2010;107(37):16066-16071. doi:10.1073/pnas.1010662107.

338. Collura A, Kemp PA, Boiteux S. Abasic sites linked to dUTP incorporation in DNA are a major cause of spontaneous mutations in absence of base excision repair and Rad17-Mec3-Ddc1 (9-1-1) DNA damage checkpoint clamp in *Saccharomyces cerevisiae*. DNA Repair. 2012;11(3):294-303. doi:10.1016/j.dnarep.2011.12.004.

339. Tishkoff DX, Boerger AL, Bertrand P, Filosi N, Gaida GM, Kane MF, Kolodner RD. Identification and characterization of Saccharomyces cerevisiae *EXO1*, a gene encoding an exonuclease that interacts with *MSH2*. Proc Natl Acad Sci USA. 1997;94(14):7487-7492.

340. Boiteux S, Jinks-Robertson S. DNA repair mechanisms and the bypass of DNA damage in *Saccharomyces cerevisiae*. Genetics. 2013;193(4):1025-1064. doi:10.1534/genetics.112.145219.

341. Galperin MY, Moroz OV, Wilson KS, Murzin AG. House cleaning, a part of good housekeeping. Mol Microbiol. 2006;59(1):5-19. doi:10.1111/j.1365-2958.2005.04950.x.

342. Masutani C, Kusumoto R, Yamada A, Dohmae N, Yokoi M, Yuasa M, Araki M, Iwai S, Takio K, Hanaoka F. The XPV (xeroderma pigmentosum variant) gene encodes human DNA polymerase η. Nature. 1999;399(6737):700-704. doi:10.1038/21447.

343. Johnson RE, Kondratick CM, Prakash S, Prakash L. hRAD30 mutations in the variant form of xeroderma pigmentosum. Science. 1999;285(5425):263-265. doi:10.1126/science.285.5425.263.

344. Morrison A, Christensen RB, Alley J, Beck AK, Bernstine EG, Lemontt JF, Lawrence CW. *REV3*, a *Saccharomyces cerevisiae* gene whose function is required for induced mutagenesis, is predicted to encode a nonessential DNA polymerase. J Bacteriol. 1989;171(10):5659-5667.

345. Degtyareva NP, Heyburn L, Sterling J, Resnick MA, Gordenin DA, Doetsch PW. Oxidative stress-induced mutagenesis in single-strand DNA occurs primarily at cytosines and is DNA polymerase ζ-dependent only for adenines and guanines. Nucleic Acids Res. 2013. doi:10.1093/nar/gkt671.

346. Goodarzi AA, Jeggo PA. The repair and signaling responses to DNA double-strand breaks. Adv Genet. 2013;82:1-45. doi:10.1016/b978-0-12-407676-1.00001-9.

347. Wood RD. Mammalian nucleotide excision repair proteins and interstrand crosslink repair. Environ Mol Mutag. 2010;51(6):520-526. doi:10.1002/em.20569.

348. Kottemann MC, Smogorzewska A. Fanconi anaemia and the repair of Watson and Crick DNA crosslinks. Nature. 2013;493(7432):356-363. doi:10.1038/nature11863.

349. Enoiu M, Jiricny J, Scharer OD. Repair of cisplatin-induced DNA interstrand crosslinks by a replication-independent pathway involving transcription-coupled repair and translesion synthesis. Nucleic Acids Res. 2012;40(18):8953-8964. doi:10.1093/nar/gks670.

350. Wang X, Peterson CA, Zheng H, Nairn RS, Legerski RJ, Li L. Involvement of nucleotide excision repair in a recombination-independent and error-prone pathway of DNA interstrand cross-link repair. Mol Cell Biol. 2001;21(3):713-720. doi:10.1128/mcb.21.3.713-720.2001.

351. Kracikova M, Akiri G, George A, Sachidanandam R, Aaronson SA. A threshold mechanism mediates p53 cell fate decision between growth arrest and apoptosis. Cell Death Differerntiation. 2013;20(4):576-588. doi:10.1038/cdd.2012.155.

352. Nam EA, Cortez D. ATR signalling: more than meeting at the fork. Biochem J. 2011;436(3):527-536. doi:10.1042/bj20102162.

353. Giannattasio M, Follonier C, Tourrière H, Puddu F, Lazzaro F, Pasero P, Lopes M, Plevani P, Muzi-Falconi M. Exo1 competes with repair synthesis, converts NER intermediates to long ssDNA Gaps, and promotes checkpoint qctivation. Mol Cell. 2010;40(1):50-62. doi:10.1016/j.molcel.2010.09.004.

354. Novarina D, Amara F, Lazzaro F, Plevani P, Muzi-Falconi M. Mind the gap: Keeping UV lesions in check. DNA Repair. 2011;10(7):751-759. doi:10.1016/j.dnarep.2011.04.030.

355. Sertic S, Pizzi S, Cloney R, Lehmann AR, Marini F, Plevani P, Muzi-Falconi M. Human exonuclease 1 connects nucleotide excision repair (NER) processing with checkpoint activation in response to UV irradiation. Proc Natl Acad Sci USA. 2011;108(33):13647-13652. doi:10.1073/pnas.1108547108.

356. Lindsey-Boltz LA, Kemp MG, Reardon JT, DeRocco V, Iyer RR, Modrich P, Sancar A. Coupling of human DNA excision repair and the DNA damage checkpoint in a defined in vitro system. J Biol Chem. 2014;289(8):5074-5082. doi:10.1074/jbc.M113.542787.

357. Woodhouse BC, Dianov GL. Poly ADP-ribose polymerase-1: an international molecule of mystery. DNA Repair. 2008;7(7):1077-1086. doi:10.1016/j.dnarep.2008.03.009.

358. Rowe LA, Degtyareva N, Doetsch PW. Yap1: a DNA damage responder in *Saccharomyces cerevisiae*. Mech Age Dev. 2012;133(4):147-156. doi:10.1016/j.mad.2012.03.009.

359. Oehler S, Muller-Hill B. High local concentration: a fundamental strategy of life. J Mol Biol. 2010;395(2):242-253. doi:10.1016/j.jmb.2009.10.056.

360. Hernandez-Verdun D. Assembly and disassembly of the nucleolus during the cell cycle. Nucleus. 2011;2(3):189-194. doi:10.4161/nucl.2.3.16246.

361. Mahen R, Venkitaraman AR. Pattern formation in centrosome assembly. Curr Opin Cell Biol. 2012;24(1):14-23. doi:10.1016/j.ceb.2011.12.012.

362. Constantin B. Dystrophin complex functions as a scaffold for signalling proteins. Biochim Biophys Acta. 2014;1838(2):635-642. doi:10.1016/j.bbamem.2013.08.023.

363. Nie J, Liu L, He F, Fu X, Han W, Zhang L. CKIP-1: a scaffold protein and potential therapeutic target integrating multiple signaling pathways and physiological functions. Ageing Res Rev. 2013;12(1):276-281. doi:10.1016/j.arr.2012.07.002.

364. Stano P, D'Aguanno E, Bolz J, Fahr A, Luisi PL. A remarkable self-organization process as the origin of primitive functional cells. Angew Chem. 2013;52(50):13397-13400. doi:10.1002/anie.201306613.

365. Wilson KL, Dawson SC. Functional evolution of nuclear structure. J Cell Biol. 2011;195(2):171-181. doi:10.1083/jcb.201103171.

366. Gilbert W. Why genes in pieces? Nature. 1978;271(5645):501. doi:10.1038/271501a0.

367. Ptak C, Aitchison JD, Wozniak RW. The multifunctional nuclear pore complex: a platform for controlling gene expression. Curr Opin Cell Biol. 2014;28c:46-53. doi:10.1016/j.ceb.2014.02.001.

368. Strambio-De-Castillia C, Niepel M, Rout MP. The nuclear pore complex: bridging nuclear transport and gene regulation. Nat Rev Mol Cell Biol. 2010;11(7):490-501. doi:10.1038/nrm2928.

369. Zuleger N, Robson MI, Schirmer EC. The nuclear envelope as a chromatin organizer. Nucleus. 2011;2(5):339-349. doi:10.4161/nucl.2.5.17846.

370. Nagai S, Davoodi N, Gasser SM. Nuclear organization in genome stability: SUMO connections. Cell Res. 2011;21(3):474-485. doi:10.1038/cr.2011.31.

371. Bukata L, Parker SL, D'Angelo MA. Nuclear pore complexes in the maintenance of genome integrity. Curr Opin Cell Biol. 2013;25(3):378-386. doi:10.1016/j.ceb.2013.03.002.

372. Nakano H, Wang W, Hashizume C, Funasaka T, Sato H, Wong RW. Unexpected role of nucleoporins in coordination of cell cycle progression. Cell Cycle. 2011;10(3):425-433. doi:10.4161/cc.10.3.14721.

373. Chatel G, Fahrenkrog B. Dynamics and diverse functions of nuclear pore complex proteins. Nucleus. 2012;3(2):162-171. doi:10.4161/nucl.19674.

374. Savas JN, Toyama BH, Xu T, Yates JR, 3rd, Hetzer MW. Extremely long-lived nuclear pore proteins in the rat brain. Science. 2012;335(6071):942. doi:10.1126/science.1217421.

375. D'Angelo MA, Raices M, Panowski SH, Hetzer MW. Age-dependent deterioration of nuclear pore complexes causes a loss of nuclear integrity in postmitotic cells. Cell. 2009;136(2):284-295. doi:10.1016/j.cell.2008.11.037.

376. Hoelz A, Debler EW, Blobel G. The structure of the nuclear pore complex. Annu Rev Biochem. 2011;80:613-643. doi:10.1146/annurev-biochem-060109-151030.

377. Denning DP, Patel SS, Uversky V, Fink AL, Rexach M. Disorder in the nuclear pore complex: the FG repeat regions of nucleoporins are natively unfolded. Proc Natl Acad Sci USA. 2003;100(5):2450-2455. doi:10.1073/pnas.0437902100.

378. Paine PL, Moore LC, Horowitz SB. Nuclear envelope permeability. Nature. 1975;254(5496):109-114.

379. Pante N, Kann M. Nuclear pore complex is able to transport macromolecules with diameters of about 39 nm. Mol Biol Cell. 2002;13(2):425-434. doi:10.1091/mbc.01-06-0308.

380. Macara IG. Transport into and out of the nucleus. Microbiol Mol Biol Rev. 2001;65(4):570-594. doi:10.1128/mmbr.65.4.570-594.2001.

381. Rout MP, Aitchison JD, Suprapto A, Hjertaas K, Zhao Y, Chait BT. The yeast nuclear pore complex: composition, architecture, and transport mechanism. J Cell Biol. 2000;148(4):635-652. doi:10.1083/jcb.148.4.635.

382. Peters R. Translocation through the nuclear pore complex: selectivity and speed by reduction-of-dimensionality. Traffic. 2005;6(5):421-427. doi:10.1111/j.1600-0854.2005.00287.x.

383. Peters R. Translocation through the nuclear pore: Kaps pave the way. Bioessays. 2009;31(4):466-477. doi:10.1002/bies.200800159.

384. Lim RY, Fahrenkrog B, Koser J, Schwarz-Herion K, Deng J, Aebi U. Nanomechanical basis of selective gating by the nuclear pore complex. Science. 2007;318(5850):640-643. doi:10.1126/science.1145980.

385. Ribbeck K, Gorlich D. Kinetic analysis of translocation through nuclear pore complexes. EMBO J. 2001;20(6):1320-1330. doi:10.1093/emboj/20.6.1320.

386. Frey S, Gorlich D. A saturated FG-repeat hydrogel can reproduce the permeability properties of nuclear pore complexes. Cell. 2007;130(3):512-523. doi:10.1016/j.cell.2007.06.024.

387. Yamada J, Phillips JL, Patel S, Goldfien G, Calestagne-Morelli A, Huang H, Reza R, Acheson J, Krishnan VV, Newsam S, Gopinathan A, Lau EY, Colvin ME, Uversky VN, Rexach MF. A bimodal distribution of two distinct categories of intrinsically disordered structures with separate functions in FG nucleoporins. Mol Cell Proteomics. 2010;9(10):2205-2224. doi:10.1074/mcp.M000035-MCP201.

388. Yang W. Distinct, but not completely separate spatial transport routes in the nuclear pore complex. Nucleus. 2013;4(3):166-175. doi:10.4161/nucl.24874.

389. Shulga N, Goldfarb DS. Binding dynamics of structural nucleoporins govern nuclear pore complex permeability and may mediate channel gating. Mol Cell Biol. 2003;23(2):534-542. doi:10.1128/MCB.23.2.534-542.2003

390. Strawn LA, Shen T, Shulga N, Goldfarb DS, Wente SR. Minimal nuclear pore complexes define FG repeat domains essential for transport. Nat Cell Biol. 2004;6(3):197-206. doi:10.1038/ncb1097.

391. Naim B, Zbaida D, Dagan S, Kapon R, Reich Z. Cargo surface hydrophobicity is sufficient to overcome the nuclear pore complex selectivity barrier. EMBO J. 2009;28(18):2697-2705. doi:10.1038/emboj.2009.225.

392. Miao L, Schulten K. Probing a structural model of the nuclear pore complex channel through molecular dynamics. Biophys J. 2010;98(8):1658-1667. doi:10.1016/j.bpj.2009.12.4305.

393. Ma J, Goryaynov A, Sarma A, Yang W. Self-regulated viscous channel in the nuclear pore complex. Proc Natl Acad Sci USA. 2012;109(19):7326-7331. doi:10.1073/pnas.1201724109.

394. Peleg O, Lim RY. Converging on the function of intrinsically disordered nucleoporins in the nuclear pore complex. Biol Chem. 2010;391(7):719-730. doi:10.1515/bc.2010.092.

395. Yang W. 'Natively unfolded' nucleoporins in nucleocytoplasmic transport: clustered or evenly distributed? Nucleus. 2011;2(1):10-16. doi:10.4161/nucl.2.1.13818.

396. Tu LC, Fu G, Zilman A, Musser SM. Large cargo transport by nuclear pores: implications for the spatial organization of FG-nucleoporins. EMBO J. 2013;32(24):3220-3230. doi:10.1038/emboj.2013.239.

397. Cook A, Bono F, Jinek M, Conti E. Structural biology of nucleocytoplasmic transport. Annu Rev Biochem. 2007;76(1):647-671. doi:10.1146/annurev.biochem.76.052705.161529.

398. Kunzler M, Hurt E. Targeting of Ran: variation on a common theme? J Cell Sci. 2001;114(18):3233-3241.

399. Quan Y, Ji ZL, Wang X, Tartakoff AM, Tao T. Evolutionary and transcriptional analysis of karyopherin β superfamily proteins. Mol Cell Proteomics. 2008;7(7):1254-1269. doi:10.1074/mcp.M700511-MCP200.

400. Goldfarb DS, Corbett AH, Mason DA, Harreman MT, Adam SA. Importin α: a multipurpose nuclear-transport receptor. Trends Cell Biol. 2004;14(9):505-514. doi:10.1016/j.tcb.2004.07.016.

401. Kutay U, Bischoff FR, Kostka S, Kraft R, Gorlich D. Export of importin α from the nucleus is mediated by a specific nuclear transport factor. Cell. 1997;90(6):1061-1071. doi:10.1016/S0092-8674(00)80372-4.

402. Hood JK, Silver PA. Cse1p is required for export of Srp1p/importin-α from the nucleus in *Saccharomyces cerevisiae*. J Biol Chem. 1998;273(52):35142-35146. doi:10.1074/jbc.273.52.35142.

403. Lange A, Mills RE, Lange CJ, Stewart M, Devine SE, Corbett AH. Classical nuclear localization signals: definition, function, and interaction with importin α. J Biol Chem. 2007;282(8):5101-5105. doi:10.1074/jbc.R600026200.

404. Twyffels L, Gueydan C, Kruys V. Transportin-1 and transportin-2: protein nuclear import and beyond. FEBS Lett. 2014. doi:10.1016/j.febslet.2014.04.023.

405. Weinmann L, Hock J, Ivacevic T, Ohrt T, Mutze J, Schwille P, Kremmer E, Benes V, Urlaub H, Meister G. Importin 8 is a gene silencing factor that targets argonaute proteins to distinct mRNAs. Cell. 2009;136(3):496-507. doi:10.1016/j.cell.2008.12.023.

406. Wei Y, Li L, Wang D, Zhang CY, Zen K. Importin 8 regulates the transport of mature microRNAs into the cell nucleus. J Biol Chem. 2014;289(15):10270-10275. doi:10.1074/jbc.C113.541417.

407. Kimura M, Imamoto N. Biological significance of the importin-β family-dependent nucleocytoplasmic transport pathways. Traffic. 2014. doi:10.1111/tra.12174.

408. Chook YM, Suel KE. Nuclear import by karyopherin-βs: recognition and inhibition. Biochim Biophys Acta. 2011;1813(9):1593-1606. doi:10.1016/j.bbamcr.2010.10.014.

409. Fung HY, Chook YM. Atomic basis of CRM1-cargo recognition, release and inhibition. Semin Cancer Biol. 2014. doi:10.1016/j.semcancer.2014.03.002.

410. Kutay U, Lipowsky G, Izaurralde E, Bischoff FR, Schwarzmaier P, Hartmann E, Gorlich D. Identification of a tRNA-specific nuclear export receptor. Mol Cell. 1998;1(3):359-369. doi:10.1016/S1097-2765(00)80036-2.

411. Hellmuth K, Lau DM, Bischoff FR, Kunzler M, Hurt E, Simos G. Yeast Los1p has properties of an exportin-like nucleocytoplasmic transport factor for tRNA. Mol Cell Biol. 1998;18(11):6374-6386.

412. Mingot JM, Kostka S, Kraft R, Hartmann E, Görlich D. Importin 13: a novel mediator of nuclear import and export. EMBO J. 2001;20(14):3685-3694. doi:10.1093/emboj/20.14.3685.

413. Gontan C, Güttler T, Engelen E, Demmers J, Fornerod M, Grosveld FG, Tibboel D, Görlich D, Poot RA, Rottier RJ. Exportin 4 mediates a novel nuclear import pathway for Sox family transcription factors. J Cell Biol. 2009;185(1):27-34. doi:10.1083/jcb.200810106.

414. Yoshida K, Blobel G. The karyopherin Kap142p/Msn5p mediates nuclear import and nuclear export of different cargo proteins. J Cell Biol. 2001;152(4):729-740. doi:10.1083/jcb.152.4.729.

415. Lange A, McLane LM, Mills RE, Devine SE, Corbett AH. Expanding the definition of the classical bipartite nuclear localization signal. Traffic. 2010;11(3):311-323. doi:10.1111/j.1600-0854.2009.01028.x.

416. Robbins J, Dilworth SM, Laskey RA, Dingwall C. Two interdependent basic domains in nucleoplasmin nuclear targeting sequence: identification of a class of bipartite nuclear targeting sequence. Cell. 1991;64(3):615-623. doi:10.1016/0092-8674(91)90245-T.

417. Dingwall C, Laskey RA. Nuclear targeting sequences—a consensus? Trends Biochem Sci. 1991;16(12):478-481.

418. Lee BJ, Cansizoglu AE, Suel KE, Louis TH, Zhang Z, Chook YM. Rules for nuclear localization sequence recognition by karyopherin β2. Cell. 2006;126(3):543-558. doi:10.1016/j.cell.2006.05.049.

419. Chuderland D, Konson A, Seger R. Identification and characterization of a general nuclear translocation signal in signaling proteins. Mol Cell. 2008;31(6):850-861. doi:10.1016/j.molcel.2008.08.007.

420. Xu D, Farmer A, Collett G, Grishin NV, Chook YM. Sequence and structural analyses of nuclear export signals in the NESdb database. Mol Biol Cell. 2012;23(18):3677-3693. doi:10.1091/mbc.E12-01-0046.

421. Marfori M, Mynott A, Ellis JJ, Mehdi AM, Saunders NF, Curmi PM, Forwood JK, Boden M, Kobe B. Molecular basis for specificity of nuclear import and prediction of nuclear localization. Biochim Biophys Acta. 2011;1813(9):1562-1577. doi:10.1016/j.bbamcr.2010.10.013.

422. Xu D, Farmer A, Chook YM. Recognition of nuclear targeting signals by karyopherin-β proteins. Curr Opin Struct Biol. 2010;20(6):782-790. doi:10.1016/j.sbi.2010.09.008.

423. Meinema AC, Poolman B, Veenhoff LM. The transport of integral membrane proteins across the nuclear pore complex. Nucleus. 2012;3(4):322-329. doi:10.4161/nucl.20439.

424. Ribbeck K, Lipowsky G, Kent HM, Stewart M, Gorlich D. NTF2 mediates nuclear import of Ran. EMBO J. 1998;17(22):6587-6598. doi:10.1093/emboj/17.22.6587.

425. Smith A, Brownawell A, Macara IG. Nuclear import of Ran is mediated by the transport factor NTF2. Curr Biol. 1998;8(25):1403-1406. doi:10.1016/S0960-9822(98)00023-2.

426. Braun IC, Herold A, Rode M, Conti E, Izaurralde E. Overexpression of TAP/p15 heterodimers bypasses nuclear retention and stimulates nuclear mRNA export. J Biol Chem. 2001;276(23):20536-20543. doi:10.1074/jbc.M100400200.

427. Kumeta M, Yamaguchi H, Yoshimura SH, Takeyasu K. Karyopherin-independent spontaneous transport of amphiphilic proteins through the nuclear pore. J Cell Sci. 2012;125(21):4979-4984. doi:10.1242/jcs.109520.

428. Sweitzer TD, Hanover JA. Calmodulin activates nuclear protein import: a link between signal transduction and nuclear transport. Proc Natl Acad Sci USA. 1996;93(25):14574-14579.

429. Hanover JA, Love DC, Prinz WA. Calmodulin-driven nuclear entry: trigger for sex determination and terminal differentiation. J Biol Chem. 2009;284(19):12593-12597. doi:10.1074/jbc.R800076200.

430. Holaska JM, Black BE, Love DC, Hanover JA, Leszyk J, Paschal BM. Calreticulin is a receptor for nuclear export. J Cell Biol. 2001;152(1):127-140. doi:10.1083/jcb.152.1.127.

431. Holaska JM, Black BE, Rastinejad F, Paschal BM. Ca^{2+}-dependent nuclear export mediated by calreticulin. Mol Cell Biol. 2002;22(17):6286-6297.

432. Wagstaff KM, Jans DA. Importins and beyond: non-conventional nuclear transport mechanisms. Traffic. 2009;10(9):1188-1198. doi:10.1111/j.1600-0854.2009.00937.x.

433. Dyall SD, Brown MT, Johnson PJ. Ancient invasions: from endosymbionts to organelles. Science. 2004;304(5668):253-257. doi:10.1126/science.1094884.

434. Gray MW, Burger G, Lang BF. Mitochondrial evolution. Science. 1999;283(5407):1476-1481. doi:10.1126/science.283.5407.1476.

435. Friedman JR, Nunnari J. Mitochondrial form and function. Nature. 2014;505(7483):335-343. doi:10.1038/nature12985.

436. Lill R, Hoffmann B, Molik S, Pierik AJ, Rietzschel N, Stehling O, Uzarska MA, Webert H, Wilbrecht C, Muhlenhoff U. The role of mitochondria in cellular iron-sulfur protein biogenesis and iron metabolism. Biochim Biophys Acta. 2012;1823(9):1491-1508. doi:10.1016/j.bbamcr.2012.05.009.

437. Andersen JL, Kornbluth S. The tangled circuitry of metabolism and apoptosis. Mol Cell. 2013;49(3):399-410. doi:10.1016/j.molcel.2012.12.026.

438. Gilkerson R, Bravo L, Garcia I, Gaytan N, Herrera A, Maldonado A, Quintanilla B. The mitochondrial nucleoid: integrating mitochondrial DNA into cellular homeostasis. Cold Spring Harb Perspect Biol. 2013;5(5):a011080. doi:10.1101/cshperspect.a011080.

439. Michel AH, Kornmann B. The ERMES complex and ER-mitochondria connections. Biochem Soc Trans. 2012;40(2):445-450. doi:10.1042/bst20110758.

440. Schwarz TL. Mitochondrial trafficking in neurons. Cold Spring Harb Perspect Biol. 2013;5(6). doi:10.1101/cshperspect.a011304.

441. Luk E, Yang M, Jensen LT, Bourbonnais Y, Culotta VC. Manganese activation of superoxide dismutase 2 in the mitochondria of *Saccharomyces cerevisiae*. J Biol Chem. 2005;280(24):22715-22720. doi:10.1074/jbc.M504257200.

442. Yogev O, Karniely S, Pines O. Translation-coupled translocation of yeast fumarase into mitochondria in vivo. J Biol Chem. 2007;282(40):29222-29229. doi:10.1074/jbc.M704201200.

443. Hill K, Model K, Ryan MT, Dietmeier K, Martin F, Wagner R, Pfanner N. Tom40 forms the hydrophilic channel of the mitochondrial import pore for preproteins. Nature. 1998;395(6701):516-521. doi:10.1038/26780.

444. Kunkele KP, Heins S, Dembowski M, Nargang FE, Benz R, Thieffry M, Walz J, Lill R, Nussberger S, Neupert W. The preprotein translocation channel of the outer membrane of mitochondria. Cell. 1998;93(6):1009-1019. doi:10.1016/S0092-8674(00)81206-4.

445. Vogtle FN, Wortelkamp S, Zahedi RP, Becker D, Leidhold C, Gevaert K, Kellermann J, Voos W, Sickmann A, Pfanner N, Meisinger C. Global analysis of the mitochondrial N-proteome identifies a processing peptidase critical for protein stability. Cell. 2009;139(2):428-439. doi:10.1016/j.cell.2009.07.045.

446. Abe Y, Shodai T, Muto T, Mihara K, Torii H, Nishikawa S, Endo T, Kohda D. Structural basis of presequence recognition by the mitochondrial protein import receptor Tom20. Cell. 2000;100(5):551-560. doi:10.1016/S0092-8674(00)80691-1.

447. Saitoh T, Igura M, Obita T, Ose T, Kojima R, Maenaka K, Endo T, Kohda D. Tom20 recognizes mitochondrial presequences through dynamic equilibrium among multiple bound states. EMBO J. 2007;26(22):4777-4787. doi:10.1038/sj.emboj.7601888.

448. Gessmann D, Flinner N, Pfannstiel J, Schlosinger A, Schleiff E, Nussberger S, Mirus O. Structural elements of the mitochondrial preprotein-conducting channel Tom40 dissolved by bioinformatics and mass spectrometry. Biochim Biophys Acta. 2011;1807(12):1647-1657. doi:10.1016/j.bbabio.2011.08.006.

449. Tamura Y, Harada Y, Shiota T, Yamano K, Watanabe K, Yokota M, Yamamoto H, Sesaki H, Endo T. Tim23-Tim50 pair coordinates functions of translocators and motor proteins in mitochondrial protein import. J Cell Biol. 2009;184(1):129-141. doi:10.1083/jcb.200808068.

450. Martin J, Mahlke K, Pfanner N. Role of an energized inner membrane in mitochondrial protein import. $\Delta\psi$ drives the movement of presequences. J Biol Chem. 1991;266(27):18051-18057.

451. Kulawiak B, Hopker J, Gebert M, Guiard B, Wiedemann N, Gebert N. The mitochondrial protein import machinery has multiple connections to the respiratory chain. Biochim Biophys Acta. 2013;1827(5):612-626. doi:10.1016/j.bbabio.2012.12.004.

452. Ungermann C, Neupert W, Cyr DM. The role of Hsp70 in conferring unidirectionality on protein translocation into mitochondria. Science. 1994;266(5188):1250-1253. doi:10.1126/science.7973708

453. Liu Q, D'Silva P, Walter W, Marszalek J, Craig EA. Regulated cycling of mitochondrial Hsp70 at the protein import channel. Science. 2003;300(5616):139-141. doi:10.1126/science.1083379.

454. Hawlitschek G, Schneider H, Schmidt B, Tropschug M, Hartl FU, Neupert W. Mitochondrial protein import: identification of processing peptidase and of PEP, a processing enhancing protein. Cell. 1988;53(5):795-806. doi:10.1016/0092-8674(88)90096-7.

455. van der Laan M, Meinecke M, Dudek J, Hutu DP, Lind M, Perschil I, Guiard B, Wagner R, Pfanner N, Rehling P. Motor-free mitochondrial presequence translocase drives membrane integration of preproteins. Nat Cell Biol. 2007;9(10):1152-1159. doi:10.1038/ncb1635.

456. Koppen M, Langer T. Protein degradation within mitochondria: versatile activities of AAA proteases and other peptidases. Crit Rev Biochem Mol Biol. 2007;42(3):221-242. doi:10.1080/10409230701380452.

457. Milenkovic D, Ramming T, Muller JM, Wenz LS, Gebert N, Schulze-Specking A, Stojanovski D, Rospert S, Chacinska A. Identification of the signal directing Tim9 and Tim10 into the intermembrane space of mitochondria. Mol Biol Cell. 2009;20(10):2530-2539. doi:10.1091/mbc.E08-11-1108.

458. Sideris DP, Petrakis N, Katrakili N, Mikropoulou D, Gallo A, Ciofi-Baffoni S, Banci L, Bertini I, Tokatlidis K. A novel intermembrane space-targeting signal docks cysteines onto Mia40 during mitochondrial oxidative folding. J Cell Biol. 2009;187(7):1007-1022. doi:10.1083/jcb.200905134.

459. Stojanovski D, Bragoszewski P, Chacinska A. The MIA pathway: a tight bond between protein transport and oxidative folding in mitochondria. Biochim Biophys Acta. 2012;1823(7):1142-1150. doi:10.1016/j.bbamcr.2012.04.014.

460. Herrmann JM, Riemer J. Mitochondrial disulfide relay: redox-regulated protein import into the intermembrane space. J Biol Chem. 2012;287(7):4426-4433. doi:10.1074/jbc.R111.270678.

461. Riemer J, Fischer M, Herrmann JM. Oxidation-driven protein import into mitochondria: Insights and blind spots. Biochim Biophys Acta. 2011;1808(3):981-989. doi:10.1016/j.bbamem.2010.06.003.

462. Vascotto C, Bisetto E, Li M, Zeef LA, D'Ambrosio C, Domenis R, Comelli M, Delneri D, Scaloni A, Altieri F, Mavelli I, Quadrifoglio F, Kelley MR, Tell G. Knock-in reconstitution studies reveal an unexpected role of Cys-65 in regulating APE1/Ref-1 subcellular trafficking and function. Mol Biol Cell. 2011;22(20):3887-3901. doi:10.1091/mbc.E11-05-0391.

463. Zhuang J, Wang PY, Huang X, Chen X, Kang JG, Hwang PM. Mitochondrial disulfide relay mediates translocation of p53 and partitions its subcellular activity. Proc Natl Acad Sci USA. 2013;110(43):17356-17361. doi:10.1073/pnas.1310908110.

464. Young JC, Hoogenraad NJ, Hartl FU. Molecular chaperones Hsp90 and Hsp70 deliver preproteins to the mitochondrial import receptor Tom70. Cell. 2003;112(1):41-50. doi:10.1016/S0092-8674(02)01250-3.

465. Wiedemann N, Pfanner N, Ryan MT. The three modules of ADP/ATP carrier cooperate in receptor recruitment and translocation into mitochondria. EMBO J. 2001;20(5):951-960. doi:10.1093/emboj/20.5.951.

466. Koehler CM, Jarosch E, Tokatlidis K, Schmid K, Schweyen RJ, Schatz G. Import of mitochondrial carriers mediated by essential proteins of the intermembrane space. Science. 1998;279(5349):369-373. doi:10.1126/science.279.5349.369.

467. Sirrenberg C, Bauer MF, Guiard B, Neupert W, Brunner M. Import of carrier proteins into the mitochondrial inner membrane mediated by Tim22. Nature. 1996;384(6609):582-585. doi:10.1038/384582a0.

468. Ferramosca A, Zara V. Biogenesis of mitochondrial carrier proteins: molecular mechanisms of import into mitochondria. Biochim Biophys Acta. 2013;1833(3):494-502. doi:10.1016/j.bbamcr.2012.11.014.

469. Qiu J, Wenz LS, Zerbes RM, Oeljeklaus S, Bohnert M, Stroud DA, Wirth C, Ellenrieder L, Thornton N, Kutik S, Wiese S, Schulze-Specking A, Zufall N, Chacinska A, Guiard B, Hunte C, Warscheid B, van der Laan M, Pfanner N, Wiedemann N, Becker T. Coupling of mitochondrial import and export translocases by receptor-mediated supercomplex formation. Cell. 2013;154(3):596-608. doi:10.1016/j.cell.2013.06.033.

470. Wenz LS, Opalinski L, Schuler MH, Ellenrieder L, Ieva R, Bottinger L, Qiu J, van der Laan M, Wiedemann N, Guiard B, Pfanner N, Becker T. The presequence pathway is involved in protein sorting to the mitochondrial outer membrane. EMBO Rep. 2014;15(6):678-685. doi:10.1002/embr.201338144.

471. Becker T, Pfannschmidt S, Guiard B, Stojanovski D, Milenkovic D, Kutik S, Pfanner N, Meisinger C, Wiedemann N. Biogenesis of the mitochondrial TOM complex: Mim1 promotes insertion and assembly of signal-anchored receptors. J Biol Chem. 2008;283(1):120-127. doi:10.1074/jbc.M706997200.

472. Schmidt O, Pfanner N, Meisinger C. Mitochondrial protein import: from proteomics to functional mechanisms. Nat Rev Mol Cell Biol. 2010;11(9):655-667. doi:10.1038/nrm2959.

473. Dudek J, Rehling P, van der Laan M. Mitochondrial protein import: common principles and physiological networks. Biochim Biophys Acta. 2013;1833(2):274-285. doi:10.1016/j.bbamcr.2012.05.028.

474. Martinou JC, Youle RJ. Mitochondria in apoptosis: Bcl-2 family members and mitochondrial dynamics. Dev Cell. 2011;21(1):92-101. doi:10.1016/j.devcel.2011.06.017.

475. Yoon HS, Hackett JD, Ciniglia C, Pinto G, Bhattacharya D. A molecular timeline for the origin of photosynthetic eukaryotes. Mol Biol Evol. 2004;21(5):809-818. doi:10.1093/molbev/msh075.

476. Shi LX, Theg SM. The chloroplast protein import system: from algae to trees. Biochim Biophys Acta. 2013;1833(2):314-331. doi:10.1016/j.bbamcr.2012.10.002.

477. Hasan S, Platta HW, Erdmann R. Import of proteins into the peroxisomal matrix. Frontiers in physiology. 2013;4:261. doi:10.3389/fphys.2013.00261.

478. Schluter A, Real-Chicharro A, Gabaldon T, Sanchez-Jimenez F, Pujol A. PeroxisomeDB 2.0: an integrative view of the global peroxisomal metabolome. Nucleic Acids Res. 2010;38(Database issue):D800-805. doi:10.1093/nar/gkp935.

479. Meinecke M, Cizmowski C, Schliebs W, Kruger V, Beck S, Wagner R, Erdmann R. The peroxisomal importomer constitutes a large and highly dynamic pore. Nat Cell Biol. 2010;12(3):273-277. doi:10.1038/ncb2027.

480. Williams C. Going against the flow: A case for peroxisomal protein export. Biochim Biophys Acta. 2014;1843(7):1386-1392. doi:10.1016/j.bbamcr.2014.04.009.

481. Kee HL, Verhey KJ. Molecular connections between nuclear and ciliary import processes. Cilia. 2013;2(1):11. doi:10.1186/2046-2530-2-11.

482. Kee HL, Dishinger JF, Blasius TL, Liu CJ, Margolis B, Verhey KJ. A size-exclusion permeability barrier and nucleoporins characterize a ciliary pore complex that regulates transport into cilia. Nat Cell Biol. 2012;14(4):431-437. doi:10.1038/ncb2450.

483. Dishinger JF, Kee HL, Jenkins PM, Fan S, Hurd TW, Hammond JW, Truong YN, Margolis B, Martens JR, Verhey KJ. Ciliary entry of the kinesin-2 motor KIF17 is regulated by importin-β2 and RanGTP. Nat Cell Biol. 2010;12(7):703-710. doi:10.1038/ncb2073.

484. Fan S, Whiteman EL, Hurd TW, McIntyre JC, Dishinger JF, Liu CJ, Martens JR, Verhey KJ, Sajjan U, Margolis B. Induction of Ran GTP drives ciliogenesis. Mol Biol Cell. 2011;22(23):4539-4548. doi:10.1091/mbc.E11-03-0267.

485. Maiuri T, Woloshansky T, Xia J, Truant R. The huntingtin N17 domain is a multifunctional CRM1 and Ran-dependent nuclear and cilial export signal. Hum Mol Genet. 2013;22(7):1383-1394. doi:10.1093/hmg/dds554.

486. Carbonaro M, Escuin D, O'Brate A, Thadani-Mulero M, Giannakakou P. Microtubules regulate hypoxia-inducible factor-1α protein trafficking and activity: implications for taxane therapy. J Biol Chem. 2012;287(15):11859-11869. doi:10.1074/jbc.M112.345587.

487. Gong X, Ming X, Deng P, Jiang Y. Mechanisms regulating the nuclear translocation of p38 MAP kinase. J Biol Chem. 2010;110(6):1420-1429. doi:10.1002/jcb.22675.

488. Roth DM, Moseley GW, Glover D, Pouton CW, Jans DA. A microtubule-facilitated nuclear import pathway for cancer regulatory proteins. Traffic. 2007;8(6):673-686. doi:10.1111/j.1600-0854.2007.00564.x.

489. Shrum CK, Defrancisco D, Meffert MK. Stimulated nuclear translocation of NF-κB and shuttling differentially depend on dynein and the dynactin complex. Proc Natl Acad Sci USA. 2009;106(8):2647-2652. doi:10.1073/pnas.0806677106.

490. Nyathi Y, Wilkinson BM, Pool MR. Co-translational targeting and translocation of proteins to the endoplasmic reticulum. Biochim Biophys Acta. 2013;1833(11):2392-2402. doi:10.1016/j.bbamcr.2013.02.021.

491. Walter P, Blobel G. Translocation of proteins across the endoplasmic reticulum. II. Signal recognition protein (SRP) mediates the selective binding to microsomal membranes of in-vitro-assembled polysomes synthesizing secretory protein. J Cell Biol. 1981;91(2):551-556. doi:10.1083/jcb.91.2.551.

492. Walter P, Blobel G. Translocation of proteins across the endoplasmic reticulum III. Signal recognition protein (SRP) causes signal sequence-dependent and site-specific arrest of chain elongation that is released by microsomal membranes. J Cell Biol. 1981;91(2):557-561. doi:10.1083/jcb.91.2.557.

493. Walter P, Ibrahimi I, Blobel G. Translocation of proteins across the endoplasmic reticulum. I. Signal recognition protein (SRP) binds to in-vitro-assembled polysomes synthesizing secretory protein. J Cell Biol. 1981;91(2):545-550. doi:10.1083/jcb.91.2.545.

494. Voss M, Schroder B, Fluhrer R. Mechanism, specificity, and physiology of signal peptide peptidase (SPP) and SPP-like proteases. Biochim Biophys Acta. 2013;1828(12):2828-2839. doi:10.1016/j.bbamem.2013.03.033.

495. Lingappa VR. Control of protein topology at the endoplasmic reticulum. Cell Biophys. 1991;19(1-3):1-15.

496. Hermesh O, Jansen RP. Take the (RN)A-train: localization of mRNA to the endoplasmic reticulum. Biochim Biophys Acta. 2013;1833(11):2519-2525. doi:10.1016/j.bbamcr.2013.01.013.

497. Aviram N, Schuldiner M. Embracing the void-how much do we really know about targeting and translocation to the endoplasmic reticulum? Curr Opin Cell Biol. 2014;29c:8-17. doi:10.1016/j.ceb.2014.02.004.

498. Hampton RY, Sommer T. Finding the will and the way of ERAD substrate retrotranslocation. Curr Opin Cell Biol. 2012;24(4):460-466. doi:10.1016/j.ceb.2012.05.010.

499. D'Arcangelo JG, Stahmer KR, Miller EA. Vesicle-mediated export from the ER: COPII coat function and regulation. Biochim Biophys Acta. 2013;1833(11):2464-2472. doi:10.1016/j.bbamcr.2013.02.003.

500. Yogev O, Pines O. Dual targeting of mitochondrial proteins: mechanism, regulation and function. Biochim Biophys Acta. 2011;1808(3):1012-1020. doi:10.1016/j.bbamem.2010.07.004.

501. Carrie C, Small I. A reevaluation of dual-targeting of proteins to mitochondria and chloroplasts. Biochim Biophys Acta. 2013;1833(2):253-259. doi:10.1016/j.bbamcr.2012.05.029.

502. Ast J, Stiebler AC, Freitag J, Bolker M. Dual targeting of peroxisomal proteins. Frontiers in physiology. 2013;4:297. doi:10.3389/fphys.2013.00297.

503. Krause K, Krupinska K. Nuclear regulators with a second home in organelles. Trends Plant Sci. 2009;14(4):194-199. doi:10.1016/j.tplants.2009.01.005.

504. Carrie C, Whelan J. Widespread dual targeting of proteins in land plants: when, where, how and why. Plant Signal Behav. 2013;8(8). doi:10.4161/psb.25034.

505. Moll T, Tebb G, Surana U, Robitsch H, Nasmyth K. The role of phosphorylation and the CDC28 protein kinase in cell cycle-regulated nuclear import of the *S. cerevisiae* transcription factor SWI5. Cell. 1991;66(4):743-758. doi:10.1016/0092-8674(91)90118-I.

506. Sbia M, Parnell EJ, Yu Y, Olsen AE, Kretschmann KL, Voth WP, Stillman DJ. Regulation of the yeast Ace2 transcription factor during the cell cycle. J Biol Chem. 2008;283(17):11135-11145. doi:10.1074/jbc.M800196200.

507. Visintin R, Craig K, Hwang ES, Prinz S, Tyers M, Amon A. The phosphatase Cdc14 triggers mitotic exit by reversal of Cdk-dependent phosphorylation. Mol Cell. 1998;2(6):709-718. doi:10.1016/S1097-2765(00)80286-5.

508. Jans DA, Moll T, Nasmyth K, Jans P. Cyclin-dependent kinase site-regulated signal-dependent nuclear localization of the SWI5 yeast transcription factor in mammalian cells. J Biol Chem. 1995;270(29):17064-17067. doi:10.1074/jbc.270.29.17064.

509. Martinez JD, Pennington ME, Craven MT, Warters RL, Cress AE. Free radicals generated by ionizing radiation signal nuclear translocation of p53. Cell Growth Differ. 1997;8(9):941-949.

510. Zhang Y, Xiong Y. A p53 amino-terminal nuclear export signal inhibited by DNA damage-induced phosphorylation. Science. 2001;292(5523):1910-1915. doi:10.1126/science.1058637.

511. Banin S, Moyal L, Shieh S, Taya Y, Anderson CW, Chessa L, Smorodinsky NI, Prives C, Reiss Y, Shiloh Y, Ziv Y. Enhanced phosphorylation of p53 by ATM in response to DNA damage. Science. 1998;281(5383):1674-1677. doi:10.1126/science.281.5383.1674.

512. Canman CE, Lim DS, Cimprich KA, Taya Y, Tamai K, Sakaguchi K, Appella E, Kastan MB, Siliciano JD. Activation of the ATM kinase by ionizing radiation and phosphorylation of p53. Science. 1998;281(5383):1677-1679. doi:10.1126/science.281.5383.1677.

513. Tibbetts RS, Brumbaugh KM, Williams JM, Sarkaria JN, Cliby WA, Shieh SY, Taya Y, Prives C, Abraham RT. A role for ATR in the DNA damage-induced phosphorylation of p53. Genes Dev. 1999;13(2):152-157.

514. Schneiderhan N, Budde A, Zhang Y, Brune B. Nitric oxide induces phosphorylation of p53 and impairs nuclear export. Oncogene. 2003;22(19):2857-2868. doi:10.1038/sj.onc.1206431.

515. Lee JH, Jeong MW, Kim W, Choi YH, Kim KT. Cooperative roles of c-Abl and Cdk5 in regulation of p53 in response to oxidative stress. J Biol Chem. 2008;283(28):19826-19835. doi:10.1074/jbc.M706201200.

516. Sakaguchi K, Saito S, Higashimoto Y, Roy S, Anderson CW, Appella E. Damage-mediated phosphorylation of human p53 threonine 18 through a cascade mediated by a casein 1-like kinase. Effect on Mdm2 binding. J Biol Chem. 2000;275(13):9278-9283. doi:10.1074/jbc.275.13.9278.

517. Shieh SY, Taya Y, Prives C. DNA damage-inducible phosphorylation of p53 at N-terminal sites including a novel site, Ser20, requires tetramerization. EMBO J. 1999;18(7):1815-1823. doi:10.1093/emboj/18.7.1815.

518. Pikuleva IA, Waterman MR. Cytochromes p450: roles in diseases. J Biol Chem. 2013;288(24):17091-17098. doi:10.1074/jbc.R112.431916.

519. Dasari VR, Anandatheerthavarada HK, Robin MA, Boopathi E, Biswas G, Fang JK, Nebert DW, Avadhani NG. Role of protein kinase C-mediated protein phosphorylation in mitochondrial translocation of mouse CYP1A1, which contains a non-canonical targeting signal. J Biol Chem. 2006;281(41):30834-30847. doi:10.1074/jbc.M510725200.

520. Dietschy T, Shevelev I, Pena-Diaz J, Huhn D, Kuenzle S, Mak R, Miah MF, Hess D, Fey M, Hottiger MO, Janscak P, Stagljar I. p300-mediated acetylation of the Rothmund-Thomson-syndrome gene product RECQL4 regulates its subcellular localization. J Cell Sci. 2009;122(8):1258-1267. doi:10.1242/jcs.037747.

521. Zhu L, Santos NC, Kim KH. Small ubiquitin-like modifier-2 modification of retinoic acid receptor-α regulates its subcellular localization and transcriptional activity. Endocrinology. 2009;150(12):5586-5595. doi:10.1210/en.2009-0868.

522. Du JX, Bialkowska AB, McConnell BB, Yang VW. SUMOylation regulates nuclear localization of Krüppel-like factor 5. J Biol Chem. 2008;283(46):31991-32002. doi:10.1074/jbc.M803612200.

523. Zehorai E, Yao Z, Plotnikov A, Seger R. The subcellular localization of MEK and ERK—a novel nuclear translocation signal (NTS) paves a way to the nucleus. Mol Cell Endocrinol. 2010;314(2):213-220. doi:10.1016/j.mce.2009.04.008.

524. Gioeli D, Black BE, Gordon V, Spencer A, Kesler CT, Eblen ST, Paschal BM, Weber MJ. Stress kinase signaling regulates androgen receptor phosphorylation, transcription, and localization. Mol Endocrinol. 2006;20(3):503-515. doi:10.1210/me.2005-0351.

525. Chen S, Kesler CT, Paschal BM, Balk SP. Androgen receptor phosphorylation and activity are regulated by an association with protein phosphatase 1. J Biol Chem. 2009;284(38):25576-25584. doi:10.1074/jbc.M109.043133.

526. Onesti S, MacNeill SA. Structure and evolutionary origins of the CMG complex. Chromosoma. 2013;122(1-2):47-53. doi:10.1007/s00412-013-0397-x.

527. Dalton S, Whitbread L. Cell cycle-regulated nuclear import and export of Cdc47, a protein essential for initiation of DNA replication in budding yeast. Proc Natl Acad Sci USA. 1995;92(7):2514-2518.

528. Liku ME, Nguyen VQ, Rosales AW, Irie K, Li JJ. CDK phosphorylation of a novel NLS-NES module distributed between two subunits of the Mcm2-7 complex prevents chromosomal rereplication. Mol Biol Cell. 2005;16(10):5026-5039. doi:10.1091/mbc.E05-05-0412.

529. Kawaguchi T, Takenoshita M, Kabashima T, Uyeda K. Glucose and cAMP regulate the L-type pyruvate kinase gene by phosphorylation/dephosphorylation of the carbohydrate response element binding protein. Proc Natl Acad Sci USA. 2001;98(24):13710-13715. doi:10.1073/pnas.231370798.

530. Ge Q, Huang N, Wynn RM, Li Y, Du X, Miller B, Zhang H, Uyeda K. Structural characterization of a unique interface between carbohydrate response element-binding protein (ChREBP) and 14-3-3β protein. J Biol Chem. 2012;287(50):41914-41921. doi:10.1074/jbc.M112.418855.

531. Sakiyama H, Wynn RM, Lee WR, Fukasawa M, Mizuguchi H, Gardner KH, Repa JJ, Uyeda K. Regulation of nuclear import/export of carbohydrate response element-binding protein (ChREBP): interaction of an α-helix of ChREBP with the 14-3-3 proteins and regulation by phosphorylation. J Biol Chem. 2008;283(36):24899-24908. doi:10.1074/jbc.M804308200.

532. Davies MN, O'Callaghan BL, Towle HC. Glucose activates ChREBP by increasing its rate of nuclear entry and relieving repression of its transcriptional activity. J Biol Chem. 2008;283(35):24029-24038. doi:10.1074/jbc.M801539200.

533. Saporita AJ, Zhang Q, Navai N, Dincer Z, Hahn J, Cai X, Wang Z. Identification and characterization of a ligand-regulated nuclear export signal in androgen receptor. J Biol Chem. 2003;278(43):41998-42005. doi:10.1074/jbc.M302460200.

534. Moye-Rowley WS, Harshman KD, Parker CS. Yeast *YAP1* encodes a novel form of the jun family of transcriptional activator proteins. Genes Dev. 1989;3(3):283-292. doi:10.1101/gad.3.3.283.

535. Yan C, Lee LH, Davis LI. Crm1p mediates regulated nuclear export of a yeast AP-1-like transcription factor. EMBO J. 1998;17(24):7416-7429. doi:10.1093/emboj/17.24.7416.

536. Wood MJ, Storz G, Tjandra N. Structural basis for redox regulation of Yap1 transcription factor localization. Nature. 2004;430(7002):917-921. doi:10.1038/nature02790.

537. Okazaki S, Naganuma A, Kuge S. Peroxiredoxin-mediated redox regulation of the nuclear localization of Yap1, a transcription factor in budding yeast. Antioxid Redox Signal. 2005;7(3-4):327-334. doi:10.1089/ars.2005.7.327.

538. Rowe LA, Degtyareva N, Doetsch PW. DNA damage-induced reactive oxygen species (ROS) stress response in *Saccharomyces cerevisiae*. Free Radical Biol Med. 2008;45(8):1167-1177. doi:10.1016/j.freeradbiomed.2008.07.018.

539. Haeseleer F, Imanishi Y, Sokal I, Filipek S, Palczewski K. Calcium-binding proteins: intracellular sensors from the calmodulin superfamily. Biochem Biophys Res Commun. 2002;290(2):615-623. doi:10.1006/bbrc.2001.6228.

540. Liao B, Paschal BM, Luby-Phelps K. Mechanism of Ca^{2+}-dependent nuclear accumulation of calmodulin. Proc Natl Acad Sci USA. 1999;96(11):6217-6222.

541. Luby-Phelps K, Hori M, Phelps JM, Won D. Ca^{2+}-regulated dynamic compartmentalization of calmodulin in living smooth muscle cells. J Biol Chem. 1995;270(37):21532-21538. doi:10.1074/jbc.270.37.21532.

542. Greber UF, Gerace L. Depletion of calcium from the lumen of endoplasmic reticulum reversibly inhibits passive diffusion and signal-mediated transport into the nucleus. J Cell Biol. 1995;128(1-2):5-14.

543. Mermelstein PG, Deisseroth K, Dasgupta N, Isaksen AL, Tsien RW. Calmodulin priming: nuclear translocation of a calmodulin complex and the memory of prior neuronal activity. Proc Natl Acad Sci USA. 2001;98(26):15342-15347. doi:10.1073/pnas.211563998.

544. Itoh K, Wakabayashi N, Katoh Y, Ishii T, Igarashi K, Engel JD, Yamamoto M. Keap1 represses nuclear activation of antioxidant responsive elements by Nrf2 through binding to the amino-terminal Neh2 domain. Genes Dev. 1999;13(1):76-86. doi:10.1101/gad.13.1.76.

545. Kang KW, Lee SJ, Park JW, Kim SG. Phosphatidylinositol 3-kinase regulates nuclear translocation of NF-E2-related factor 2 through actin rearrangement in response to oxidative stress. Mol Pharmacol. 2002;62(5):1001-1010. doi:10.1124/mol.62.5.1001.

546. Kang MI, Kobayashi A, Wakabayashi N, Kim SG, Yamamoto M. Scaffolding of Keap1 to the actin cytoskeleton controls the function of Nrf2 as key regulator of cytoprotective phase 2 genes. Proc Natl Acad Sci USA. 2004;101(7):2046-2051. doi:10.1073/pnas.0308347100.

547. Buckley BJ, Marshall ZM, Whorton AR. Nitric oxide stimulates Nrf2 nuclear translocation in vascular endothelium. Biochem Biophys Res Commun. 2003;307(4):973-979. doi:10.1016/S0006-291X(03)01308-1.

548. Itoh K, Wakabayashi N, Katoh Y, Ishii T, O'Connor T, Yamamoto M. Keap1 regulates both cytoplasmic-nuclear shuttling and degradation of Nrf2 in response to electrophiles. Genes Cells. 2003;8(4):379-391. doi:10.1046/j.1365-2443.2003.00640.x.

549. Hsieh CY, Hsiao HY, Wu WY, Liu CA, Tsai YC, Chao YJ, Wang DL, Hsieh HJ. Regulation of shear-induced nuclear translocation of the Nrf2 transcription factor in endothelial cells. J Biomed Sci. 2009;16:12. doi:10.1186/1423-0127-16-12.

550. Kaspar JW, Niture SK, Jaiswal AK. Nrf2:INrf2 (Keap1) signaling in oxidative stress. Free Radical Biol Med. 2009;47(9):1304-1309. doi:10.1016/j.freeradbiomed.2009.07.035.

551. Niture SK, Jain AK, Jaiswal AK. Antioxidant-induced modification of INrf2 cysteine 151 and PKCδ-mediated phosphorylation of Nrf2 serine 40 are both required for stabilization and nuclear translocation of Nrf2 and increased drug resistance. J Cell Sci. 2009;122(24):4452-4464. doi:10.1242/jcs.058537.

552. Tsukimoto M, Tamaishi N, Homma T, Kojima S. Low-dose γ-ray irradiation induces translocation of Nrf2 into nuclear in mouse macrophage RAW264.7 cells. J Radiat Res. 2010;51(3):349-353. doi:10.1269/jrr.10002.

553. Niture SK, Khatri R, Jaiswal AK. Regulation of Nrf2—an update. Free Radical Biol Med. 2014;66:36-44. doi:10.1016/j.freeradbiomed.2013.02.008.

554. Armstrong EH, Goswami D, Griffin PR, Noy N, Ortlund EA. Structural basis for ligand regulation of the fatty acid binding protein 5, peroxisome proliferator activated receptor β/δ (FABP5-PPARβ/δ) signaling pathway. J Biol Chem. 2014. doi:10.1074/jbc.M113.514646.

555. Roth DM, Moseley GW, Pouton CW, Jans DA. Mechanism of microtubule-facilitated "fast track" nuclear import. J Biol Chem. 2011;286(16):14335-14351. doi:10.1074/jbc.M110.210302.

556. Salman H, Abu-Arish A, Oliel S, Loyter A, Klafter J, Granek R, Elbaum M. Nuclear localization signal peptides induce molecular delivery along microtubules. Biophys J. 2005;89(3):2134-2145. doi:10.1529/biophysj.105.060160.

557. Moseley GW, Roth DM, DeJesus MA, Leyton DL, Filmer RP, Pouton CW, Jans DA. Dynein light chain association sequences can facilitate nuclear protein import. Mol Biol Cell. 2007;18(8):3204-3213. doi:10.1091/mbc.E07-01-0030.

558. Wysolmerski JJ. Parathyroid hormone-related protein: an update. J Clin Endocrinol Metab. 2012;97(9):2947-2956. doi:doi:10.1210/jc.2012-2142.

559. Lam MH, Thomas RJ, Loveland KL, Schilders S, Gu M, Martin TJ, Gillespie MT, Jans DA. Nuclear transport of parathyroid hormone (PTH)-related protein is dependent on microtubules. Mol Endocrinol. 2002;16(2):390-401. doi:10.1210/mend.16.2.0775.

560. Cingolani G, Bednenko J, Gillespie MT, Gerace L. Molecular basis for the recognition of a nonclassical nuclear localization signal by importin β. Mol Cell. 2002;10(6):1345-1353. doi:10.1016/S1097-2765(02)00727-X.

561. Lam MH, House CM, Tiganis T, Mitchelhill KI, Sarcevic B, Cures A, Ramsay R, Kemp BE, Martin TJ, Gillespie MT. Phosphorylation at the cyclin-dependent kinases site (Thr85) of parathyroid hormone-related protein

negatively regulates its nuclear localization. J Biol Chem. 1999;274(26):18559-18566. doi:10.1074/jbc.274.26.18559

562. Wong-Riley MT, Besharse JC. The kinesin superfamily protein KIF17: one protein with many functions. Biomol Concepts. 2012;3(3):267-282. doi:10.1515/bmc-2011-0064.

563. Song W, Nadeau P, Yuan M, Yang X, Shen J, Yankner BA. Proteolytic release and nuclear translocation of Notch-1 are induced by presenilin-1 and impaired by pathogenic presenilin-1 mutations. Proc Natl Acad Sci USA. 1999;96(12):6959-6963. doi:10.1073/pnas.96.12.6959.

564. Berezovska O, Jack C, McLean P, Aster JC, Hicks C, Xia W, Wolfe MS, Weinmaster G, Selkoe DJ, Hyman BT. Rapid Notch1 nuclear translocation after ligand binding depends on presenilin-associated γ-secretase activity. Ann N Y Acad Sci. 2000;920:223-226. doi:10.1111/j.1749-6632.2000.tb06926.x.

565. Shimizu K, Chiba S, Hosoya N, Kumano K, Saito T, Kurokawa M, Kanda Y, Hamada Y, Hirai H. Binding of Delta1, Jagged1, and Jagged2 to Notch2 rapidly induces cleavage, nuclear translocation, and hyperphosphorylation of Notch2. Mol Cell Biol. 2000;20(18):6913-6922. doi:10.1128/MCB.20.18.6913-6922.2000.

566. Baron M, Aslam H, Flasza M, Fostier M, Higgs JE, Mazaleyrat SL, Wilkin MB. Multiple levels of Notch signal regulation. Mol Membr Biol. 2002;19(1):27-38.

567. Fortini ME. Notch signaling: the core pathway and its posttranslational regulation. Dev Cell. 2009;16(5):633-647. doi:10.1016/j.devcel.2009.03.010.

568. Wen W, Zhu F, Zhang J, Keum YS, Zykova T, Yao K, Peng C, Zheng D, Cho YY, Ma WY, Bode AM, Dong Z. MST1 promotes apoptosis through phosphorylation of histone H2AX. J Biol Chem. 2010;285(50):39108-39116. doi:10.1074/jbc.M110.151753.

569. Ura S, Masuyama N, Graves JD, Gotoh Y. Caspase cleavage of MST1 promotes nuclear translocation and chromatin condensation. Proc Natl Acad Sci USA. 2001;98(18):10148-10153. doi:10.1073/pnas.181161698.

570. Fry AM, O'Regan L, Sabir SR, Bayliss R. Cell cycle regulation by the NEK family of protein kinases. J Cell Sci. 2012;125(19):4423-4433. doi:10.1242/jcs.111195.

571. Wu W, Baxter JE, Wattam SL, Hayward DG, Fardilha M, Knebel A, Ford EM, da Cruz e Silva EF, Fry AM. Alternative splicing controls nuclear translocation of the cell cycle-regulated Nek2 kinase. J Biol Chem. 2007;282(36):26431-26440. doi:10.1074/jbc.M704969200.

572. Sharp DJ, Ross JL. Microtubule-severing enzymes at the cutting edge. J Cell Sci. 2012;125(11):2561-2569. doi:10.1242/jcs.101139.

573. Claudiani P, Riano E, Errico A, Andolfi G, Rugarli EI. Spastin subcellular localization is regulated through usage of different translation start sites and active export from the nucleus. Exp Cell Res. 2005;309(2):358-369. doi:10.1016/j.yexcr.2005.06.009.

574. Levadoux M, Mahon C, Beattie JH, Wallace HM, Hesketh JE. Nuclear import of metallothionein requires its mRNA to be associated with the perinuclear cytoskeleton. J Biol Chem. 1999;274(49):34961-34966. doi:10.1074/jbc.274.49.34961.

575. Gadir N, Haim-Vilmovsky L, Kraut-Cohen J, Gerst JE. Localization of mRNAs coding for mitochondrial proteins in the yeast *Saccharomyces cerevisiae*. Rna. 2011;17(8):1551-1565. doi:10.1261/rna.2621111.

576. Zipor G, Haim-Vilmovsky L, Gelin-Licht R, Gadir N, Brocard C, Gerst JE. Localization of mRNAs coding for peroxisomal proteins in the yeast, *Saccharomyces cerevisiae*. Proc Natl Acad Sci USA. 2009;106(47):19848-19853. doi:10.1073/pnas.0910754106.

577. Holt CE, Schuman EM. The central dogma decentralized: new perspectives on RNA function and local translation in neurons. Neuron. 2013;80(3):648-657. doi:10.1016/j.neuron.2013.10.036.

578. Bouche G, Gas N, Prats H, Baldin V, Tauber JP, Teissie J, Amalric F. Basic fibroblast growth factor enters the nucleolus and stimulates the transcription of ribosomal genes in ABAE cells undergoing G0→G1 transition. Proc Natl Acad Sci USA. 1987;84(19):6770-6774.

579. Baldin V, Roman AM, Bosc-Bierne I, Amalric F, Bouche G. Translocation of bFGF to the nucleus is G1 phase cell cycle specific in bovine aortic endothelial cells. EMBO J. 1990;9(5):1511-1517.

580. Imamura T, Engleka K, Zhan X, Tokita Y, Forough R, Roeder D, Jackson A, Maier JA, Hla T, Maciag T. Recovery of mitogenic activity of a growth factor mutant with a nuclear translocation sequence. Science. 1990;249(4976):1567-1570. doi:10.1126/science.1699274.

581. Amalric F, Baldin V, Bosc-Bierne I, Bugler B, Couderc B, Guyader M, Patry V, Prats H, Roman AM, Bouche G. Nuclear translocation of basic fibroblast growth factor. Ann N Y Acad Sci. 1991;638:127-138. doi:10.1111/j.1749-6632.1991.tb49023.x.

582. Lobie PE, Mertani H, Morel G, Morales-Bustos O, Norstedt G, Waters MJ. Receptor-mediated nuclear translocation of growth hormone. J Biol Chem. 1994;269(33):21330-21339.

583. Mertani HC, Raccurt M, Abbate A, Kindblom J, Tornell J, Billestrup N, Usson Y, Morel G, Lobie PE. Nuclear translocation and retention of growth hormone. Endocrinology. 2003;144(7):3182-3195. doi:10.1210/en.2002-221121.

584. Curtis BM, Widmer MB, deRoos P, Qwarnstrom EE. IL-1 and its receptor are translocated to the nucleus. J Immunol. 1990;144(4):1295-1303.

585. Ookawara T, Kizaki T, Takayama E, Imazeki N, Matsubara O, Ikeda Y, Suzuki K, Li Ji L, Tadakuma T, Taniguchi N, Ohno H. Nuclear translocation of extracellular superoxide dismutase. Biochem Biophys Res Commun. 2002;296(1):54-61. doi:.1016/S0006-291X(02)00804-5.

586. Cheung ZH, Ip NY. Cdk5: a multifaceted kinase in neurodegenerative diseases. Trends Cell Biol. 2012;22(3):169-175. doi:10.1016/j.tcb.2011.11.003.

587. Asada A, Yamamoto N, Gohda M, Saito T, Hayashi N, Hisanaga S. Myristoylation of p39 and p35 is a determinant of cytoplasmic or nuclear localization of active cyclin-dependent kinase 5 complexes. J Neurochem. 2008;106(3):1325-1336. doi:10.1111/j.1471-4159.2008.05500.x.

588. Martin DD, Beauchamp E, Berthiaume LG. Post-translational myristoylation: fat matters in cellular life and death. Biochimie. 2011;93(1):18-31. doi:10.1016/j.biochi.2010.10.018.

589. Fu X, Choi Y-K, Qu D, Yu Y, Cheung NS, Qi RZ. Identification of nuclear import mechanisms for the neuronal Cdk5 activator. J Biol Chem. 2006;281(51):39014-39021. doi:10.1074/jbc.M512663200.

590. Asada A, Saito T, Hisanaga S. Phosphorylation of p35 and p39 by Cdk5 determines the subcellular location of the holokinase in a phosphorylation-site-specific manner. J Cell Sci. 2012;125(14):3421-3429. doi:10.1242/jcs.100503.

591. Patrick GN, Zukerberg L, Nikolic M, de la Monte S, Dikkes P, Tsai LH. Conversion of p35 to p25 deregulates Cdk5 activity and promotes neurodegeneration. Nature. 1999;402(6762):615-622. doi:10.1038/45159.

592. Saito T, Konno T, Hosokawa T, Asada A, Ishiguro K, Hisanaga S-i. p25/Cyclin-dependent kinase 5 promotes the progression of cell death in nucleus of endoplasmic reticulum-stressed neurons. J Neurochem. 2007;102(1):133-140. doi:10.1111/j.1471-4159.2007.04540.x.

593. Chang KH, Multani PS, Sun KH, Vincent F, de Pablo Y, Ghosh S, Gupta R, Lee HP, Lee HG, Smith MA, Shah K. Nuclear envelope dispersion triggered by deregulated Cdk5 precedes neuronal death. Mol Biol Cell. 2011;22(9):1452-1462. doi:10.1091/mbc.E10-07-0654.

594. Stommel JM, Marchenko ND, Jimenez GS, Moll UM, Hope TJ, Wahl GM. A leucine-rich nuclear export signal in the p53 tetramerization domain: regulation of subcellular localization and p53 activity by NES masking. EMBO J. 1999;18(6):1660-1672. doi:10.1093/emboj/18.6.1660.

595. Komlodi-Pasztor E, Trostel S, Sackett D, Poruchynsky M, Fojo T. Impaired p53 binding to importin: a novel mechanism of cytoplasmic sequestration identified in oxaliplatin-resistant cells. Oncogene. 2009;28(35):3111-3120. doi:10.1038/onc.2009.166.

596. Carter S, Bischof O, Dejean A, Vousden KH. C-terminal modifications regulate MDM2 dissociation and nuclear export of p53. Nat Cell Biol. 2007;9(4):428-435. doi:10.1038/ncb1562.

597. Yuan J, Luo K, Zhang L, Cheville JC, Lou Z. USP10 regulates p53 localization and stability by deubiquitinating p53. Cell. 2010;140(3):384-396. doi:10.1016/j.cell.2009.12.032.

598. Li Q, Martinez JD. P53 is transported into the nucleus via an Hsf1-dependent nuclear localization mechanism. Mol Carcinog. 2011;50(2):143-152. doi:10.1002/mc.20713.

599. Giannakakou P, Sackett DL, Ward Y, Webster KR, Blagosklonny MV, Fojo T. p53 is associated with cellular microtubules and is transported to the nucleus by dynein. Nat Cell Biol. 2000;2(10):709-717. doi:10.1038/35036335.

600. De S, Kumari J, Mudgal R, Modi P, Gupta S, Futami K, Goto H, Lindor NM, Furuichi Y, Mohanty D, Sengupta S. RECQL4 is essential for the transport of p53 to mitochondria in normal human cells in the absence of exogenous stress. J Cell Sci. 2012;125(10):2509-2522. doi:10.1242/jcs.101501.

601. Briscoe J, Therond PP. The mechanisms of Hedgehog signalling and its roles in development and disease. Nat Rev Mol Cell Biol. 2013;14(7):416-429. doi:10.1038/nrm3598.

602. Hatayama M, Aruga J. Gli protein nuclear localization signal. Vitam Horm. 2012;88:73-89. doi:10.1016/B978-0-12-394622-5.00004-3.

603. Shi Q, Han Y, Jiang J. Suppressor of fused impedes Ci/Gli nuclear import by opposing Trn/Kapβ2 in Hedgehog signaling. J Cell Sci. 2014;127(5):1092-1103. doi:10.1242/jcs.142828.

604. Robbins DJ, Nybakken KE, Kobayashi R, Sisson JC, Bishop JM, Therond PP. Hedgehog elicits signal transduction by means of a large complex containing the kinesin-related protein costal2. Cell. 1997;90(2):225-234. doi:10.1016/S0092-8674(00)80331-1.

605. Sisson JC, Ho KS, Suyama K, Scott MP. Costal2, a novel kinesin-related protein in the Hedgehog signaling pathway. Cell. 1997;90(2):235-245. doi:10.1016/S0092-8674(00)80332-3.

606. Liem KF, Jr., He M, Ocbina PJ, Anderson KV. Mouse Kif7/Costal2 is a cilia-associated protein that regulates Sonic hedgehog signaling. Proc Natl Acad Sci USA. 2009;106(32):13377-13382. doi:10.1073/pnas.0906944106.

607. Sheng T, Chi S, Zhang X, Xie J. Regulation of Gli1 localization by the cAMP/protein kinase A signaling axis through a site near the nuclear localization signal. J Biol Chem. 2006;281(1):9-12. doi:10.1074/jbc.C500300200.

608. Monnier V, Dussillol F, Alves G, Lamour-Isnard C, Plessis A. Suppressor of fused links fused and Cubitus interruptus on the hedgehog signalling pathway. Curr Biol. 1998;8(10):583-586.

609. Wang QT, Holmgren RA. The subcellular localization and activity of *Drosophila* cubitus interruptus are regulated at multiple levels. Development. 1999;126(22):5097-5106.

610. Sisson BE, Ziegenhorn SL, Holmgren RA. Regulation of Ci and Su(fu) nuclear import in *Drosophila*. Dev Biol. 2006;294(1):258-270. doi:10.1016/j.ydbio.2006.02.050.

611. Aza-Blanc P, Ramirez-Weber FA, Laget MP, Schwartz C, Kornberg TB. Proteolysis that is inhibited by hedgehog targets Cubitus interruptus protein to the nucleus and converts it to a repressor. Cell. 1997;89(7):1043-1053. doi:10.1016/S0092-8674(00)80292-5.

612. Zhang Y, Mao F, Lu Y, Wu W, Zhang L, Zhao Y. Transduction of the Hedgehog signal through the dimerization of Fused and the nuclear translocation of Cubitus interruptus. Cell Res. 2011;21(10):1436-1451. doi:10.1038/cr.2011.136.

613. Ohlmeyer JT, Kalderon D. Hedgehog stimulates maturation of Cubitus interruptus into a labile transcriptional activator. Nature. 1998;396(6713):749-753. doi:10.1038/25533.

614. Potter VR, Reif AE. Inhibition of an electron transport component by antimycin A. J Biol Chem. 1952;194(1):287-297.

615. Mitra S, Izumi T, Boldogh I, Bhakat KK, Chattopadhyay R, Szczesny B. Intracellular trafficking and regulation of mammalian AP-endonuclease 1 (APE1), an essential DNA repair protein. DNA Repair. 2007;6(4):461-469. doi:10.1016/j.dnarep.2006.10.010.

616. Qu J, Liu GH, Huang B, Chen C. Nitric oxide controls nuclear export of APE1/Ref-1 through *S*-nitrosation of cysteines 93 and 310. Nucleic Acids Res. 2007;35(8):2522-2532. doi:10.1093/nar/gkl1163.

617. Nakabeppu Y. Regulation of intracellular localization of human MTH1, OGG1, and MYH proteins for repair of oxidative DNA damage. Prog Nucleic Acid Res Mol Biol. 2001;68:75-94.

618. da Costa NM, Hautefeuille A, Cros MP, Melendez ME, Waters T, Swann P, Hainaut P, Pinto LF. Transcriptional regulation of thymine DNA glycosylase (TDG) by the tumor suppressor protein p53. Cell Cycle. 2012;11(24):4570-4578. doi:10.4161/cc.22843.

619. Fisher LA, Bessho M, Wakasugi M, Matsunaga T, Bessho T. Role of interaction of XPF with RPA in nucleotide excision repair. J Mol Biol. 2011;413(2):337-346. doi:10.1016/j.jmb.2011.08.034.

620. Goto M, Shinmura K, Igarashi H, Kobayashi M, Konno H, Yamada H, Iwaizumi M, Kageyama S, Tsuneyoshi T, Tsugane S, Sugimura H. Altered expression of the human base excision repair gene *NTH1* in gastric cancer. Carcinogenesis. 2009;30(8):1345-1352. doi:10.1093/carcin/bgp108.

621. Koketsu S, Watanabe T, Nagawa H. Expression of DNA repair protein: MYH, NTH1, and MTH1 in colorectal cancer. Hepatogastroenterology. 2004;51(57):638-642.

Chapter 2
Sequence Analysis of Base Excision and Strand Incision Repair Proteins

The functions of a protein are fully defined by the sequence of amino acid residues of which it is composed. This sequence determines the protein's three-dimensional structure and its ability to interact with other proteins. Chemical modifications to this sequence allow the function of a protein to be rapidly modulated based on cellular conditions. Extensive work has been done to characterize the sequences and structural characteristics important for determining the functions, interactions, and modifications that a protein may undergo. This work has been combined with the power of computational analysis to predict the structures and functions of previously uncharacterized proteins based solely on primary sequence and by homology to other proteins with known structures and functions. These tools enable the detection of potential functionalities and provide a starting point for experimental work examining their role in the protein's overall function. Before starting the project in earnest, I ran the core set of *Saccharomyces cerevisiae* base excision and strand incision repair (BESIR) protein sequences (except the RNase H2 complex) through available prediction algorithms for localization signals and post-translational modifications. Since that time, new tools have been developed and old ones updated. In this chapter, I present and briefly discuss the updated results of these sequence analyses, which I have expanded to the human proteins. These analyses provided insight into potential modes of regulation, including localization, contributing an important resource for future work.

Sequence Analysis Algorithms

Various laboratory groups have developed software for sequence analysis. The standard practice over the last decade has been to release the tools on a web site that may be accessed freely. However, some tools are only available for local use, and a minority may only be made available by request. As a result, these tools are scattered around the Internet and

in the literature and are not readily discoverable or accessible. Many web services and software packages have disappeared without a trace (e.g. for acetylation: LysAcet (1), PredMod (2), PLMLA (3); for sumoylation: PSFS-SUMO (4), SUMOpre (5)), while still others are broken in some way (e.g. for acetylation: EnsemblePail (6), LAceP (7); for sumoylation: PCI-SUMO (8)). What's especially troubling is that these losses are occurring in less than a decade since their publication. Local software, often written for Linux, is plagued by poor user interfaces and intricate dependencies on other packages and manual settings, forming a barrier to use that vastly limit the impact of the algorithms they implement. Availability by request depends on being able to track down the researchers, hope that they can find the software, and hope that enough documentation was written for you to be able to understand it. Scientific software more broadly suffers from: 1) development by biologists who do some informal programming or by programmers who do not fully understand the needs of biologists; and 2) scientists facing pressure to focus on doing novel research, rather than maintaining and improving software related to past research.

These factors form a significant barrier preventing the broader biological research community from benefiting from the advances coming out of bioinformatics and computational biology. The field desperately needs standardization of the way local sequence information is stored, analyzed, and annotated, just as is slowly being done for genomic sequence databases. These problems are beyond the scope of this work, and must simply be accepted. This section lists the prediction algorithms I employed and describes how results from multiple predictors for the same motifs were combined.

Modification site predictors

Combining scores generally

Scores from modification site predictors were normalized to a scale from 0 to 1. Algorithms which output None/Low/Medium/High categories were transcoded to 0.000

/0.333/0.667/1.000, respectively. Algorithms which produce a binary answer with a confidence value of 0 to 1 were transcoded to -1 (definitely not modified) to 1 (definitely modified) and then to 0 to 1. Algorithms which do positive-only detection were scored as 1 or were not included. The normalized scores were averaged to produce a combined prediction score, and the thresholds for medium and high probability are modification-dependent. I want to emphasize that this method is not by any means robust or statistically sound. Rather, it is only being used to get a rough picture of the consensus predictions of these algorithms.

Acetylation

Internal lysine acetylation was predicted with PAIL (9), the predictor available at the PHOSIDA database (10), BPBPHKA (11), and the Protein Peptide Scanner (PPS; http://pps.biocuckoo.org). PAIL produces scores which it ranks as no probability (score ≤ 0.1), low probability (score > 0.1), medium probability (score > 0.2) or high probability (score > 0.5). PPS provides positive-only detection of acetylation motifs. The average of the normalized values was computed, and averages ≥ 0.33 were considered to be medium-probability sites and those ≥ 0.67 were considered to be high-probability sites.

Methylation

Lysine and arginine methylation was predicted by MASA (12), which directly produces a probability on a scale of 0 to 1. A probability of ≥ 0.5 was considered to indicate medium-probability sites and those ≥ 0.75 were considered to be high-probability sites.

Palmitoylation

Palmitoylation was predicted by CSS-Palm 4.0 (13), which bins results as Low, Medium, or High probability, with sequence context-dependent thresholds.

Phosphorylation

Phosphorylation was predicted by NetPhos 2.0 (14) for human proteins and

NetPhosYeast 1.0 (15) for yeast proteins, and the predictor available at PHOSIDA (16), along with positive-only detection by PPS. A probability of ≥ 0.5 was considered to indicate medium-probability sites and those ≥ 0.75 were considered to be high-probability sites.

Sumoylation

SUMO modification of lysines was predicted by SUMOplot (Abgent, http://www.abgent.com/sumoplot) and GPS-SUMO (17). PCI-SUMO was used for a subset of proteins when it was available (8), and it produces a binary answer with a confidence of 0 to 1. PPS provided positive-only detection. A probability of ≥ 0.5 was considered to indicate medium-probability sites and those ≥ 0.75 were considered to be high-probability sites.

Ubiquitination

UbPred (18) was used to predict ubiquitin-modified lysines. It considers values between 0.62 and 0.69 to be low-confidence, 0.69 and 0.84 to be medium-confidence, and 0.84 and 1.00 to be high-confidence.

Localization signal predictors

Nuclear localization signal

Nuclear localization signals (NLSs) were detected by the NUCDISC component of PSORT II (19).

Nuclear export signal

Nuclear export signals (NESs) were detected by NetNES 1.1 (20). If present, the high-scoring residues around a residue crossing the NES threshold were noted (generally, > 0.100). Otherwise, a high-scoring region which otherwise did not reach the NES threshold was also recorded as a weak NES. Some proteins were also analyzed with the Linux program NESsential 1.0, which indicates short NES-like sequences and assigns a confidence score of 0 to 1. Those with a score of > 0.5 were taken as positive hits (21).

Mitochondrial matrix targeting signal

Mitochondrial matrix targeting signals (MTSs) were predicted by the MITDISC and Gavel components of PSORT II (19), iPSORT (22), MitoProt II (23), and TargetP 1.1 (24). A qualitative comparison was used to synthesize the results.

Domains and other motif predictor

Functional motifs were identified by ProSite (25).

Results and Discussion

The sequence analysis results are displayed in Figure 2-1. Only the high-probability sites for small modifications (phosphoryl, acetyl, methyl, palmitoyl) and both medium- and high-probability sites for large modifications (ubiquitin, SUMO) are shown. The raw results and calculations are available in supplemental files identified in the Appendix.

All but a handful of the analyzed proteins have both a predicted MTS, at least one predicted NLS, and at least one predicted NES. Proteins which have only been detected in the nucleus generally do not possess a predicted MTS, including Ntg2, Mag1, Apn2, NEIL3, SMUG1, MBD4, and TDG. However, there are some surprises. The two isoforms of UNG are the nuclear UNG2 and mitochondrial UNG1, which contain different N-termini from alternate promoters. UNG2 has been reported as only nuclear, yet three of the four MTS algorithms (PSORTII/MITDISC, MitoProt II, TargetP) predict the presence of an MTS which scores higher than the MTS of UNG1. The newer algorithm iPSORT, however, predicts no MTS in UNG2 because it lacks a particular pattern of amino acids in the N-terminal region even though it has a high net positive charge. Perhaps this result simply highlights the imperfect nature of these predictors, and that everything predicted by them must be experimentally validated. In an opposite case, APEX1 has no predicted MTS, and yet it has been observed in mitochondria. Since APEX1 contains cysteine residues, it may be a candidate for import into mitochondria by a MIA-like mechanism, similar to that used by TP53 (26). Meanwhile, NEIL2, SMUG2, and

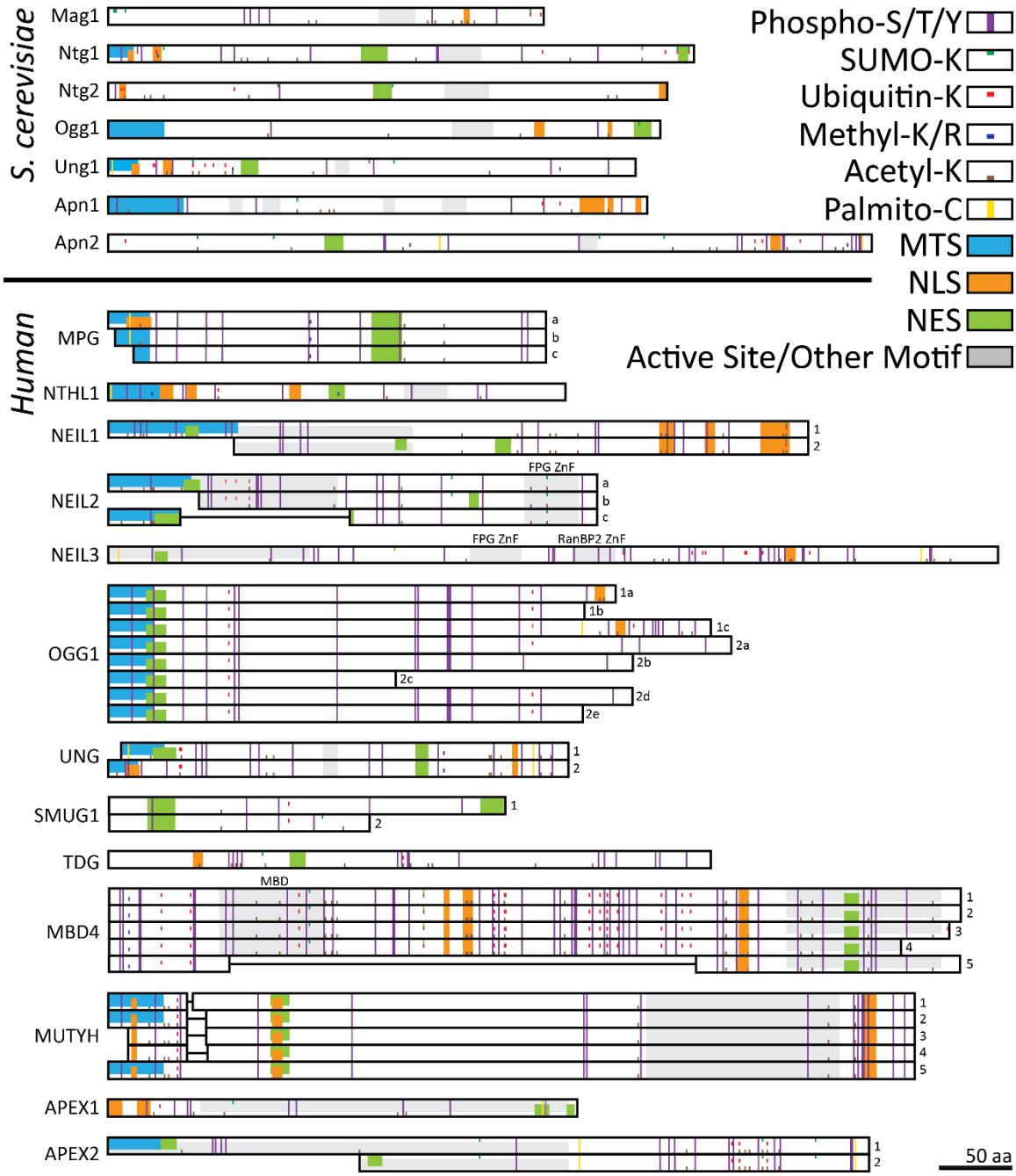

Figure 2-1. Schematic of predicted localization signals and modification sites in *S. cerevisiae* and human BESIR proteins. High-probability small modifications (methyl, acetyl, palmitoyl, phosphoryl) and medium- to high-probability large modifications (ubiquitin, SUMO) are shown. All proteins are to scale. Isoforms are named to the right of the diagram and are aligned by common domain and length. Gray regions which indicate other named motifs are noted above the region.

APEX2 have no predicted NLS, though the NLS could be conformational rather than sequential. Some of these signals have been experimentally validated. The two predicted NLS sequences in Ntg1 were recently determined to be a single bipartite NLS with an unusually

long linker (27,28), and of the three predicted NLSs of NTHL1, the most C-terminal one was found to be dispensable for nuclear localization (29). The C-terminal region of Apn1 was confirmed to contain its NLS (30). The N-terminal region of Ung1 was confirmed to contain its MTS (31). Lastly, the localization signals of OGG1 and MUTYH were discussed in (32).

Few of the dual-localized proteins which have alternate isoforms appear to have exclusively localized isoforms. Many of the alternate isoforms retain both predicted MTS and NLS sequences, including UNG1, MPGa, NEIL1-1, MYH-1/2/5, and APEX2-1. Several of these share the intriguing overlapping NLS and MTS topology observed with Ntg1 and Ung1, including NTHL1, MPGa, and MYH-1/2/5. Such overlaps could invite competition between localization signal receptors and provide a direct point of regulation for the localization of these proteins.

Intriguingly, many of the analyzed proteins which localize to mitochondria have strongly predicted palmitoylation sites, especially the human mitochondrial proteins: Ung1, NTHL1, UNG1, MPGa/b, OGG1-1c, APEX1, APEX2-1/2. Recent studies have identified palmitoylation or myristoylation of several proteins as necessary for proper localization to the mitochondrial membrane, intermembrane space, or matrix (33-35). It is tempting to speculate that the putative palmitoylation sites on the analyzed proteins may promote association with the large surface of the mitochondrial membrane as a first step in mitochondrial import.

Most of the analyzed proteins contain putative sumoylation and ubiquitination sites. Ntg1 has been experimentally verified as a SUMO-modified protein at 5 different sites (36,37) with up to approximately 5% of the total Ntg1 in a cell sumoylated under oxidative stress. SUMO modification of TDG is involved in the turnover step of its glycosylase activity. SUMO has also been identified modifying several other DNA repair proteins, and it has been proposed that these modifications may coordinate the formation of repair complexes. However,

the functions of sumoylation of these proteins are only beginning to be probed.

Putative acetylation and phosphorylation sites have been identified on every protein, and most proteins have several of these sites lying within or near predicted localization signals. As modifications in these positions are known to be able to modify the binding affinity of the localization signal for it receptor, they are prime candidates for modulating the localization of their containing protein. Only a few of these marks have been experimentally verified thus far. For example, two of these sites, at S^{12} and S^{14} in UNG2, lie within the predicted MTS discussed above. S^{12} is part of a motif recognized by CK2, and S^{14} is part of a motif recognized by CDK1 and CDK2. Part of the reason this region was scored as a likely MTS is the positive charge in the region. Phosphorylation of these residues would reduce the net charge. Analyzing the UNG2 sequence with phosphomimetic substitutions of S12D and S14D results in a vastly reduced probability of mitochondrial localization. These phosphorylation sites also lie at the N-terminus of a predicted NLS, which could disrupt importin binding. However, running the phosphomimetic mutant through the PSORT II/NUCDISC algorithm does not change the location or strength of the predicted NLSs, and even if these phosphorylation marks did negatively impact the putative NLS, the C-terminal putative NLS would still be available.

Little else can be said directly from these analyses, but they form an important resource for understanding experimental results and formulating new hypotheses regarding the localization and function of the BESIR proteins.

References

1. Li S, Li H, Li M, Shyr Y, Xie L, Li Y. Improved prediction of lysine acetylation by support vector machines. Protein Peptide Lett. 2009;16(8):977-983.

2. Basu A, Rose KL, Zhang J, Beavis RC, Ueberheide B, Garcia BA, Chait B, Zhao Y, Hunt DF, Segal E, Allis CD, Hake SB. Proteome-wide prediction of acetylation substrates. Proc Natl Acad Sci USA. 2009;106(33):13785-13790. doi:10.1073/pnas.0906801106.

3. Shi SP, Qiu JD, Sun XY, Suo SB, Huang SY, Liang RP. PLMLA: prediction of lysine methylation and lysine acetylation by combining multiple features. Molecular bioSystems. 2012;8(5):1520-1527. doi:10.1039/c2mb05502c.

4. Liu B, Li S, Wang Y, Lu L, Li Y, Cai Y. Predicting the protein SUMO modification sites based on Properties Sequential Forward Selection (PSFS). Biochem Biophys Res Commun. 2007;358(1):136-139. doi:10.1016/j.bbrc.2007.04.097.

5. 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. doi:10.1186/1471-2105-9-8.

6. Xu Y, Wang XB, Ding J, Wu LY, Deng NY. Lysine acetylation sites prediction using an ensemble of support vector machine classifiers. J Theor Biol. 2010;264(1):130-135. doi:10.1016/j.jtbi.2010.01.013.

7. Hou T, Zheng G, Zhang P, Jia J, Li J, Xie L, Wei C, Li Y. LAceP: lysine acetylation site prediction using logistic regression classifiers. PLoS One. 2014;9(2):e89575. doi:10.1371/journal.pone.0089575.

8. Jones DT. Protein secondary structure prediction based on position-specific scoring matrices. J Mol Biol. 1999;292(2):195-202. doi:10.1006/jmbi.1999.3091.

9. Li A, Xue Y, Jin C, Wang M, Yao X. Prediction of *N*ε-acetylation on internal lysines implemented in Bayesian Discriminant Method. Biochem Biophys Res Commun. 2006;350(4):818-824. doi:10.1016/j.bbrc.2006.08.199.

10. Gnad F, Ren S, Choudhary C, Cox J, Mann M. Predicting post-translational lysine acetylation using support vector machines. Bioinformatics. 2010;26(13):1666-1668. doi:10.1093/bioinformatics/btq260.

11. Shao J, Xu D, Hu L, Kwan YW, Wang Y, Kong X, Ngai SM. Systematic analysis of human lysine acetylation proteins and accurate prediction of human lysine acetylation through bi-relative adapted binomial score Bayes feature representation. Mol BioSyst. 2012;8(11):2964-2973. doi:10.1039/c2mb25251a.

12. Shien DM, Lee TY, Chang WC, Hsu JB, Horng JT, Hsu PC, Wang TY, Huang HD. Incorporating structural characteristics for identification of protein methylation sites. J Comp Chem. 2009;ren30(9):1532-1543. doi:10.1002/jcc.21232.

13. Ren J, Wen L, Gao X, Jin C, Xue Y, Yao X. CSS-Palm 2.0: an updated software for palmitoylation sites prediction. Protein Eng Des Sel. 2008;21(11):639-644. doi:10.1093/protein/gzn039.

14. Blom N, Gammeltoft S, Brunak S. Sequence and structure-based prediction of eukaryotic protein phosphorylation sites. J Mol Biol. 1999;294(5):1351-1362. doi:10.1006/jmbi.1999.3310.

15. Ingrell CR, Miller ML, Jensen ON, Blom N. NetPhosYeast: prediction of protein phosphorylation sites in yeast. Bioinformatics. 2007;23(7):895-897. doi:10.1093/bioinformatics/btm020.

16. Gnad F, Gunawardena J, Mann M. PHOSIDA 2011: the posttranslational modification database. Nucleic Acids Res. 2011;39(Database issue):D253-260. doi:10.1093/nar/gkq1159.

17. Zhao Q, Xie Y, Zheng Y, Jiang S, Liu W, Mu W, Liu Z, Zhao Y, Xue Y, Ren J. GPS-SUMO: a tool for the prediction of sumoylation sites and SUMO-interaction motifs. Nucleic Acids Res. 2014;42(Web Server issue):W325-330. doi:10.1093/nar/gku383.

18. Radivojac P, Vacic V, Haynes C, Cocklin RR, Mohan A, Heyen JW, Goebl MG, Iakoucheva LM. Identification, analysis, and prediction of protein ubiquitination sites. Proteins. 2010;78(2):365-380. doi:10.1002/prot.22555.

19. Nakai K, Horton P. PSORT: a program for detecting sorting signals in proteins and predicting their subcellular localization. Trends Biochem Sci. 1999;24(1):34-35. doi:10.1016/S0968-0004(98)01336-X.

20. la Cour T, Kiemer L, Molgaard A, Gupta R, Skriver K, Brunak S. Analysis and prediction of leucine-rich nuclear export signals. Protein Eng Des Sel. 2004;17(6):527-536. doi:10.1093/protein/gzh062.

21. Fu SC, Imai K, Horton P. Prediction of leucine-rich nuclear export signal containing proteins with NESsential. Nucleic Acids Res. 2011;39(16):e111. doi:10.1093/nar/gkr493.

22. Bannai H, Tamada Y, Maruyama O, Nakai K, Miyano S. Extensive feature detection of N-terminal protein sorting signals. Bioinformatics. 2002;18(2):298-305. doi:10.1093/bioinformatics/18.2.298.

23. Claros MG, Vincens P. Computational method to predict mitochondrially imported proteins and their targeting sequences. Eur J Biochem. 1996;241(3):779-786.

24. Guda C, Subramaniam S. pTARGET: a new method for predicting protein subcellular localization in eukaryotes. Bioinformatics. 2005;21(21):3963-3969. doi:10.1093/bioinformatics/bti650.

25. Sigrist CJ, de Castro E, Cerutti L, Cuche BA, Hulo N, Bridge A, Bougueleret L, Xenarios I. New and continuing developments at PROSITE. Nucleic Acids Res. 2013;41(Database issue):D344-347. doi:10.1093/nar/gks1067.

26. Zhuang J, Wang PY, Huang X, Chen X, Kang JG, Hwang PM. Mitochondrial disulfide relay mediates translocation of p53 and partitions its subcellular activity. Proc Natl Acad Sci USA. 2013;110(43):17356-17361. doi:10.1073/pnas.1310908110.

27. 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. doi:10.1093/nar/gkq108.

28. Lange A, McLane LM, Mills RE, Devine SE, Corbett AH. Expanding the definition of the classical bipartite nuclear localization signal. Traffic. 2010;11(3):311-323. doi:10.1111/j.1600-0854.2009.01028.x.

29. 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. 2002;1(10):847-854. doi:10.1016/S1568-7864(02)00145-3.

30. Vongsamphanh R, Fortier PK, Ramotar D. Pir1p mediates translocation of the yeast Apn1p endonuclease into the mitochondria to maintain genomic stability. Mol Cell Biol. 2001;21(5):1647-1655. doi:10.1128/MCB.21.5.1647-1655.2001.

31. Chatterjee A, Singh KK. Uracil-DNA glycosylase–deficient yeast exhibit a mitochondrial mutator phenotype. Nucleic Acids Res. 2001;29(24):4935-4940. doi:10.1093/nar/29.24.4935.

32. Nakabeppu Y. Regulation of intracellular localization of human MTH1, OGG1, and MYH proteins for repair of oxidative DNA damage. Prog Nucleic Acid Res Mol Biol. 2001;68:75-94.

33. Darshi M, Trinh KN, Murphy AN, Taylor SS. Targeting and import mechanism of coiled-coil helix coiled-coil helix domain-containing protein 3 (ChChd3) into the mitochondrial intermembrane space. J Biol Chem. 2012;287(47):39480-39491. doi:10.1074/jbc.M112.387696.

34. Lynes EM, Bui M, Yap MC, Benson MD, Schneider B, Ellgaard L, Berthiaume LG, Simmen T. Palmitoylated TMX and calnexin target to the mitochondria-associated membrane. EMBO J. 2012;31(2):457-470. doi:10.1038/emboj.2011.384.

35. Merrick BA, Dhungana S, Williams JG, Aloor JJ, Peddada S, Tomer KB, Fessler MB. Proteomic profiling of *S*-acylated macrophage proteins identifies a role for palmitoylation in mitochondrial targeting of phospholipid scramblase 3. Mol Cell Proteomics. 2011;10(10):M110.006007. doi:10.1074/mcp.M110.006007.

36. 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. doi:10.1128/mcb.01357-08.

37. Swartzlander DB. Regulation of base excision repair in response to genotoxic stress [Dissertation]. Atlanta, GA: Emory University; 2012.

Chapter 3
AUTOMATED QUANTIFICATION OF THE SUBCELLULAR LOCALIZATION OF MULTICOMPARTMENT PROTEINS VIA Q-SCAN

Nicholas C. Bauer[1,2], Anita H. Corbett[1,3,*] and Paul W. Doetsch[1,3,4,5,*]

[1] Department of Biochemistry, Emory University School of Medicine, Atlanta, GA 30322, USA
[2] Graduate Program in Biochemistry, Cell, and Developmental Biology, Emory University School of Medicine, Atlanta, GA 30322, USA
[3] Winship Cancer Institute, Emory University School of Medicine, Atlanta, GA 30322, USA
[4] Department of Radiation Oncology, Emory University School of Medicine, GA 30322, USA
[5] Department of Hematology and Medical Oncology, Emory University School of Medicine, Atlanta, GA 30322, USA

* Corresponding authors: Anita H. Corbett, acorbe2@emory.edu and Paul W. Doetsch, medpwd@emory.edu

Published in *Traffic*, Volume 14, Issue 12, Pages 1200–1208
doi: 10.1111/tra.12118

Abstract

In eukaryotic cells, proteins can occupy multiple intracellular compartments and even move between compartments to fulfill critical biological functions or respond to cellular signals. Examples include transcription factors that reside in the cytoplasm but are mobilized to the nucleus as well as dual-purpose DNA repair proteins that are charged with simultaneously maintaining the integrity of both the nuclear and mitochondrial genomes. While numerous methods exist to study protein localization and dynamics, automated methods to quantify the relative amounts of proteins that occupy multiple subcellular compartments have not been extensively developed. To address this need, we present a rapid, automated method termed quantitative subcellular compartmentalization analysis (Q-SCAn). To develop this method, we exploited the facile molecular biology of the budding yeast, *Saccharomyces cerevisiae*. Individual subcellular compartments are defined by a fluorescent marker protein and the intensity of a target GFP-tagged protein is then quantified within each compartment. To validate Q-SCAn, we analyzed relocalization of the transcription factor Yap1 following oxidative stress and then extended the approach to multicompartment localization by examining two DNA repair proteins critical for the base excision repair pathway, Ntg1 and Ung1. Our findings demonstrate the utility of Q-SCAn for quantitative analysis of the subcellular distribution of multicompartment proteins.

Introduction

In eukaryotic cells, subcellular compartments are functionally defined by the proteins present within them. While many proteins are localized to a single compartment, some

proteins are localized to multiple compartments either constitutively or in response to cellular signals. For example, many transcription factors are localized to the cytoplasm until a signal triggers nuclear import (1,2). Such mobilization of proteins is an efficient strategy to alter local function in response to stimuli because transport of a preexisting pool of protein is more rapid than the *de novo* synthesis and localization of a comparable amount of protein. Other proteins simultaneously play roles in multiple compartments such as some DNA repair proteins that are localized to both the nucleus and mitochondria to maintain the integrity of the genomes in each of these cellular compartments (3).

Despite the biological significance of localizing proteins to multiple subcellular compartments, tools for quantifying the relative subcellular distribution of multi-compartment proteins have not been extensively developed. Many protein localization studies employ manual scoring from microscopy data, relying on the heterogeneity of the cell population and human visual detection to provide a useful threshold (4-7). However, these implicit thresholds are subjective and the process can be very labor-intensive. In addition, manual methods are only semi-quantitative as they are based on qualitative data. True quantification can be achieved by manually tracing the boundaries of the compartments of interest and then quantifying pixels within each compartment, but the laborious nature of this type of analysis means the number of cells that can be analyzed is effectively limited. Colocalization analysis (8), which has advanced greatly over the last decade and is widely available in image analysis software, is more suited to addressing questions about whether proteins and markers are spatially linked rather than about the distribution of a protein among distinct compartments. Photobleaching (9) and photoactivation techniques can be employed to examine dynamics (10); however, these techniques require highly specialized experimental setups and are also limited to larger cells amenable to such techniques. Biochemical fractionation techniques can also provide quantifiable compartmentalization

information on a population of cells (4,11,12), but microscopy-based techniques are superior to fractionation because micrographs preserve the spatial relationships and yield information on the single-cell level, not just the population average.

The limitations of the above techniques form a critical impediment to analyzing the steady-state distribution of proteins localized to multiple compartments. Development of advanced, automatable techniques that provide unbiased quantification of protein localization on a per-cell basis is becoming an active area of research. We have developed an approach to quantifying protein distribution among multiple compartments, which we term quantitative subcellular compartmentalization analysis (Q-SCAn). This microscopy-based method uses brightfield DIC images to identify cells, relies on a set of fluorescent markers to define subcellular compartments, and provides information about the amount of a protein of interest, marked by a third fluorescent reporter, within the identified compartments. By comparing the fluorescence intensities for each compartment, a localization index is calculated for each cell, yielding a quantitative measure of protein localization. Furthermore, the distribution of these localization indices can be compared between different cell types, conditions and time points to address the regulation of protein localization.

Here we describe the development of Q-SCAn in *Saccharomyces cerevisiae* and demonstrate its utility in measuring the single-cell localization of proteins by following the oxidative stress–induced relocalization of the transcription factor Yap1 (13). Next, we extend the approach to multicompartment localization by examining the nucleomitochondrial base excision repair (BER) protein Ntg1 (14). Finally, we apply the method to evaluate the localization of another nucleomitochondrial BER protein, Ung1 (15), which has not been previously analyzed in any quantitative manner. Our analysis of Ung1 provides new biological information about mechanisms of localization of Ung1 and thus insight into regulation of the BER pathway, demonstrating the utility of Q-SCAn for such studies. This work presents

a novel method for quantifying the subcellular distribution of multicompartment proteins which can be immediately put to use and extended without specialized equipment or programming experience.

Materials and Methods

Yeast strains, media and growth conditions

Saccharomyces cerevisiae strains used in this study are listed in Table 3-1. *S. cerevisiae* cells were cultured at 30 °C in rich YPD medium (1% yeast extract, 2% peptone, 2% dextrose and 2% agar for plates) or synthetic defined drop-out media for selection (uracil$^-$/leucine$^-$) and/or transcription induction (methionine$^-$). Plasmids or integrated constructs were transformed into cells by a modified lithium acetate method (16).

Plasmid construction

Plasmids used in this study are listed in Table 3-1. Both a nuclear marker gene (NLS–tdTom) and a mitochondrial marker gene (MTS–mCer) were constructed and inserted into an *S. cerevisiae* integration plasmid (pRS305) (17). Each marker gene was expressed from a constitutive promoter (low-level *CYC1* (18) and high-level *TEF1* (18), respectively, chosen to equalize the brightness of the marker protein) and terminated by a generic transcription terminator from the *NUF2* gene (19). The marker genes themselves are composed of a localization sequence fused to a fluorescent protein (NLS–tdTom: bipartite SV40 nuclear localization signal (NLS) (20) and tdTomato (21); MTS–mCer: *Neurospora crassa* Su9 mitochondrial matrix targeting sequence (MTS) (22) and mCerulean (23)). The pRS305 *LEU2* gene contains a unique EcoRV site to linearize the plasmid and thus enhance integration efficiency. Localization of these reporters in *S. cerevisiae* was confirmed by co-staining with DAPI or MitoTracker Red CMXRos (Invitrogen) and performing colocalization analysis on the fluorescence micrographs (8,24). Mitochondrial and cytoplasmic GFP localization control plasmids were constructed by replacing the NLS of pNLS–GFP (20) with the Su9 MTS (22) or

the nuclear export sequence (NES) of mammalian PKI (25), respectively. The Ung1–GFP plasmid was constructed by inserting the *UNG1* gene (ORF + 914 bp upstream) amplified by PCR from wild type *S. cerevisiae* (FY86), an in-frame C-terminal GFP (26), and *NUF2* terminator into pRS426 (27). The predicted classical NLS (28) and MTS (29) of Ung1 were mutated (NLS1: K18A, R19A; NLS2: K40A, K41A; MTS: R5A, R6A) using the QuikChange II Site-Directed Mutagenesis Kit (Agilent). All constructs were verified by sequencing.

Table 3-1. Strains and plasmids used in this study.

Name	Genotype	Ref
FY86 (ACY193)	*MATa ura3-52 leu2Δ1 his3Δ200*	(30)
Q-SCAn (DSC569)	*MATa ura3-52 leu2Δ1 his3Δ200 NLS–tdTom MTS–mCer*	This study
pRS305	*LEU2*, Amp^R	(17)
pRS426	*2μ*, *URA3*, Amp^R	(17)
pNLS–tdTom (pD0464)	*pTEF1-BPSV40NLS–tdTomato*, *LEU2*, Amp^R	This study
pMTS–mCer (pD0465)	*pCYC1-Su9MTS–mCerulean*, *LEU2*, Amp^R	This study
pQ-SCAn (pD0466)	*pTEF1-BPSV40NLS–tdTomato*, *pCYC1-Su9MTS–mCerulean*, *LEU2*, Amp^R	This study
pNLS–GFP_2	*pMET17-BPSV40NLS–GFP-GFP*, *CEN*, *URA3*, Amp^R	(20)
pMTS–GFP (pD0467)	*pMET17-Su9MTS–GFP-GFP*, *CEN*, *URA3*, Amp^R	This study
pNES–GFP (pD0468)	*pMET17-NES–GFP-GFP*, *CEN*, *URA3*, Amp^R	This study
pYap1–GFP	*YAP1–GFP*, *CEN*, *URA3*, Amp^R	(7)
pNtg1–GFP	*NTG1–GFP*, *2μ*, *URA3*, Amp^R	(4)
pNtg1$_{NLS2}$–GFP	*ntg1$_{NLS2}$–GFP*, *2μ*, *URA3*, Amp^R	(5)
pNtg1$_{NLS1/2}$–GFP	*ntg1$_{NLS1/2}$–GFP*, *2μ*, *URA3*, Amp^R	(5)
pNtg1$_{MTS}$–GFP	*ntg1$_{MTS}$–GFP*, *2μ*, *URA3*, Amp^R	(5)
pUng1–GFP (pD0419)	*UNG1–GFP*, *2μ*, *URA3*, Amp^R	This study
pUng1$_{NLS1}$–GFP (pD0469)	*ung1$_{NLS1}$–GFP*, *2μ*, *URA3*, Amp^R	This study
pUng1$_{NLS2}$–GFP (pD0470)	*ung1$_{NLS2}$–GFP*, *2μ*, *URA3*, Amp^R	This study
pUng1$_{NLS1/2}$–GFP (pD0471)	*ung1$_{NLS1/2}$–GFP*, *2μ*, *URA3*, Amp^R	This study
pUng1$_{MTS}$–GFP (pD0472)	*ung1$_{MTS}$–GFP*, *2μ*, *URA3*, Amp^R	This study

Confocal microscopy

All fluorescence micrographs were obtained using a Zeiss LSM 510 confocal fluorescence microscope using a Plan-Apochromat 100x/1.4 NA Oil DIC objective. Each fluorescence channel was imaged sequentially, with the transmission brightfield DIC collected simultaneously with mCerulean. See Table 3-2 for parameters used. Fluorescence intensities from a 90.0 μm square area were encoded into a 1024 × 1024 pixel, 12-bit image file for near-optimal Nyquist sampling (31) and enhanced discrimination of fluorescence intensities, respectively. The laser dwell time was 1.60 μs/pixel. All pinholes were set to 168.00 μm

(1 Airy unit for 543-nm laser line), resulting in an optical slice of 0.8 μm. Expression of each fluorescent protein alone was used to control for crosstalk between fluorescence channels. Greater than 100 cells were imaged for each condition. The median projections of 10 uniform brightfield images and 10 dark images per channel were used to correct images for shading and dark current, respectively.

Table 3-2. Confocal microscopy parameters.

	tdTomato	GFP	mCerulean	DIC
Excitation laser line (nm)	543 nm	488 nm	405 nm	405 nm
Laser power (mW)	1.0 mW	15.0 mW	25.0 mW	25.0 mW
Laser transmission (%)	10%	30%	65%	65%
Detected λ range (nm)	>560 nm	505–545 nm	420–490 nm	—
Detector gain (V)	750 V	725 V	750 V	260 V[a]
Amplifier offset	−0.01	−0.01	−0.03	0.00

[a] Sometimes had to be adjusted slightly to maintain an image solidly within the dynamic range.

Yap1-GFP hydrogen peroxide treatment

Overnight late-log phase (approximately 1×10^8 cells/mL) cultures of *S. cerevisiae* cells expressing the Q-SCAn reporter and Yap1–GFP were washed twice with H_2O, resuspended in 1 mL H_2O, counted via hemocytometer, and adjusted to a density of $2–4 \times 10^7$ cells/mL. Cells were treated with 5 mM H_2O_2 and imaged by confocal microscopy both before treatment and within three time windows following the start of treatment: 1–10 min, 30–45 min or 60–80 min of start of treatment. Cells were visually scored for nuclear only, cytoplasmic only, or nuclear and cytoplasmic Yap1–GFP localization.

Data analysis

Colocalization analysis of the marker proteins was conducted using the commercial software Volocity 5.6.2 (PerkinElmer). ImageJ 1.46r was used to export channels from the raw images and to compute the median shading and dark current correction images (32). CellProfiler 2.0 (Developer build) with custom plugins was used to process the images, identify *S. cerevisiae* cells and compartments, and measure fluorescence intensities (33). CellProfiler Analyst 2.0 (r11710) was used to develop filtering thresholds for the

algorithm (34). Statistical analysis was performed with Stata 11.2. The localization index for a single cell is the mean nuclear GFP intensity divided by the sum of the mean nuclear and mean mitochondrial GFP intensities, producing a value in the range of 0 (mitochondrial) to 1 (nuclear). Cells whose sums of intensities were below the noise threshold were excluded. Shifts in localization were assessed by the Kruskal–Wallis test and the post-hoc multiple comparisons test with correction for ties (35), with $\alpha = 0.05$.

Results

Automated quantification of subcellular protein localization: Q-SCAn

To address a critical gap in the techniques available to rapidly and reproducibly provide quantitative information about protein distribution among multiple cellular compartments, we have developed a method termed Q-SCAn. To develop this system, we exploited the budding yeast *S. cerevisiae* and focused on developing a system that could be used to quantify protein levels in the nucleus and mitochondria.

We designed a dual reporter for *S. cerevisiae* that could be integrated into the genome to create a reporter yeast strain with constitutively labeled nuclei and mitochondria. The reporter encodes spectrally distinct fluorescent proteins targeted to the nucleus and mitochondria via well-characterized targeting signals (Figure 3-1A). The nuclear reporter protein comprises the strong, artificial SV40 bipartite nuclear localization signal (NLS) (20) fused to the tandem dimer red fluorescent protein tdTomato (21). This reporter protein is expressed from the low-level *CYC1* constitutive promoter (18) and terminated by the *NUF2* terminator (19). The mitochondrial reporter protein is composed of the highly efficient *Neurospora crassa* Su9 mitochondrial matrix targeting signal (MTS) (22) fused to the cyan fluorescent protein mCerulean (23). This reporter protein is expressed from the high-level *TEF1* constitutive promoter (18) due to the relative dimness of mCerulean as compared to tdTomato, and is also terminated by the *NUF2* terminator (19) (Figure 3-1A). These reporters were

integrated into the *S. cerevisiae LEU2* locus to create a constitutively labeled yeast reporter strain.

To validate the localization of these fluorescent reporter proteins to the targeted compartments, we co-stained cells expressing either NLS–tdTomato or MTS–mCerulean with DAPI to label chromatin and mitochondrial nucleoids or MitoTracker Red to label mitochondrial proteins, respectively. As shown in the micrographs in Figure 3-1B and validated by colocalization analysis, the marker proteins localized to their intended subcellular compartments consistently and with high specificity, a necessary requirement for computer-assisted image analysis, whereas these standard dyes comparatively displayed highly variable and non-specific staining.

To carry out Q-SCAn, fluorescent micrographs of cells expressing the marker proteins were obtained (Figure 3-1C). The Q-SCAn method, which is implemented in the open source CellProfiler software package (the CellProfiler pipelines necessary to run Q-SCAn are available for download at http://www.biochem.emory

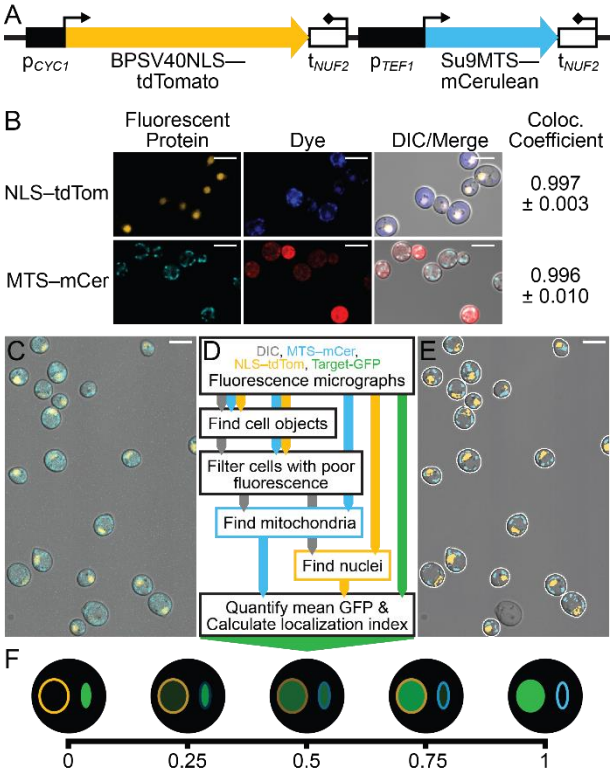

Figure 3-1. Quantitative subcellular compartmentalization analysis (Q-SCAn). A) Schematic of the tandem nuclear and mitochondrial marker protein construct, which is integrated into the *Saccharomyces cerevisiae LEU2* locus as described in Results. B) The Q-SCAn marker proteins specifically and uniformly localize to their respective compartments. The nuclear and mitochondrial fluorescent marker proteins (NLS–tdTomato, MTS–mCerulean) were expressed independently in S. cerevisiae cells and co-stained with DAPI (dye, upper panel) or MitoTracker (dye, lower panel), respectively, and subjected to colocalization analysis to generate a colocalization coefficient (Coloc. Coefficient: mean ± SD over five images). C) Example merged image of cells co-expressing both fluorescent marker proteins. Scale bar is 5 µm. D) Diagram of the flow of information in Q-SCAn. E) Result of the compartmentalization of the sample image in (C). Identified cells are outlined in white, identified nuclei are marked in orange, and identified mitochondria in cyan. Overlapping portions of compartments are excluded, and cells that do not have both markers expressed and visible in the optical slice are excluded. Scale bar is 5 µm. F) Schematic of the mito–nuclear localization index. The index is generated by dividing the mean nuclear GFP intensity by the sum of the mean nuclear and mean mitochondrial GFP intensities. A value of 0 indicates that all of the signal is in the mitochondria, while a value of 1 indicates that all of the signal is in the nucleus, as depicted by the cell diagrams (nucleus outlined in orange, mitochondria in blue).

.edu/doetsch/qscan.html [and included as a supplemental file identified in the Appendix]), was then applied to the images collected as depicted in Figure 3-1D to: (i) identify the location of cells from the DIC image utilizing information from the marker channels; (ii) filter out cells that have poor marker fluorescence that would prevent robust identification of compartments; (iii) identify, within each cell, the location of nuclei and mitochondria using information from the tdTomato and mCerulean channels, respectively; (iv) subtract the nuclei and mitochondria from the cell to define the cytoplasm; (v) remove overlapping parts of nuclei and mitochondria; and (vi) quantify mean GFP intensity within each cell compartment (Figure 3-1D). The result of the compartmentalization is shown in Figure 3-1E. A measure indicating the relative distribution of the protein, a localization index, is then constructed, where the mean nuclear GFP intensity is divided by the sum of the mean nuclear and mean mitochondrial GFP intensities. This analysis results in a measure ranging from 0 (100% mitochondrial) to 1 (100% nuclear) (Figure 3-1F). (See Supporting Information for technical details of the analysis.)

To calibrate the system, we constructed and expressed localization sequence–fused GFP proteins targeted to the nucleus, NLS–GFP_2 (NucGFP) (20) and mitochondria, MTS–GFP_2 (MitoGFP), as well as one designed to remain in the cytoplasm, NES–GFP_2 (CytoGFP) (Figure 3-2A). The thresholds and corrections used in the algorithm were tuned such that the nuclear and mitochondrial GFP protein–

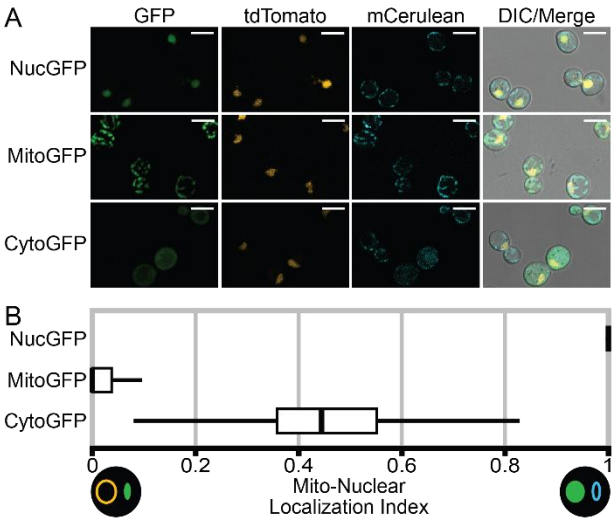

Figure 3-2. Q-SCAn calibrated to detect distinct nuclear and mitochondrial localization. *Saccharomyces cerevisiae* cells co-expressing integrated NLS–tdTomato and MTS–mCerulean, and nuclear, mitochondrial, or cytoplasmic GFP (NucGFP, MitoGFP, CytoGFP) were imaged. Greater than 200 cells were analyzed per experiment. A) Representative images of cells expressing a localized GFP. Scale bar is 5 µm. B) Quantification of the distribution of localization controls presented as mito–nuclear localization index (0 = mitochondrial to 1 = nuclear). Vertical black line indicates the sample median, white box the interquartile range, and horizontal black line the adjacent range.

expressing cells resulted in an exclusively nuclear or mitochondrial localization index, respectively (Figure 3-2B), though mitochondrial localization is somewhat less sharply defined than nuclear localization. A protein which does not localize to either compartment, cytoplasmic GFP, has a neutral and diffuse localization index distribution. These data demonstrate that the nucleomitochondrial localization index is valid for quantifying proteins with distributions in between these two extremes.

Q-SCAn used to analyze nuclear localization

To assess the practical utility of Q-SCAn to examine a classical signaling paradigm for multicompartment protein localization, the movement of a transcription factor from the cytoplasm to the nucleus (1,2), we analyzed the localization of the *S. cerevisiae* transcription factor, Yap1 (13). Under normal cellular conditions, Yap1 shuttles between the nucleus and the cytoplasm but shows diffuse, steady-state localization primarily to the cytoplasm, as the rate of nuclear export exceeds the rate of nuclear import (36). However, upon exposure to oxidative stress, although steady-state levels of Yap1 do not change significantly (37), Yap1 nuclear export is blocked, resulting in rapid nuclear accumulation (7,36).

We expressed Yap1–GFP in *S. cerevisiae* expressing the nuclear marker protein. As expected, in the absence of any oxidative stress (Figure 3-3A, no treatment), Yap1–GFP was diffusely localized throughout each cell (37). Cultures were then treated with a mild dose of hydrogen peroxide and samples were imaged within three windows of time following exposure (Figure 3-3A). On applying Q-SCAn using a nucleocytoplasmic localization index (cytoplasmic = 0, nuclear = 1), the rapid mobilization of Yap1 into the nucleus was readily apparent (Figure 3-3B). The time-course and extent of relocalization corresponds well with previously published work (7) as well as with manual scoring of the same data (Figure 3-3C). Q-SCAn also allowed us to obtain new biological information about the heterogeneity of the distribution of Yap1 in the population. Initially the variance in Yap1 localization is very wide,

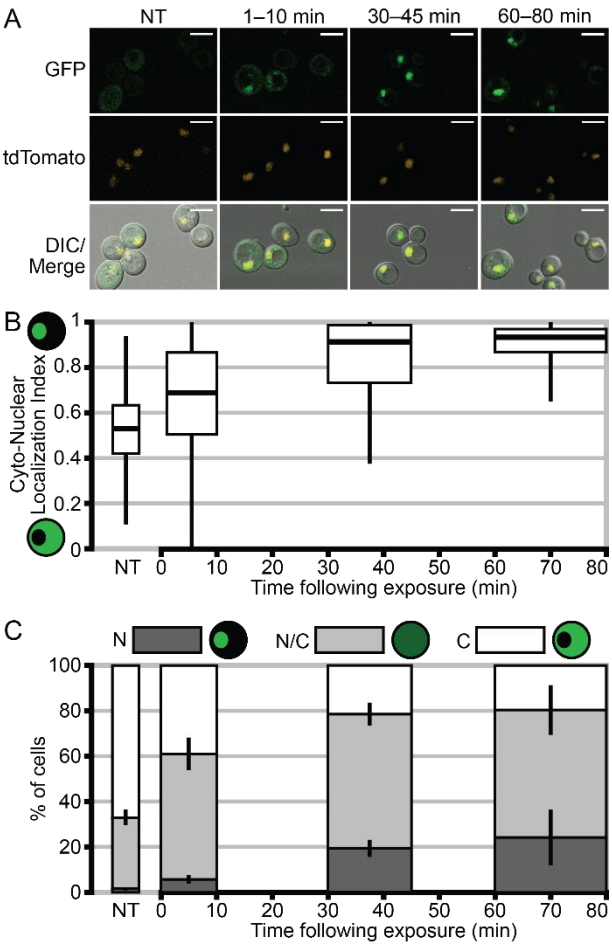

Figure 3-3. Yap1–GFP redistributes to the nucleus under oxidative stress, as revealed by Q-SCAn. *Saccharomyces cerevisiae* cells co-expressing integrated NLS–tdTomato and MTS–mCerulean and Yap1–GFP were exposed to 5 mM hydrogen peroxide and imaged within three windows of time after exposure, 1–10 min, 30–45 min and 60–80 min. Across three independent replicates, >200 cells were analyzed per time window. A) Representative images of cells expressing Yap1–GFP before treatment (no treatment: NT) and within each time window following exposure. Scale bar is 5 µm. B) Quantification of the distribution of Yap1–GFP (0 = cytoplasmic to 1 = nuclear) within each time window following exposure. Horizontal black line indicates the sample median, white box the interquartile range, and vertical black line the adjacent range. C) Percentage of cells visually scored for cytoplasmic (C, white), nuclear and cytoplasmic (N/C, light gray), or nuclear (N, dark gray) Yap1–GFP localization within each time window following exposure. Boundaries are mean ± SEM.

but the distribution quickly tightens during oxidative stress as nuclear localization increases. The comparison with manual scoring also allowed direct comparison of the efficiency benefits of using Q-SCAn. Analyzing the same dataset (from end of image collection to start of data analysis) of over 200 images manually required approximately 6 h of working time while setting up and beginning the Q-SCAn analysis for the same number of cells required less than 30 min, with minimal interaction over the course of the analysis. With a small program we developed to automate the process, even this minim interaction is eliminated (Supporting Information).

Q-SCAn used to analyze the dual-localized DNA repair protein Ntg1

To extend the utility of Q-SCAn beyond localization to a single compartment, we next analyzed a system where a protein is localized to two specific compartments. For this purpose, we examined a DNA repair protein critical for the BER pathway, the yeast endonuclease III–like DNA *N*-glycosylase/AP lyase, Ntg1, which localizes to both the nucleus and mitochondria (14) to differing extents depending upon the DNA damage status of each genome (4).

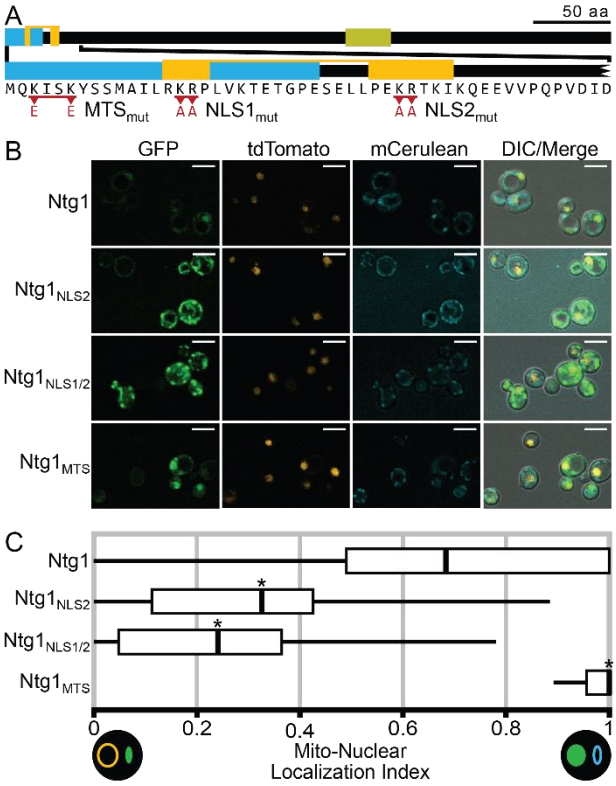

Figure 3-4. Localization of Ntg1 analyzed by Q-SCAn. *Saccharomyces cerevisiae* cells co-expressing integrated NLS–tdTomato and MTS–mCerulean, and an Ntg1–GFP variant on a 2µ plasmid were imaged. A) Schematic of Ntg1 showing nuclear localization signals (NLS, orange), mitochondrial matrix targeting signal (MTS, blue) and catalytic site (yellow), indicating the amino acid changes introduced within each targeting signal. B) Representative images of cells expressing Ntg1–GFP variants. Scale bar is 5 µm. C) Quantification of the distribution of Ntg1–GFP fusion proteins (0 = mitochondria to 1 = nucleus). Vertical black line indicates the sample median, white box the interquartile range, and horizontal black line the adjacent range. Asterisk denotes distributions determined to be significantly different from wild type Ntg1 ($p < 0.05$). Across three independent replicates, >350 cells were analyzed per variant analyzed.

To assess the utility of Q-SCAn for detecting differences in the distribution of a protein between the nucleus and mitochondria, we exploited previously characterized mutants of Ntg1 with altered targeting to each compartment (5) as indicated in Figure 4A and compared their localization with wild type Ntg1. We expressed these Ntg1 localization sequence mutants (Figure 3-4A) in *S. cerevisiae* containing the Q-SCAn nuclear and mitochondrial reporter proteins. These cells were imaged and analyzed via Q-SCAn. Micrographs of representative fields of cells are shown in Figure 3-4B. Q-SCAn revealed that wild type Ntg1 is relatively evenly distributed between the nucleus and mitochondria, with most cells containing slightly more nuclear Ntg1 relative to mitochondrial Ntg1. Disrupting the MTS resulted in nearly complete loss of mitochondrial localization (Figure 3-4C). Altering both segments of the bipartite NLS within Ntg1 (NLS1 and NLS2) resulted in loss of nearly all nuclear localization of Ntg1 (Figure 3-4C). Consistent with a previous study that defined the contributions of NLS1 and NLS2 to Ntg1 localization (5), disruption of NLS2 also resulted in significant loss of nuclear localization, but, notably, Q-SCAn revealed that this Ntg1 variant retained some nuclear localization as compared to the NLS1/2 double mutant. These data

demonstrate that Q-SCAn can be readily used to experimentally quantify the localization of nucleomitochondrial proteins with different distributions between cellular compartments, revealing even modest changes such as the difference between altering both NLS1 and NLS2 vs. NLS2 alone.

Application of Q-SCAn to define targeting signals within the DNA repair protein Ung1

Finally, to assess the utility of Q-SCAn to obtain novel biological insights, we analyzed another DNA repair protein, the *S. cerevisiae* uracil–DNA glycosylase, Ung1. Similar to Ntg1, Ung1 localizes to both the nucleus and mitochondria (38); however, neither the distribution of the protein between the two compartments has been analyzed, nor have the intracellular targeting signals been experimentally defined. The predicted nuclear and mitochondrial localization signal sequences within Ung1 were identified via the NUCDISC component of PSORT II (28) and iPSORT (29), respectively. As shown in Figure 3-5A, these predicted targeting motifs within Ung1 are arrayed in an overlapping manner similar to the organization of these sequences in Ntg1 (see Figure 3-4A).

Localization of wild type Ung1

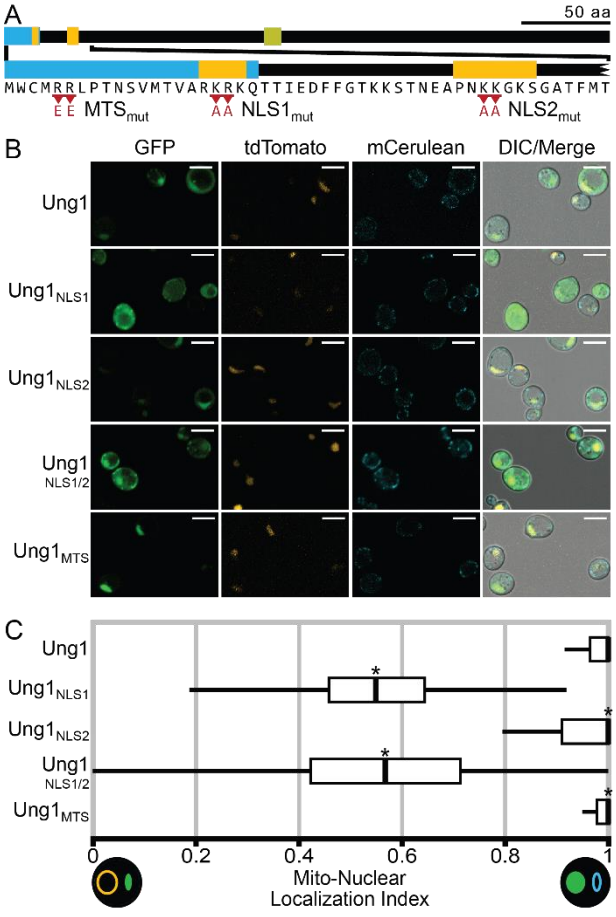

Figure 3-5. Functional analysis of Ung1 targeting signals revealed by Q-SCAn. *Saccharomyces cerevisiae* cells co-expressing integrated NLS–tdTomato and MTS–mCerulean, and an Ung1–GFP variant on a 2μ plasmid were imaged. A) Schematic of Ung1 showing nuclear localization signals (NLS, orange, identified via PSORT II (28)), mitochondrial matrix targeting signal (MTS, blue, identified via iPSORT (29)) and catalytic site (yellow), indicating the amino acid changes introduced within each targeting signal. B) Representative images of cells expressing Ung1–GFP variants. Scale bar is 5 μm. C) Quantification of the distribution of Ung1–GFP fusion proteins (0 = mitochondria to 1 = nucleus). Vertical black line indicates the sample median, white box the interquartile range, and horizontal black line the adjacent range. Asterisk denotes distributions determined to be significantly different from wild type Ntg1 ($p < 0.05$). Across three independent replicates, >250 cells were analyzed per variant analyzed.

reveals that this BER protein has a distinct intracellular distribution compared to Ntg1 (Figure 3-5B,C) as Ung1 displays greater nuclear concentration than Ntg1. This assessment can be made by comparing the distribution of localization indices for the two proteins (median, Ung1: 1.0 vs. Ntg1: 0.68). To extend the analysis and gain insight into the functional targeting signals in Ung1, we disrupted the three predicted localization sequences (MTS, NLS1 and NLS2) within Ung1 by site-directed mutagenesis (Figure 3-5A). We compared the localization of these Ung1 variants to wild type Ung1 using Q-SCAn. Micrographs of representative fields are shown in Figure 3-5B. These results reveal that NLS1 is necessary and sufficient for nuclear localization of Ung1 (Figure 3-5C), defining the functional NLS within Ung1. NLS2 may have a minor contribution to nuclear localization, but the distributions of the NLS1 mutant and NLS1/2 double mutant are statistically identical. The Ung1 variant with the altered MTS lost all mitochondrial localization as would be predicted if this motif is a functional targeting signal (Figure 3-5C). These data demonstrate that Q-SCAn can provide novel information about the extent of intermediate shifts in protein localization that may not be readily apparent by visual inspection.

Discussion

The goal of the work presented here was to develop a rapid and quantitative method that could be used to assess the relative amount of a protein present in multiple compartments. Our results demonstrate the utility of Q-SCAn for such studies. Given the automated and rapid data analysis feasible with Q-SCAn, biological questions that examine changes in the localization of a dual-compartment protein can be readily addressed quantitatively. For example, a change in compartment-specific localization in response to a cellular signal over time could be analyzed as presented here for Yap1. In addition, quantitative information regarding the contribution of intracellular targeting signals to specific cellular compartments, such as described here for the DNA repair proteins Ntg1 and

Ung1, can be readily collected and analyzed. Thus, Q-SCAn expands the tools available to address questions of protein distribution in a quantitative manner. Furthermore, Q-SCAn has been implemented in such a way as to ensure that the method can be readily employed by those that wish to analyze any protein of interest.

This study also demonstrates how Q-SCAn can be employed to gather new biological information about a protein of interest. While accumulation of Yap1 within the nucleus upon oxidative stress is a well-established biological response (7,13,36), our analysis reveals information about the extent of nuclear localization with time. Furthermore, we employed Q-SCAn to define the functional localization sequences of the DNA repair protein Ung1, which is localized both to the nucleus and mitochondria (38). Our analysis reveals that the NLS1 sequence constitutes the functional classical NLS in Ung1. However, altering both predicted classical NLS motifs did not eliminate all nuclear localization, indicating that Ung1 likely exploits a non-classical nuclear import pathway to ensure access to the nucleus. Such a mechanism using dual pathways to access the nucleus has also been observed for Ntg1 (5). Thus, in addition to facilitating experiments through rapid and automated data analysis, the quantitative results obtained using Q-SCAn provide novel information that is not revealed through conventional qualitative scoring methods.

A major strength of Q-SCAn is that the approach has been developed so that it can be easily applied by a user without any need for highly specialized equipment or software. Q-SCAn employs the open-source CellProfiler software package (33). This package is user-friendly and highly modular, allowing for any aspect of the Q-SCAn algorithm to be modified for different circumstances or to couple with other analyses. The CellProfiler package is also under active development and improvement at the Broad Institute, and technical assistance is readily available if required (www.cellprofiler.org).

There are several points that must be taken into consideration when implementing

Q-SCAn. First, a marker protein for the compartment or compartments of interest will need to be designed. The fluorescence spectrum of fluorescent markers should not overlap with the fluorescence spectrum of the analyzed protein. Here we selected tdTomato and mCerulean coupled with GFP based on the spectral properties of these fluorescent proteins (21). Another consideration is that, as described here, the localization examined is not that of the endogenous protein, but rather that of a fluorescently tagged protein. Such tagged proteins are commonly used to obtain information about protein localization (39,40) but both the presence of the tag and the level of the tagged-protein relative to endogenous protein need to be taken into account when interpreting results. Fluorescence noise should also be monitored. Although several noise-reduction methods are implemented in Q-SCAn, some noise still remains. As a result, as the target protein level approaches the noise level, the localization index will approach equality. Finally, care must be taken to ensure that cells in the imaged fields are well-separated. While the cell-finding algorithm can readily separate cells that touch, there can be difficulty in separating cell clusters, which could inadvertently be treated as a single cell.

Development of methods to analyze protein distribution is an active area of research. A recent publication employed a dye to mark the *S. cerevisiae* plasma membrane to quantify recruitment of YFP-Ste5 to the plasma membrane upon pheromone signaling (41). Using this approach, the authors were able to calculate both the pheromone dose–response for Ste5 recruitment to the membrane and the dissociation constant for Ste5 from the membrane. While elements of this quantitative approach are similar to Q-SCAn, there are some significant differences. These authors employed a dye to mark the target compartment, but such dyes are not readily available to specifically mark all cellular compartments. In addition, the method was implemented using a special-purpose software, Cell-ID (42). While this program, like CellProfiler, is open-source, significant expertise would be required to use or

modify the program. Another key distinction from Q-SCAn is that the authors employed defocused brightfield images to identify cells. While this approach facilitates cell-finding, obtaining these images on a non-automated microscope would be laborious and employing such images result in a loss of information about internal cellular morphology and cell fitness. Given the different biological questions addressed by the two approaches, recruitment of a protein to the membrane (3) and movement of proteins between intracellular compartments (this report), these two methods complement one another to obtain quantitative biological information.

As a general method, there are numerous ways that Q-SCAn could be developed and extended, which is facilitated by implementation in CellProfiler. For example, the cell-finding algorithm could be replaced to identify mammalian cells, and the compartment-finding algorithms could be adapted to different types of markers including immunofluorescence. A third fluorescent protein from the deep red spectral range such as mPlum (43) could be introduced to mark a third compartment of interest. More than three fluorescent protein compartment markers could be employed if spectral imaging/linear unmixing is used to separate the signals. The localization distribution information could also be coupled to other information, such as cell size, compartment morphology or measures of the protein distribution within a compartment. With modifications, Q-SCAn could also be adapted to analyze localization within three-dimensional datasets (44). In addition to analyzing single images, Q-SCAn could further be employed to extract localization data from frames of a video, allowing quantitative analysis of the dynamics of protein relocalization within single cells. Q-SCAn could also have applications in diagnostics and therapeutics, as mislocalization of proteins is associated with disease processes including cancer, autoimmune disorders and degenerative disorders.

This study reports the development and utilization of Q-SCAn as a facile, quantitative

analytical tool for providing broader and more detailed analysis of the localization of multicompartment proteins when compared to the current approaches available.

Acknowledgments

We would like to thank the Emory Custom Cloning Core Facility and the Emory University Integrated Cellular Imaging Core of the Winship Cancer Institute (National Cancer Institute Cancer Center Support Grant P30CA138292), especially Debbie Martinson and Dr. Adam Marcus. We would also like to thank the members of the Doetsch and Corbett laboratories, especially Drs. Dan Swartzlander and Natalya Degtyareva, for their constructive feedback in designing these experiments and in the preparation of this manuscript. Finally, we thank the Imaging Platform at the Broad Institute for help with CellProfiler. This work was supported by the National Institutes of Health [grant numbers P01ES011163 to P. W. D., F31CA168272 to N. C. B.]. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Supporting Information

Preparing the method and determining thresholds

1. Select fluorescence probes that selectively mark the compartments that the protein of interest localizes to. This could be a set of fluorescent proteins fused to a localization sequence, antibodies to proteins for immunofluorescence, or some dyes.
2. Select a third marked protein that localizes exclusively to one compartment or the other. This could be a third fluorescent protein fused to a localization sequence or another immunofluorescence antibody, but the detection method should be the same as for the protein of interest.
3. Collect images of cells with each marker alone and combined, and also with the third marker for the protein to be analyzed. (1, 2, 3, 1+2, 1+2+3)
4. For the compartment markers alone, determine if there is crosstalk between the channels for each compartment by doing a linear regression analysis to compare the pixel intensities for the cells with marker 1 or 2 between channel 1 (or 2) and channels 2 (or 1) and 3. Low-level crosstalk can be ignored and dealt with via the compartment-finding

parameters. High-level cross-talk should be corrected by linear unmixing prior to compartment-finding.

5. Also measure any cross-talk that exists between marker 1 or 2 and channel 3 by the same method. This signal should be removed by linear unmixing before the quantification of channel 3.
6. If one or more of the markers has low-level diffuse signal which is used to assist cell-finding, adjust those parameters.
7. Applying the cross-talk corrections above, adjust the compartment-finding parameters by running the 1+2+3 compartment-control images through the analysis. The optimal parameters will be those that minimize signal detected in one compartment from the test marker localized to the second compartment.
8. Use CellProfiler Analyst to determine filter rules that will allow removal of cells with extremely low and/or diffuse marker fluorescence.
9. Using these parameters, analyze the 1+2 cells to measure the mean channel 3 intensity in each compartment of each cell. This represents autofluorescence and noise. The mean value for each compartment is used to correct for this signal by subtracting them from the mean channel 3 intensities.
10. Applying the autofluorescence/noise correction, analyze the 1+2+3 compartment-control cells for the amount of stray signal detected in the unmarked compartment, which can be due to optics imperfections, marker variation, cell/compartment movement during image collection, or conservative compartment-finding parameters. Perform a linear regression across cells for each compartment to determine how much signal leaks into the other. This will highly vary from cell to cell, so empirically we've found that using half the slope to do linear unmixing between compartments strikes a good balance between losing too much signal and accounting for stray signal.

11. Apply these final parameters to an analysis of the 1+2+3 compartment control cells. Assuming a localization index of compartment A / (compartment A + B), the control for compartment A localization should have most cells at 1.0, while the control for compartment B localization should have most cells at 0.0. Determine a low-intensity threshold for the denominator to remove cells from the analysis that primarily represent noise.
12. The method is ready to apply to your protein of interest.

Running the Q-SCAn method

1. Perform background correction of collected images using the median of 10 dark current images for each channel and 10 empty brightfield images for shading correction. (See Figure 3-S1)
2. Identify cells from the corrected brightfield images, using nonspecific fluorescence information from the marker proteins to assist. (See Figure 3-S2)

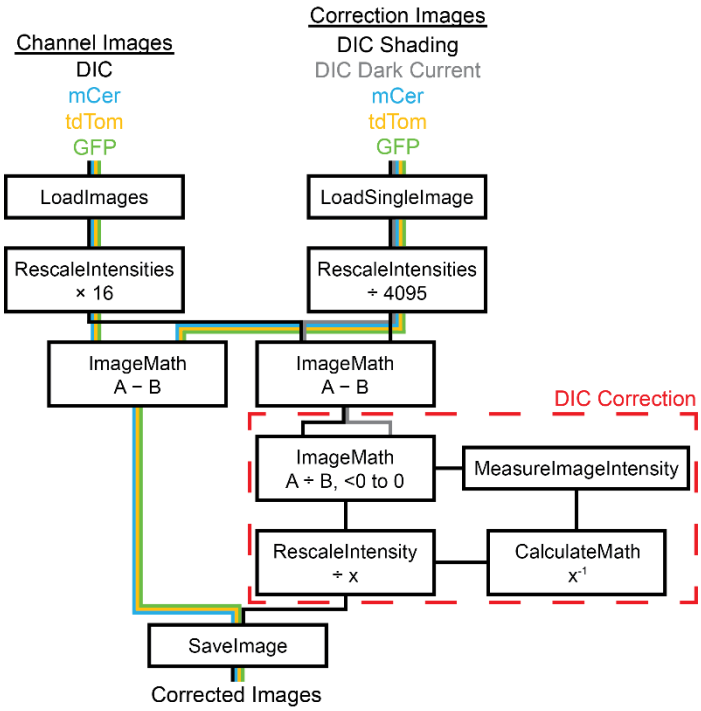

Figure 3-S1. Diagram of CellProfiler Pipeline 0 – Background Correction. Each colored line represents flow (top to bottom) of information through the various modules from an input image to the corrected output images.

3. Find the non-cell-containing background by finding the inverse of an expanded version of the cells.
4. Measure the mean fluorescence intensity in the background of the image for each marker, and subtract 5 times this value from the entire image. This helps remove lingering noise that can interfere with compartment-finding. (See Figure 3-S3)

5. Measure the fluorescence intensity within the cells for each marker and calculate the mean across the entire image set. This value is then used to normalize the mean marker

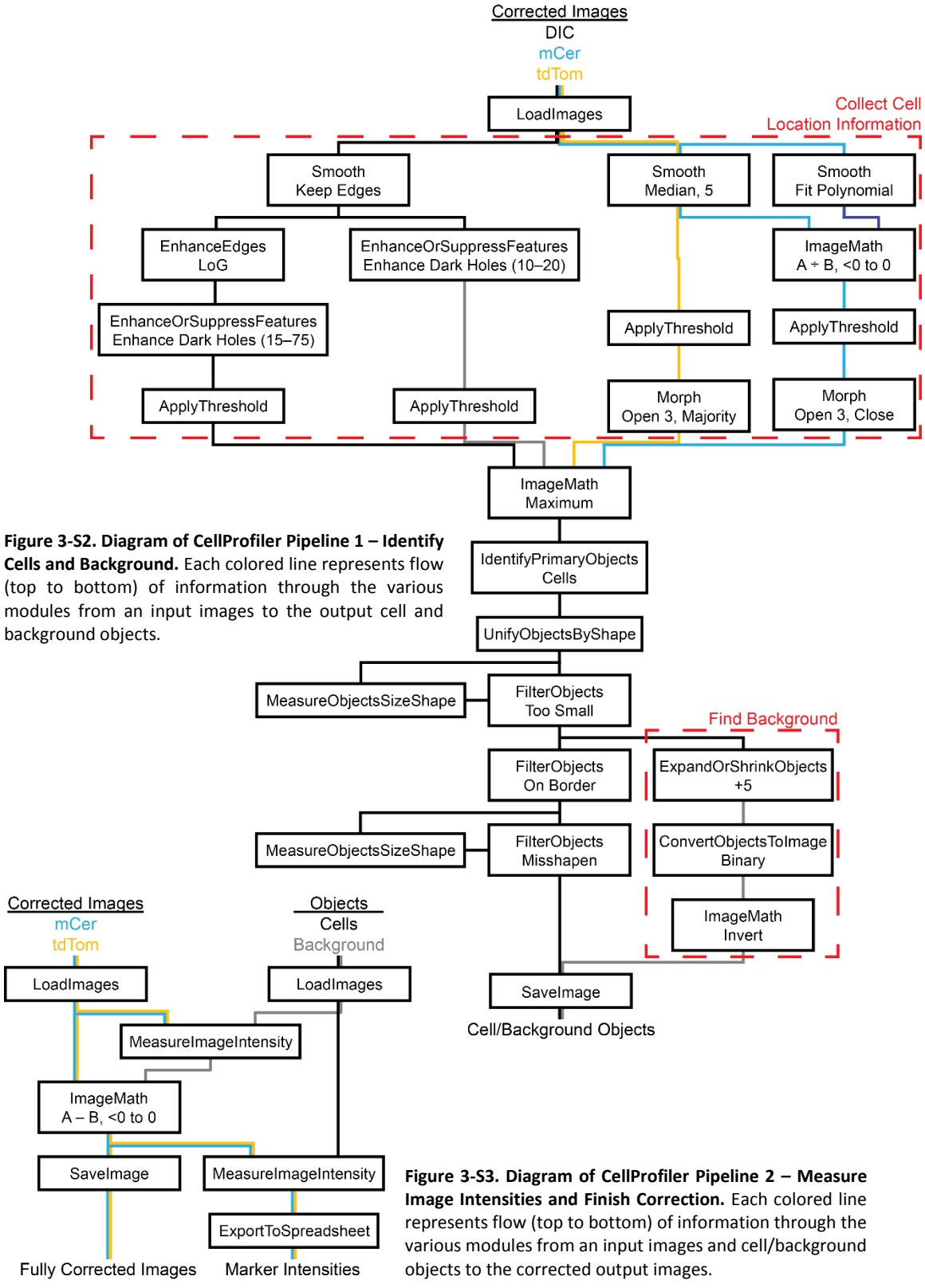

Figure 3-S2. Diagram of CellProfiler Pipeline 1 – Identify Cells and Background. Each colored line represents flow (top to bottom) of information through the various modules from an input images to the output cell and background objects.

Figure 3-S3. Diagram of CellProfiler Pipeline 2 – Measure Image Intensities and Finish Correction. Each colored line represents flow (top to bottom) of information through the various modules from an input images and cell/background objects to the corrected output images.

fluorescence intensities to a value of 0.01, so that the compartment thresholding is independent of experimental variability in marker fluorescence.

6. Filter out cells that have very low or diffuse marker fluorescence. These rules are determined empirically as described above. (See Figure 3-S4)
7. Within each cell, use thresholding to find the compartments. These thresholds are determined empirically as described above.
8. Filter out cells that have extremely large compartments.
9. Temporarily expand each compartment slightly and subtract the expanded compartment from the normal

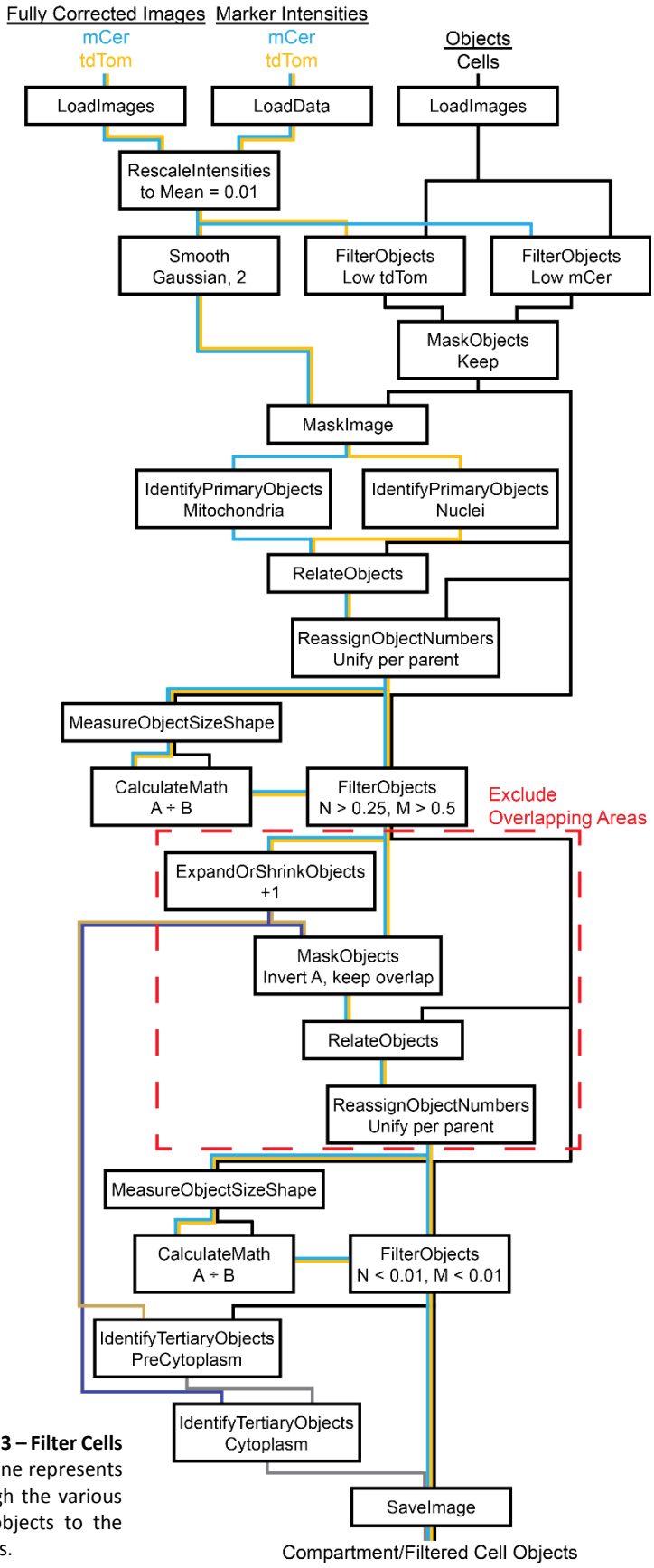

Figure 3-S4. Diagram of CellProfiler Pipeline 3 – Filter Cells and Identify Compartments. Each colored line represents flow (top to bottom) of information through the various modules from an input images and cell objects to the output filtered cell and compartment objects.

other compartment to find a conservative non-overlapping region for each compartment.

10. Filter out the cells that have overly small non-overlapped compartments and those that have none of a compartment.
11. If desired, the remainder of the cell can be grouped into a third compartment.
12. Crosstalk-correct mildly smoothed target channel images, measure per-cell, per-compartment level of the target marker mean intensity, correct this value for autofluorescence and leakage, as empirically determined above. (See Figure 3-S5)
13. Filter out cells with an overly low denominator for the localization index (noise-level signal), as empirically determined above.

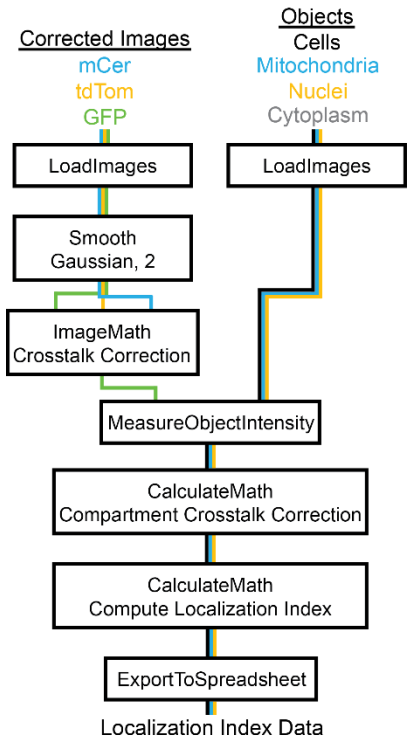

Figure 3-S5. Diagram of CellProfiler Pipeline 4 – Calculate Localization Index and Export Intensity Data. Each colored line represents flow (top to bottom) of information through the various modules from an input image to the output localization data.

14. Analyze the distributions among the samples using Kruskal–Wallis test and the post-hoc multiple comparisons test with correction for ties.

Q-SCAn fully automated

The method as implemented above uses two ImageJ macros and four CellProfiler pipelines, with an intermediate step in between pipelines 2 and 3. We have written a small program for use in Windows to coordinate this process and fully automate it. This program and its source code is distributed with the pipelines at http://www.biochem.emory.edu/doetsch/qscan.html [and included as a supplemental file identified in the Appendix].

References

1. Ye W, Lin W, Tartakoff AM, Tao T. Karyopherins in nuclear transport of homeodomain proteins during development. Biochim Biophys Acta. 2011;1813(9):1654-1662. doi:10.1016/j.bbamcr.2011.01.013.

2. Kumar S, Saradhi M, Chaturvedi NK, Tyagi RK. Intracellular localization and nucleocytoplasmic trafficking of steroid receptors: an overview. Mol Cell Endocrinol. 2006;246(1-2):147-156. doi:10.1016/j.mce.2005.11.028.

3. Larsen NB, Rasmussen M, Rasmussen LJ. Nuclear and mitochondrial DNA repair: similar pathways? Mitochondrion. 2005;5(2):89-108. doi:10.1016/j.mito.2005.02.002.

4. 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. doi:10.1128/mcb.01357-08.

5. 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. doi:10.1093/nar/gkq108.

6. Fernández-Cid A, Riera A, Herrero P, Moreno F. Glucose levels regulate the nucleo-mitochondrial distribution of Mig2. Mitochondrion. 2012;12(3):370-380. doi:10.1016/j.mito.2012.02.001.

7. Rowe LA, Degtyareva N, Doetsch PW. DNA damage-induced reactive oxygen species (ROS) stress response in *Saccharomyces cerevisiae*. Free Radical Biol Med. 2008;45(8):1167-1177. doi:10.1016/j.freeradbiomed.2008.07.018.

8. Zinchuk V, Zinchuk O, Okada T. Quantitative colocalization analysis of multicolor confocal immunofluorescence microscopy images: pushing pixels to explore biological phenomena. Acta Histochem Cytochem. 2007;40(4):101-111. doi:10.1267/ahc.07002.

9. Gonzalez-Gonzalez IM, Jaskolski F, Goldberg Y, Ashby MC, Henley JM. Measuring membrane protein dynamics in neurons using fluorescence recovery after photobleach. Methods Enzymol. 2012;504:127-146. doi:10.1016/b978-0-12-391857-4.00006-9.

10. Gauthier A, Brandt R. Live cell imaging of cytoskeletal dynamics in neurons using fluorescence photoactivation. Biol Chem. 2010;391(6):639-643. doi:10.1515/bc.2010.060.

11. Mah W, Deme JC, Watkins D, Fung S, Janer A, Shoubridge EA, Rosenblatt DS, Coulton JW. Subcellular location of MMACHC and MMADHC, two human proteins central to intracellular vitamin B12 metabolism. Mol Genet Metab. 2013;108(2):112-118. doi:10.1016/j.ymgme.2012.11.284.

12. Pronsato L, Boland R, Milanesi L. Non-classical localization of androgen receptor in the C2C12 skeletal muscle cell line. Arch Biochem Biophys. 2013;530(1):13-22. doi:10.1016/j.abb.2012.12.011.

13. Moye-Rowley WS, Harshman KD, Parker CS. Yeast *YAP1* encodes a novel form of the jun family of transcriptional activator proteins. Genes Dev. 1989;3(3):283-292. doi:10.1101/gad.3.3.283.

14. You HJ, Swanson RL, Harrington C, Corbett AH, Jinks-Robertson S, Sentürker S, Wallace SS, Boiteux S, Dizdaroglu M, Doetsch PW. *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. doi:10.1021/bi991121i.

15. Burgers PM, Klein MB. Selection by genetic transformation of a *Saccharomyces cerevisiae* mutant defective for the nuclear uracil-DNA-glycosylase. J Bacteriol. 1986;166(3):905-913.

16. Gietz RD, Woods RA. Transformation of yeast by lithium acetate/single-stranded carrier DNA/polyethylene glycol method. Methods Enzymol. 2002;350:87-96. doi:10.1016/S0076-6879(02)50957-5.

17. 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.

18. Mumberg D, Müller R, Funk M. Yeast vectors for the controlled expression of heterologous proteins in different genetic backgrounds. Gene. 1995;156(1):119-122. doi:10.1016/0378-1119(95)00037-7.

19. Kahana JA, Schnapp BJ, Silver PA. Kinetics of spindle pole body separation in budding yeast. Proc Natl Acad Sci USA. 1995;92(21):9707-9711.

20. Hodel MR, Corbett AH, Hodel AE. Dissection of a nuclear localization signal. J Biol Chem. 2001;276(2):1317-1325. doi:10.1074/jbc.M008522200.

21. Shaner NC, Campbell RE, Steinbach PA, Giepmans BNG, Palmer AE, Tsien RY. Improved monomeric red, orange and yellow fluorescent proteins derived from *Discosoma* sp. red fluorescent protein. Nat Biotechnol. 2004;22(12):1567-1572. doi:10.1038/nbt1037.

22. Pfanner N, Muller HK, Harmey MA, Neupert W. Mitochondrial protein import: involvement of the mature part of a cleavable precursor protein in the binding to receptor sites. EMBO J. 1987;6(11):3449-3454.

23. Rizzo MA, Springer GH, Granada B, Piston DW. An improved cyan fluorescent protein variant useful for FRET. Nat Biotechnol. 2004;22(4):445-449. doi:10.1038/nbt945.

24. Manders MM, Verbeek PJ, Aten JA. Measurement of co-localization of objects in dual colour confocal images. J Microsc. 1993;169(3):375-382.

25. Wen W, Meinkotht JL, Tsien RY, Taylor SS. Identification of a signal for rapid export of proteins from the nucleus. Cell. 1995;82(3):463-473. doi:10.1016/0092-8674(95)90435-2.

26. Anderson MT, Tjioe IM, Lorincz MC, Parks DR, Herzenberg LA, Nolan GP, Herzenberg LA. Simultaneous fluorescence-activated cell sorter analysis of two distinct transcriptional elements within a single cell using engineered green fluorescent proteins. Proc Natl Acad Sci USA. 1996;93(16):8508-8511.

27. Christianson TW, Sikorski RS, Dante M, Shero JH, Hieter P. Multifunctional yeast high-copy-number shuttle vectors. Gene. 1992;110(1):119-122. doi:10.1016/0378-1119(92)90454-W.

28. Nakai K, Horton P. PSORT: a program for detecting sorting signals in proteins and predicting their subcellular localization. Trends Biochem Sci. 1999;24(1):34-35. doi:10.1016/S0968-0004(98)01336-X.

29. Bannai H, Tamada Y, Maruyama O, Nakai K, Miyano S. Extensive feature detection of N-terminal protein sorting signals. Bioinformatics. 2002;18(2):298-305. doi:10.1093/bioinformatics/18.2.298.

30. Winston F, Dollard C, Ricupero-Hovasse SL. Construction of a set of convenient *Saccharomyces cerevisiae* strains that are isogenic to S288C. Yeast. 1995;11(1):53-55. doi:10.1002/yea.320110107.

31. Scriven DRL, Lynch RM, Moore EDW. Image acquisition for colocalization using optical microscopy. Am J Physiol Cell Physiol. 2008;294(5):C1119-C1122. doi:10.1152/ajpcell.00133.2008.

32. Schneider CA, Rasband WS, Eliceiri KW. NIH Image to ImageJ: 25 years of image analysis. Nat Methods. 2012;9(7):671-675. doi:10.1038/nmeth.2089.

33. Kamentsky L, Jones TR, Fraser A, Bray M-A, Logan DJ, Madden KL, Ljosa V, Rueden C, Eliceiri KW, Carpenter AE. Improved structure, function and compatibility for CellProfiler: modular high-throughput image analysis software. Bioinformatics. 2011;27(8):1179-1180. doi:10.1093/bioinformatics/btr095.

34. Jones T, Kang I, Wheeler D, Lindquist R, Papallo A, Sabatini D, Golland P, Carpenter A. CellProfiler Analyst: data exploration and analysis software for complex image-based screens. BMC Bioinformatics. 2008;9(1):482. doi:10.1186/1471-2105-9-482.

35. Zar JH. Biostatistical Analysis. 4th edition. Upper Saddle River, NJ: Prentice Hall; 1999. 195-199, 223-225.

36. Yan C, Lee LH, Davis LI. Crm1p mediates regulated nuclear export of a yeast AP-1-like transcription factor. EMBO J. 1998;17(24):7416-7429. doi:10.1093/emboj/17.24.7416.

37. Gulshan K, Lee SS, Moye-Rowley WS. Differential oxidant tolerance determined by the key transcription factor Yap1 is controlled by levels of the Yap1-binding protein, Ybp1. J Biol Chem. 2011;286(39):34071-34081. doi:10.1074/jbc.M111.251298.

38. Chatterjee A, Singh KK. Uracil-DNA glycosylase–deficient yeast exhibit a mitochondrial mutator phenotype. Nucleic Acids Res. 2001;29(24):4935-4940. doi:10.1093/nar/29.24.4935.

39. Huh W-K, Falvo JV, Gerke LC, Carroll AS, Howson RW, Weissman JS, O'Shea EK. Global analysis of protein localization in budding yeast. Nature. 2003;425(6959):686-691. doi:10.1038/nature02026.

40. Wu B, Piatkevich KD, Lionnet T, Singer RH, Verkhusha VV. Modern fluorescent proteins and imaging technologies to study gene expression, nuclear localization, and dynamics. Curr Opin Cell Biol. 2011;23(3):310-317. doi:10.1016/j.ceb.2010.12.004.

41. Bush A, Colman-Lerner A. Quantitative measurement of protein relocalization in live cells. Biophys J. 2013;104(3):727-736. doi:10.1016/j.bpj.2012.12.030.

42. Gordon A, Colman-Lerner A, Chin TE, Benjamin KR, Yu RC, Brent R. Single-cell quantification of molecules and rates using open-source microscope-based cytometry. Nat Methods. 2007;4(2):175-181.

43. Wang L, Jackson WC, Steinbach PA, Tsien RY. Evolution of new nonantibody proteins via iterative somatic hypermutation. Proc Natl Acad Sci USA. 2004;101(48):16745-16749. doi:10.1073/pnas.0407752101.

44. Caster AH, Kahn RA. Computational method for calculating fluorescence intensities within three-dimensional structures in cells. Cell Logist. 2012;2(4):176-188. doi:10.4161/cl.23150.

Chapter 4
Characterization of Ung1 Uracil–DNA Glycosylase: Bisulfite-Induced Deamination, Spontaneous Mutagenesis, and Reactive Oxygen Species Levels

Abstract

The *Saccharomyces cerevisiae* DNA repair protein Ntg1, a DNA *N*-glycosylase/AP lyase, localizes to both nuclei and mitochondria. Recently, it was discovered that the distribution of Ntg1 shifts when exposed to oxidative DNA damage that preferentially targets each organelle. Many of the other components of the pathway are shared by both compartments, opening up a potentially important and novel mode of regulation for DNA repair proteins: dynamic compartmentalization. Two of the key questions prompted by this finding were: 1) whether this mode of regulation is restricted to Ntg1 or if it is general to BESIR; and 2) what signaling pathways may be involved in causing Ntg1 to change localization. I hypothesized that dynamic compartmentalization is a general mode of regulation, that abasic sites are responsible for producing base damage–dependent reactive oxygen species (ROS) signaling, and that this signaling is responsible for dynamic compartmentalization. To partially address question 1, the uracil–DNA glycosylase Ung1 was selected. To determine if the chemical agent bisulfite is suitable for assessing changes in Ung1 localization, cells with varied Ung1 activity were treated with bisulfite and then toxicity and mutagenesis was measured. To address question 2, ROS levels and spontaneous mutation frequencies were measured in repair-deficient strains lacking abasic site processing (*ntg1 ntg2 apn1* or in combination with *rad1*) were combined with *ung1Δ* to vary the level of abasic sites. The results demonstrate that bisulfite is not a suitable agent for studying Ung1 in vivo due to its high toxicity and low mutagenicity. The results further strongly support the hypothesis that abasic sites are responsible for base damage–induced ROS signaling. However, further work will be needed to determine whether ROS signaling is responsible for dynamic compartmentalization of Ntg1 or other BESIR proteins.

Introduction

The *Saccharomyces cerevisiae* DNA repair protein Ntg1, a DNA *N*-glycosylase/AP lyase responsible for excising oxidized pyrimidines, localizes to both nuclei and mitochondria as part of the base excision and strand incision repair (BESIR) pathway. Recently, it was discovered that the distribution of Ntg1 shifts between the nucleus and mitochondria when exposed to oxidative DNA damaging agents (nucleus: hydrogen peroxide; mitochondria:

hydrogen peroxide plus antimycin A) that preferentially target each organelle (2,3). As BESIR is a critical pathway and many of the components of the pathway are shared by both compartments, these results opened up a potentially important and novel mode of regulation for DNA repair proteins: dynamic compartmentalization. One of the key questions prompted by this finding was whether dynamic compartmentalization is restricted to this protein or whether it applies to the other shared repair machinery. To address this question, the uracil–DNA glycosylase Ung1 was selected. Ung1 has several properties that made it an attractive target. Unlike Ntg1, Ung1 does not possess AP lyase activity, and Ung1 is approximately 10-fold more abundant per cell (4,5). But like Ntg1, Ung1 possesses a very similar overlapping N-terminal targeting signal structure (6,7), which could readily create competition for localization signal receptors.

Hydrogen peroxide and antimycin A have been established tools in the field for inducing in vivo oxidative stress and oxidation-dependent mutagenesis in nuclei and mitochondria (9). MMS is available for inducing alkylation DNA damage and is also widely used in vivo. However, very little work has been done on agents for inducing the in vivo deamination of cytosine to uracil. Sodium bisulfite and nitrous acid have long been known to deaminate cytosine (10-12), with nitrous acid preferentially also deaminating adenine and guanine (12). The primary application of sodium bisulfite has been for in vitro methods to detect 5-methylcytosine (13) or in vitro experiments requiring uracil–DNA (14). The only study to use bisulfite in vivo in *S. cerevisiae* was a pioneer study identifying an Ung1 mutant allele conferring sensitivity to bisulfite (15). There are a few scattered reports using bisulfite in several model systems, including *C. elegans* (16) and *E. coli* (17), though the latter reported that the glycosylase was inactivated by bisulfite. Because bisulfite is relatively specific for cytosine deamination, it was a promising agent to be used for this purpose. The utility of bisulfite was tested by characterizing its effects in *S. cerevisiae* cells with varying levels of Ung1 activity.

Another key question prompted by the Ntg1 dynamic compartmentalization finding is what signal pathways may be involved in causing Ntg1 to change localization. The same study also found that mitochondrial DNA was required for Ntg1 to shift localization to the mitochondria, suggesting that DNA damage is a required part of the localization signaling pathway (2). Methyl methanesulfonate (MMS), a DNA alkylating agent which creates lesions that are not a target for Ntg1 glycosylase activity (though which may lead to substrates via the alkylpurine glycosylase Mag1), spurs Ntg1 to shift to the nucleus (2). MMS treatment also indirectly causes production of reactive oxygen species (ROS) (2,18). Importantly, ROS levels increase in strains with defective abasic site processing, which is achieved by knocking out the genes for the major AP endonuclease Apn1 and AP lyases Ntg1 and Ntg2 in the BESIR pathway; an enhanced defect is obtained by additionally knocking out the backup nucleotide excision repair (NER) pathway (18). ROS has been established as part of the base damage signaling pathway, which triggers the transcription factor Yap1 to accumulate in the nucleus and direct the transcription of BESIR genes and free radical scavengers (19). ROS were therefore a good candidate to act as the signal for dynamic compartmentalization. However, the initiating lesions for ROS signaling was still unknown. I hypothesized that abasic sites, a common BESIR intermediate, were the lesions responsible for generating ROS, rather than the myriad of possible base lesions. Abasic sites are a good candidate to initiate signaling not only because they are common to all BESIR-recognized damage, but also because they are highly toxic and mutagenic in their own right and thus sensing their accumulation should be a high priority for the cell.

To address the signaling component, I collaborated with an Emory undergraduate Xi Jiang to generate and employ strains knocked out for BESIR or BESIR and NER to reduce abasic site repair, and thus allow abasic sites to accumulate. Because Ung1 is a fairly abundant monofunctional glycosylase, and uracil lesions are common in *S. cerevisiae* (7), these strains

were combined with *UNG1* deletion to reduce the generation and accumulation of abasic sites. This system allowed us to directly test the role of abasic sites in the generation of ROS. The results we obtained strongly support the model that abasic sites are a major proximal DNA base damage signal and result in the generation of ROS. However, bisulfite proved to be a very poor agent for in vivo deamination. Other approaches will have to be developed to increase the levels of uracil in the genome and to probe the activity and regulation of Ung1.

Materials and Methods

Yeast strains, media, and growth conditions

Saccharomyces cerevisiae strains and plasmids used in this study are listed in Table 4-1. *S. cerevisiae* cells were cultured at 30 °C in rich YPD medium (1% yeast extract, 2% peptone, and 2% dextrose; plus 2% agar for plates), synthetic defined arginine drop-out medium supplemented with canavanine (SD arg^- can^+) for the mutation frequency assay, and YPD supplemented with drug (G418 for *kanMX*, hygromycin for *hphMX4*, blastocidin for *BSD*, nourseothricin for *natNT2*) and SD drop-out media ($uracil^-$, $tryptophan^-$) for selection. Plasmids were transformed into cells by a modified lithium acetate method (20). The *UNG1* locus was replaced in DSC226 and hDNP119 with a nourseothricin-resistance cassette (*NAT*) by homology-directed recombination. hDNP119 was subsequently induced to sporulate, and the asci were dissected, grown, and screened for genotype by plating on selective media. Multiple spores for each genotype were isolated to reduce the effects of additional mutations resulting from the DNA repair deficiency.

Sodium bisulfite treatment

Sodium bisulfite can be a problematic compound to employ in biological studies, and special consideration must be taken when determining its concentration in solution. Solid sodium bisulfite ($NaHSO_3$, 104.06 g/mol) only exists as a mixture with sodium metabisulfite ($Na_2S_2O_5$, 190.11 g/mol), but the exact proportion of the mixture is typically unknown. In

Table 4-1. Strains and plasmids used in this study.

Name	Genotype	Ref
hDNP119	*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 ung1Δ::NAT/UNG1*	(1)
WT	*MATα his7-1 lys2Δ5'::LEU-lys2Δ3' ade5-1 trp1-289 ura3-52*	This study
BESIR$^-$	WT *ntg1::hphMX4 ntg2::BSD apn1::TRP1*	This study
BESIR$^-$/NER$^-$	WT *ntg1::hphMX4 ntg2::BSD apn1::TRP1 rad1::kanMX*	This study
WT ung1Δ	WT *ung1Δ::natNT2*	This study
BESIR$^-$ ung1Δ	BESIR$^-$ *ung1Δ::natNT2*	This study
BESIR$^-$/NER$^-$ ung1Δ	BESIR$^-$/NER$^-$ *ung1Δ::natNT2*	This study
FY86 (ACY193)	*MATα ura3-52 leu2Δ1 his3Δ200*	(8)
DSC499	FY86 *ung1Δ::natNT2*	This study
DSC500	DSC499 pD0419	This study
pD0419	*UNG1–GFP, 2μ, URA3, Amp^R*	(6)

aqueous solution, 1 mol of metabisulfite decomposes to approximately 2 mol of bisulfite. The purity of the solid and the ratio of its components is expressed as the percent of the solid mass released as SO_2 by iodometry titration. This value should be determined by the chemical manufacturer and noted on the label. Since the amount of bisulfite released in solution by a given mass of solid is approximately 1:1 with the amount of sulfur atoms, an effective molecular weight of the solid can be calculated as follows:

$$MW_{Eff} = MW_{\mathrm{SO_2}} \div \%\mathrm{SO_2}$$

For example:

$$EMW = 64.06\frac{\mathrm{g}}{\mathrm{mol}} \div 66.9\% = 95.75\frac{\mathrm{g}}{\mathrm{mol}}$$

Working with bisulfite is further complicated by its pH-dependent behavior (pK_{a1} = 1.81; pK_{a2} = 6.36) (21). Above pH 4, the majority of the compound exists as soluble anions HSO_3^- or SO_3^{2-}. Below pH 4, it converts to gaseous SO_2 + H_2O, which can then escape the solution. This characteristic requires bisulfite solutions to be mixed fresh each time, especially those at low pH. For use with cells, bisulfite has typically been buffered with 0.1 M sodium citrate (15), but the pH varies. In this study, pH 3.6 and 6.0 were used and approximately 2×10^7 cells/mL were exposed to a 0–30 mM bisulfite solution for 1 hour at 30 °C with aeration.

Cytotoxicity and canavanine resistance mutagenesis assays

Cytotoxicity and canavanine resistance mutagenesis assays were carried out as previously described (22). In brief, overnight late–log phase (approximately 1×10^8 cells/mL) cultures of *S. cerevisiae* cells were washed twice with H_2O, resuspended in 1 mL H_2O, counted via hemocytometer, and adjusted to a density of $2–4 \times 10^7$ cells/mL. For bisulfite-induced deamination experiments, the cells were treated with bisulfite as described above and then washed twice with H_2O. Cells were then diluted and inoculated onto YPD and SD arg^- can^+ agar plates and incubated for 3 days, at which point the number of colonies were counted and mutation frequency calculated. "Jackpot" cultures with extraordinarily high mutation frequencies were excluded.

Measurement of ROS levels by flow cytometry

Overnight late–log phase (approximately 1×10^8 cells/mL) cultures of *S. cerevisiae* cells were treated for 2 hours with superoxide fluorescent probe dihydroethidium (DHEt, Life Technologies) as previously described (18). In brief, cells were washed twice with H_2O and resuspended in phosphate-buffered saline (137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4). Cells (10,000/sample) were analyzed with an LSR II flow cytometer (BD Biosciences) using the 488-nm laser, 550-nm longpass dichroic mirror, and 560–590 nm bandpass filter.

Data analysis

Three to six independent colonies were used for each experiment, and at least three independent experiments for each strain were completed. The most consistent data from two of three independent isolates of the same genotype were retained. FlowJo 5.7.2 was used to process the raw cytometry data, with gating for single unclumped cells and cells with dye. The geometric mean of the fluorescence intensity distribution was obtained for each sample and corrected for background fluorescence intensity with an undyed sample. Two-sample

T-tests with correction for multiple comparisons were used to analyze the bisulfite experiments ($\alpha = 0.05$). Two-way ANOVA with Tukey–Cramer post hoc multiple comparisons was applied to assess the effect of Ung1 knockout in BESIR$^-$ and BESIR$^-$/NER$^-$ backgrounds on ROS levels and spontaneous mutation frequencies ($\alpha = 0.05$). Statistical analyses were performed with Stata 11.2.

Results

Uracil glycosylase activity sensitizes cells to bisulfite-induced deamination

To determine whether bisulfite was a suitable agent to create in vivo lesions recognized by Ung1 (i.e. uracil), the cytotoxicity and mutagenesis of bisulfite (pH 3.6) on *S. cerevisiae* cells containing *UNG1*, *ung1Δ*, and 2μ-overexpressed *UNG1–GFP* fusion protein were examined. These data demonstrated that bisulfite is a fairly potent agent over a small range of concentrations: only 59% of wild type cells survived 15 mM bisulfite, and 6% survived 25 mM bisulfite (Figure 4-1A). Strikingly, bisulfite toxicity was enhanced in the 2μ *UNG1–GFP* overexpression cells, with 31% and 0.9% ($p < 0.01$), respectively, while *ung1Δ* cells were as hardy as wild type cells (Figure 4-1A).

Despite the toxicity of bisulfite in wild-type cells, no increase in mutagenesis was apparent across any of the three doses (Figure 4-1B). The mutation frequency of *ung1Δ* cells

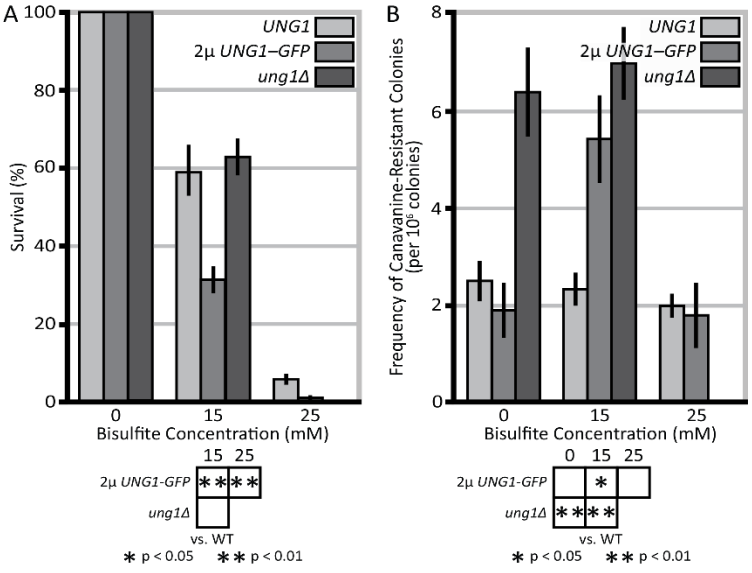

Figure 4-1. *S. cerevisiae* cells overexpressing Ung1–GFP are sensitive to bisulfite toxicity and mutagenesis. *S. cerevisiae* cells containing *UNG1*, *ung1Δ*, and *UNG1–GFP* on a high-copy 2μ plasmid (2μ *UNG1–GFP*) were exposed to 0, 15, or 25 mM sodium bisulfite in 0.1 M sodium citrate at pH 3.6 for 1 hour at 30 °C. Cells were grown for 2–3 days on rich and canavanine-containing medium. Values are mean ± SEM from at least 3 independent experiments, n = 3–6 independent colonies for each experiment. Pairwise statistical significance is indicated in the tables below the graphs. A) Percent of cells surviving the bisulfite treatment. B) Frequency of canavanine-resistant colonies after bisulfite treatment.

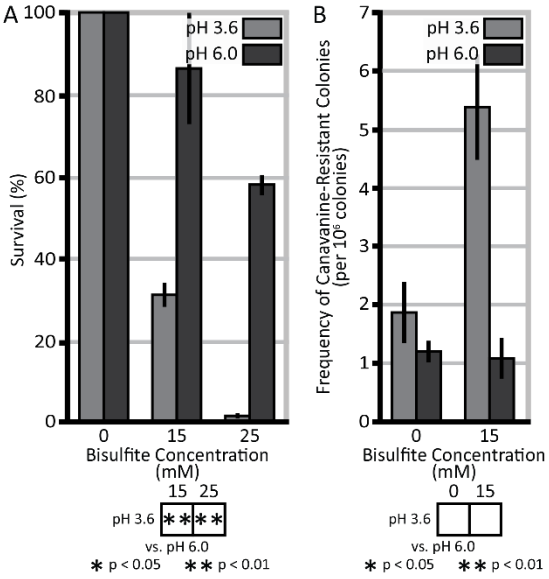

Figure 4-2. Bisulfite sensitivity of *S. cerevisiae* cells overexpressing Ung1–GFP is dependent on low pH. *S. cerevisiae* cells containing *UNG1–GFP* on a high-copy 2μ plasmid (2μ *UNG1–GFP*) were exposed to 0, 15, or 25 mM sodium bisulfite in 0.1 M sodium citrate at pH 3.6 or pH 6.0 for 1 hour at 30 °C. Cells were grown for 2–3 days on rich and canavanine-containing medium. Values are mean ± SEM from at least 3 independent experiments, n = 3–6 independent colonies for each experiment. Pairwise statistical significance is indicated in the tables below the graphs. A) Percent of cells surviving the bisulfite treatment. B) Frequency of canavanine-resistant colonies after bisulfite treatment.

was similarly unaffected by bisulfite exposure, but they did display a 1.6-fold increase in spontaneous mutagenesis over wild type ($p < 0.01$) (Figure 4-1B). In contrast, 2μ Ung1–GFP overexpression cells had a 1.9-fold induction of mutagenesis upon 15 mM bisulfite treatment ($p < 0.05$); however, this induction disappeared in the 25 mM treatment (Figure 4-1B).

Effects of bisulfite-induced deamination depend on low pH

Bisulfite is a weak acid and can interconvert between SO_2, HSO_3^-, and SO_3^{2-}. Prior work with yeast largely used pH 3.6 buffer, at which point a small fraction of bisulfite readily converts to the evolvable gas SO_2. Sulfur dioxide gas is used as an antimicrobial agent in wine making (23) and may have non-mutagenic modes of toxicity. To determine whether less toxicity, but more mutagenesis, could be obtained by treating at a higher pH where the primary equilibrium would be between HSO_3^- and SO_3^{2-}, the 2μ *UNG1–GFP* overexpression cells, which displayed the most bisulfite sensitivity, were exposed to bisulfite at both pH 3.6 and pH 6.0. The results clearly demonstrate that both bisulfite-induced cytotoxicity and mutagenesis are dependent on the acidic buffer ($p < 0.01$ and marginally nonsignificant, respectively) (Figure 4-2). No difference in survival was observed for buffer pH alone (data not shown).

Uracil glycosylase activity is required for elevated ROS level and mutation frequency in base damage repair–compromised cells

To test the hypothesis that abasic sites are responsible for the increased ROS level

and mutation frequency in base damage repair–deficient strains, knockouts of *UNG1* in repair-proficient, BESIR$^-$, and BESIR$^-$/NER$^-$ strain backgrounds were generated. Measurements were taken of their spontaneous mutation frequencies (Figure 4-3A) and relative ROS levels (Figure 4-3B). As expected, BESIR$^-$ was associated with a mild increase in both ROS (44% increase, $p < 0.01$) and mutation frequency (26-fold increase, $p < 0.01$) over wild type, and these effects were further enhanced over BESIR$^-$ by adding NER$^-$ to the genotype for both ROS (64% increase, $p < 0.01$) and mutation frequency (12-fold increase, $p < 0.01$). Ung1 deletion had no significant effect on either biological endpoint in the repair-proficient background, nor on the ROS levels in the BESIR$^-$ background. However, *UNG1* deletion dramatically reduced the mutation frequency in the BESIR$^-$ background (82% reduction, $p < 0.05$) and both the ROS level (62% reduction, $p < 0.01$) and mutation frequency (85% reduction, $p < 0.01$) in the BESIR$^-$/NER$^-$ strain background. These results indicate that glycosylase excision of spontaneous or misincorporated uracil lesions is a major cause of elevated mutagenesis and ROS levels in cells deficient in abasic site processing. A preliminary experiment to rescue the *ung1Δ* phenotypes by transformation of the 2μ *UNG1–GFP* plasmid was conducted, but the 2μ vector alone appeared to impact ROS levels (data not shown).

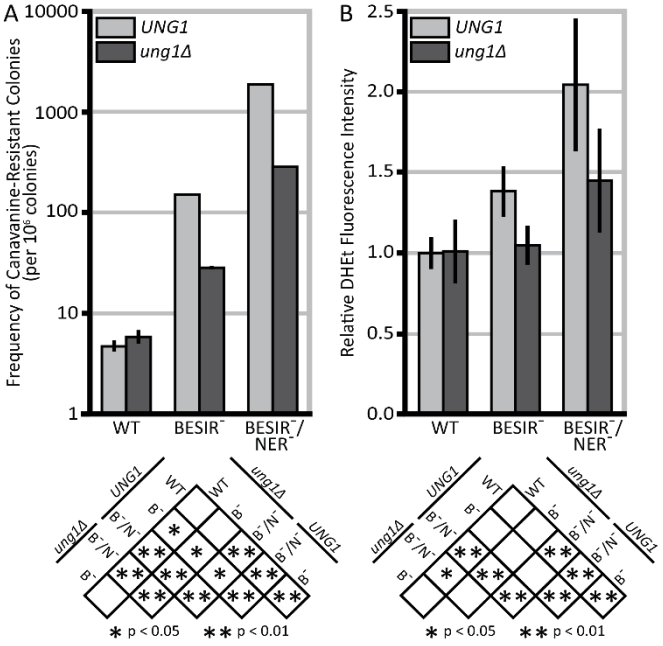

Figure 4-3. Spontaneous mutation frequency and heightened ROS levels in base damage repair–deficient *S. cerevisiae* strains are dependent on Ung1 activity. *S. cerevisiae* cells containing *ntg1 ntg2 apn1* (BESIR-) and additionally with *rad1* (BESIR-/NER-) in combination with *UNG1* or *ung1Δ* were grown overnight to late-log phase. A) Frequency of canavanine-resistant colonies. Cells were plated, and grown for 2–3 days on rich and canavanine-containing medium. B) Relative fluorescence of ROS-dependent dye dihydroethidium (DHEt). 10,000 cells were analyzed by flow cytometry per sample, data was gated for single, unclumped cells with dye, and the geometric mean was obtained to represent the distribution. Values are mean ± SEM from at least 5 independent experiments and at least 2 independent strain isolates, n = 3–6 independent colonies for each experiment. Pairwise statistical significance is indicated in the tables below the graphs.

Discussion

The goal of the work presented in this chapter was twofold. First, to characterize the utility of bisulfite as an agent to introduce in vivo uracil lesions recognizable by Ung1 in a manner similar to the use of hydrogen peroxide to generate oxidative lesions for Ntg1. Second, to test the hypothesis that abasic sites or other downstream BESIR intermediates are responsible for base damage–induced ROS signaling.

The cytotoxic and mutagenic effect of bisulfite on cells with varied amounts of Ung1 activity was measured, and bisulfite was found to be highly toxic but poorly mutagenic in repair-proficient cells. By contrast, hydrogen peroxide induces a 10-fold increase in mutation frequency in repair-proficient cells at a concentration that kills only approximately half of the cells exposed (24). These data suggest that the major mode of bisulfite toxicity is largely independent of cytosine deamination, potentially by reacting with cysteine residues on proteins and depleting the glutathione pool (25). These findings are buttressed by the observation that cells without Ung1 activity were no more sensitive to bisulfite than those with Ung1 activity even while displaying modestly increased spontaneous mutagenesis. However, this observation is at odds with an early study of Ung1, which identified a mutant of Ung1 (*ung1-1*) which conferred bisulfite sensitivity and depleted uracil excision activity from the cell (15). (While that study describes residual mitochondrial uracil excision activity, the author has communicated that this activity was later determined to be an artifact.) The difference could be explained by the nature of the mutant allele in that study, but no sequence information is available for the mutant.

The one condition exhibiting sensitivity to bisulfite-induced effects was *UNG1–GFP* expressed from a high-copy 2μ plasmid, presenting with both increased cytotoxicity and a modest increase in mutation frequency. This result indicates that bisulfite is inducing at least a small amount of cytosine deamination and that these uracils are being processed by Ung1

into abasic sites. While a uracil paired with guanine is mutagenic, an abasic site is both mutagenic and can block polymerase progression, which can result in cell death if left unrepaired (26). DNA repair proteins are typically present in low quantities (4) and their activities are highly regulated because repair intermediates, including abasic sites, can be more toxic than the initial base lesion (27). Overexpressing a glycosylase can result in the generation of more of these toxic intermediates than can be efficiently handled by the down-stream repair enzymes, and this has previously been observed in mitochondrial DNA with overexpression of Ung1 (7). Curiously, the induction effect in this strain is lost at the 25 mM treatment level. This loss of effect may be a result of heavy selection against cells containing the plasmid in this highly toxic exposure, such that the remaining cells thus produce less Ung1 and generate fewer abasic sites.

While bisulfite has previously been used with Ung1 in yeast at pH 3.6 (15), I wanted to determine whether increased pH and concomitant reduced production of SO_2 gas would produce a lower toxicity–mutation frequency ratio. The two pH levels were compared with the most bisulfite-sensitive strain, 2μ *UNG1–GFP*. However, all bisulfite-induced effects were substantially reduced in the pH 6 buffer, and this finding agrees with prior work (15). These results suggest that the mechanism of entry of bisulfite into *S. cerevisiae* cells is dependent on the chemical species. The HSO_3^- anion may not be able to efficiently penetrate the yeast cell wall and membrane, while the neutral SO_2 can readily diffuse across the barrier. Once inside the cell's cytosol in a neutral environment, SO_2 would rapidly protonate back to HSO_3^- and SO_3^{2-}, effectively trapping bisulfite inside the cell and allowing it to accumulate. Collectively, these results demonstrate that bisulfite is not a suitable agent to induce in vivo cytosine deamination. The only other chemical agent which can deaminate cytosine, but which prefers adenine and guanine, is nitrous acid (12). However, a preliminary trial showed high toxicity and morphological irregularities by microscopic inspection (data not shown). Thus, it

appears that chemical agents are wholly unsuitable for inducing in vivo deamination lesions. A different approach is needed to effectively study the regulation of uracil–DNA glycosylases. One suggestion would be to express a deaminase, such as the activation-induced deaminase (AID) involved in immune system gene recombination. A hypermutagenic AID from sea lamprey has been expressed in yeast and its deamination activity has been characterized (28), making it a good candidate for this work.

The previous experiments with bisulfite also highlight the comparatively high toxicity and mutagenicity of the common repair intermediates in the BESIR pathway, and supports the model that these intermediates, and not the base lesions themselves, would be responsible for mediating base damage signaling, and specifically ROS signaling. To further test this hypothesis, we constructed strains deficient in BESIR and both BESIR and NER (because NER can partially cover for loss of BESIR (22,29)), with or without Ung1 activity. The expected increased mutation frequencies in the cells lacking the core pathways was dramatically reduced without Ung1 activity. This result indicates that abasic sites derived from misincorporated uracil are a major component of spontaneous mutagenesis in these base damage repair–deficient strains. This finding was confirmed by independent work published soon after the bulk of these experiments had been completed (30). What was particularly striking, however, is that the expected high levels of ROS observed in the base damage repair–deficient cells were also dramatically reduced in the absence of Ung1 activity. These data strongly support our hypothesis that base damage–induced ROS signaling is a result of the presence of abasic sites and not of the initial base lesions.

Further work is needed to characterize this signaling mechanism. The next logical step is to determine how abasic sites are responsible for generating these signals. One experiment that could provide some clues would be to use a modified chromatin immunoprecipitation method. Biotin-tagged DNA containing abasic sites could be generated

in vitro and mixed with nuclear extract, along with a generous amount of carrier DNA to delay nuclease activity. Any proteins bound to the abasic site could be crosslinked, and then the biotin-labeled abasic site–containing DNA could be pulled down, crosslinks reversed, and analyzed by mass spectrometry to determine which proteins bound to abasic DNA but not undamaged DNA and vice versa.

Another important aspect is to determine which oxidase(s) are responsible for generating the base damage–induced ROS signal. Unfortunately, the best method to detect a DNA damage–responsive oxidase, assuming it does not bind directly to the abasic site itself, would be to look for reduced ROS levels in a base damage repair–deficient strain. Unfortunately, a high-throughput fluorescence-activated cell sorting–based screen is infeasible because ROS measurements are highly variable, both due to inherent cellular heterogeneity and inconsistent dye uptake.

Acknowledgements

I must first extend gratitude to a former undergraduate student, Xi Jiang. He carried out much of the work on the repair-deficient strains, and he wrote and defended an Honors Thesis on this work in 2013. That portion of this chapter has been written independently but is based on the same data. I would also like to thank the members of the Doetsch laboratory and especially Dr. Natalya Degtyareva, for their constructive support in developing these experiments, and the Emory University Flow Cytometry core facility for providing the flow cytometer and training.

References

1. 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. doi:10.1128/mcb.00307-08.

2. 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. doi:10.1128/mcb.01357-08.

3. 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. doi:10.1093/nar/gkq108.

4. Ghaemmaghami S, Huh WK, Bower K, Howson RW, Belle A, Dephoure N, O'Shea EK, Weissman JS. Global analysis of protein expression in yeast. Nature. 2003;425(6959):737-741. doi:10.1038/nature02046.

5. Swartzlander D. Regulation of base excision repair in response to genotoxic stress [Dissertation]. Atlanta, GA: Emory University; 2012.

6. Bauer NC, Corbett AH, Doetsch PW. Automated quantification of the subcellular localization of multicompartment proteins via Q-SCAn. Traffic. 2013;14(12):1200-1208. doi:10.1111/tra.12118.

7. Chatterjee A, Singh KK. Uracil-DNA glycosylase–deficient yeast exhibit a mitochondrial mutator phenotype. Nucleic Acids Res. 2001;29(24):4935-4940. doi:10.1093/nar/29.24.4935.

8. Winston F, Dollard C, Ricupero-Hovasse SL. Construction of a set of convenient *Saccharomyces cerevisiae* strains that are isogenic to S288C. Yeast. 1995;11(1):53-55. doi:10.1002/yea.320110107.

9. Doudican NA, Song B, Shadel GS, Doetsch PW. Oxidative DNA damage causes mitochondrial genomic instability in *Saccharomyces cerevisiae*. Mol Cell Biol. 2005;25(12):5196-5204. doi:10.1128/mcb.25.12.5196-5204.2005.

10. Shapiro R. Genetic effects of bisulfite (sulfur dioxide). Mutat Res. 1977;39(2):149-175. doi:10.1016/0165-1110(77)90020-3.

11. Shapiro R, DiFate V, Welcher M. Deamination of cytosine derivatives by bisulfite. Mechanism of the reaction. J Am Chem Soc. 1974;96(3):906-912.

12. Shapiro R. The reaction of ribonucleosides with nitrous acid. Side products and kinetics. Biochemistry. 1968;7(1):448-455.

13. Hayatsu H, Negishi K, Shiraishi M, Tsuji K, Moriyama K. Chemistry of bisulfite genomic sequencing; advances and issues. Nucleic Acids Symp Series. 2007;51(51):47-48.

14. Prorok P, Alili D, Saint-Pierre C, Gasparutto D, Zharkov DO, Ishchenko AA, Tudek B, Saparbaev MK. Uracil in duplex DNA is a substrate for the nucleotide incision repair pathway in human cells. Proc Natl Acad Sci USA. 2013;110(39):E3695-3703. doi:10.1073/pnas.1305624110.

15. Burgers PM, Klein MB. Selection by genetic transformation of a *Saccharomyces cerevisiae* mutant defective for the nuclear uracil-DNA-glycosylase. J Bacteriol. 1986;166(3):905-913.

16. Nakamura N, Morinaga H, Kikuchi M, Yonekura S, Ishii N, Yamamoto K, Yonei S, Zhang QM. Cloning and characterization of uracil-DNA glycosylase and the biological consequences of the loss of its function in the nematode *Caenorhabditis elegans*. Mutagenesis. 2008;23(5):407-413. doi:10.1093/mutage/gen030.

17. Duncan BK, Weiss B. Specific mutator effects of Ung (uracil-DNA glycosylase) mutations in *Escherichia coli*. J Bacteriol. 1982;151(2):750-755.

18. Rowe LA, Degtyareva N, Doetsch PW. DNA damage-induced reactive oxygen species (ROS) stress response in *Saccharomyces cerevisiae*. Free Radical Biol Med. 2008;45(8):1167-1177. doi:10.1016/j.freeradbiomed.2008.07.018.

19. Rowe LA, Degtyareva N, Doetsch PW. Yap1: a DNA damage responder in *Saccharomyces cerevisiae*. Mech Age Dev. 2012;133(4):147-156. doi:10.1016/j.mad.2012.03.009.

20. Gietz RD, Woods RA. Transformation of yeast by lithium acetate/single-stranded carrier DNA/polyethylene glycol method. Methods Enzymol. 2002;350:87-96. doi:10.1016/S0076-6879(02)50957-5.

21. Munson JW, Hussain A, Bilous R. Precautionary note for use of bisulfite in pharmaceutical formulations. J Pharm Sci. 1977;66(12):1775-1776.

22. 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.

23. Schneider M, Turke A, Fischer WJ, Kilmartin PA. Determination of the wine preservative sulphur dioxide with cyclic voltammetry using inkjet printed electrodes. Food Chem. 2014;159:428-432. doi:10.1016/j.foodchem.2014.03.049.

24. 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. doi:10.1093/nar/gkh696.

25. Menzel DB, Keller DA, Leung KH. Covalent reactions in the toxicity of SO_2 and sulfite. Adv Exp Med Biol. 1986;197:477-492.

26. Boiteux S, Guillet M. Abasic sites in DNA: repair and biological consequences in *Saccharomyces cerevisiae*. DNA Repair. 2004;3(1):1-12. doi:10.1016/j.dnarep.2003.10.002.

27. Posnick LM, Samson LD. Imbalanced base excision repair increases spontaneous mutation and alkylation sensitivity in *Escherichia coli*. J Bacteriol. 1999;181(21):6763-6771.

28. Lada AG, Dhar A, Boissy RJ, Hirano M, Rubel AA, Rogozin IB, Pavlov YI. AID/APOBEC cytosine deaminase induces genome-wide kataegis. Biol Direct. 2012;7:47. doi:10.1186/1745-6150-7-47.

29. Doetsch PW, Morey NJ, Swanson RL, Jinks-Robertson S. Yeast base excision repair: interconnections and networks. Prog Nucleic Acid Res Mol Biol. 2001;68:29-39.

30. Collura A, Kemp PA, Boiteux S. Abasic sites linked to dUTP incorporation in DNA are a major cause of spontaneous mutations in absence of base excision repair and Rad17-Mec3-Ddc1 (9-1-1) DNA damage checkpoint clamp in *Saccharomyces cerevisiae*. DNA Repair. 2012;11(3):294-303. doi:10.1016/j.dnarep.2011.12.004.

Chapter 5
Analysis of Comparison Between Q-SCAn and Manual Scoring of Ntg1 Dynamic Compartmentalization

Abstract

The *Saccharomyces cerevisiae* DNA repair protein Ntg1, a DNA *N*-glycosylase/AP lyase responsible for excising oxidized pyrimidines, localizes to both nuclei and mitochondria as part of the base excision and strand incision repair (BESIR) pathway. Recently, it was discovered that the distribution of Ntg1 shifts between these compartments when exposed to oxidative DNA damage agents that preferentially target each organelle (nucleus: hydrogen peroxide; mitochondria: hydrogen peroxide plus antimycin A). These results opened up a potentially important and novel mode of regulation for DNA repair proteins: dynamic compartmentalization. A novel technique, Q-SCAn has been developed to objectively and automatically quantify the distribution of a protein within cells as a replacement for the prior subjective and laborious manual scoring method. Q-SCAn was applied to *S. cerevisiae* cells expressing Ntg1–GFP and treated with oxidative DNA damage agents. Q-SCAn results were directly compared against manual scoring results from the same images. Results from this analysis revealed the large impact of subjectivity in the manual technique, and that Q-SCAn results do not agree with the manual scoring results. This discrepancy is likely due to subtle differences in what each technique measures. Closer analysis of the compartment GFP intensity data suggested that the hydrogen peroxide plus antimycin A condition caused a loss of Ntg1–GFP fluorescence in both compartments, and that this effect is stronger in mitochondria than in the nucleus. Thus, these chemical agents may be producing secondary effects which confound attempts to quantify the distribution of Ntg1–GFP.

Introduction

The *Saccharomyces cerevisiae* DNA repair protein Ntg1, a DNA *N*-glycosylase/AP lyase responsible for excising oxidized pyrimidines, localizes to both nuclei and mitochondria as part of the base excision and strand incision repair (BESIR) pathway. Recently, it was discovered that the distribution of Ntg1 shifts between these compartments when exposed to oxidative DNA damage agents (nucleus: hydrogen peroxide; mitochondria: hydrogen peroxide and the electron transport chain decoupler antimycin A (1)) that preferentially target each organelle (2,3). As BESIR is a critical pathway and many of the other components

of the pathway are shared by both compartments, these results opened up a potentially important and novel mode of regulation for DNA repair proteins: dynamic compartmentalization.

These findings spurred the development of an automated image analysis method to quantify the distribution of multi-compartment DNA repair proteins, Q-SCAn (4). After the Q-SCAn technique was developed, an important next step was to employ Q-SCAn to quantify the distribution of Ntg1–GFP in the same experimental setup in which dynamic compartmentalization was discovered, and to directly compare the Q-SCAn results to those resulting from the manual scoring technique. In the process, I found that these two methods produce different results, which is likely because they are measuring subtly different aspects of the distribution of the protein. Knowing this will enable the laboratory to better interpret the data being obtained. Additionally, I discovered that these differences are in part because oxidation-inducing treatments, especially the combination of hydrogen peroxide plus antimycin A, are likely confounding the microscopic quantification of GFP-tagged proteins.

Materials and Methods

Yeast strains, media, and growth conditions

Saccharomyces cerevisiae strains and plasmids used in this study are listed in Table 5-1. *S. cerevisiae* cells were cultured at 30 °C in rich YPD medium (1% yeast extract, 2% peptone, and 2% dextrose; plus 2% agar for plates) or synthetic defined drop-out media for selection (uracil$^-$). Plasmids or integrated constructs were transformed into cells by a modified lithium acetate method (5).

Table 5-1. Strains and plasmids used in this study.

Name	Genotype	Ref
Q-SCAn (DSC569)	*MATa ura3-52 leu2Δ1 his3Δ200 NLS–tdTom MTS–mCer*	(4)
DSC574	Q-SCAn (*NTG1–GFP*, *2μ*, *URA3*, *Amp^R*)	(4)

Oxidative stress induction and confocal microscopy

Overnight late–log phase (approximately 1×10^8 cells/mL) cultures of *S. cerevisiae* cells expressing both the Q-SCAn reporter and Ntg1–GFP were washed twice with H_2O,

resuspended in 1 mL H_2O, counted via hemocytometer, and adjusted to a density of 2–4 × 10^7 cells/mL. Cells were treated with 20 mM H_2O_2 or 20 mM H_2O_2 and 10 μg/mL antimycin A for 1 hour in the dark at 30 °C with aeration. Cells were washed twice with H_2O and imaged by confocal microscopy. Microscopy parameters were as previously described (4).

Data analysis

Microscopy images were randomly assigned to two groups. The first group was analyzed by the manual scoring method (2), with blinding, by two separate individuals, while both groups were analyzed using Q-SCAn (4). In brief, the manual scoring method involved examining the micrograph for each cell and determining whether GFP fluorescence overlapped with either the mitochondrial marker, the nuclear marker, or both. Q-SCAn involved measuring the fluorescence intensity within each compartment of a cell and computing a localization index to indicate the balance of localization between compartments. Statistical testing for shifts in localization using manual scoring were assessed by the χ^2 goodness-of-fit test, with the untreated frequencies as the expected values, with $\alpha = 0.05$. Statistical testing for shifts in localization using Q-SCAn were assessed by the Kruskal–Wallis test and the post-hoc multiple comparisons test with correction for ties (6), with $\alpha = 0.05$. A computerized approximation of the manual scoring technique used the Q-SCAn–measured compartment mean GFP intensity values and categorized cells based on whether that value was greater than 1.5 times the cytoplasm mean GFP intensity.

Results and Discussion

Inter-analyst variability affects manual scoring, but trends remain consistent

One of the concerns about the manual scoring technique previously employed (2,3) was subjectivity resulting from different people exercising their own judgment about the localization within each cell. To examine this issue, two different researchers scored

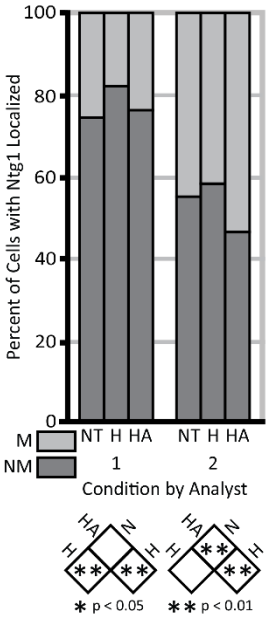

Figure 5-1. Manual scoring reveals dynamic compartmentalization of Ntg1–GFP in response to oxidative stress and demonstrates inter-analyst variability. *S. cerevisiae* cells co-expressing integrated NLS–tdTomato and MTS–mCerulean and 2μ Ntg1–GFP were exposed to 20 mM hydrogen peroxide (H) or 20 mM hydrogen peroxide and 10 μg/mL antimycin A (HA), or were untreated (NT), and were subsequently imaged by fluorescence microscopy. Two different analysts scored the same images for cells with mitochondrial-only (M), nuclear-only (none scored as this), or nuclear and mitochondrial (NM) localization of Ntg1. Pairwise statistical significance is indicated in the tables below the graph.

fluorescence micrographs of the same cells expressing Ntg1–GFP and exposed to hydrogen peroxide (nuclear oxidative stress) or hydrogen peroxide plus antimycin A (mitochondrial oxidative stress). Differences between the analysts are readily apparent, with approximately 20% of cells placed in different bins by each (Figure 5-1). However, what is striking is that the relative pattern of shifting localization is preserved between the two analysts, with approximately 5% of total cells moving from mitochondrial to nuclear and mitochondrial in the hydrogen peroxide condition, and then a shift back towards mitochondrial when antimycin A was included (Figure 5-1). However, the variability leads to some conflicting results. For analyst 1, the hydrogen peroxide treatment was significantly more nuclear than both the untreated and antimycin A–treated samples, but neither of the latter were significantly different from each other (Figure 5-1). For analyst 2, the hydrogen peroxide treatment was not significantly different from untreated, but the hydrogen peroxide plus antimycin A condition resulted in significantly more nuclear localization than both untreated and hydrogen peroxide alone (Figure 5-1). High variability in these manual scoring experiments is well known (2,3), and these data have been only obtained from two independent experiments. Therefore, the trends are more important than the exact differences in the quantification. These results suggest that through some combination of analyst bias, eyesight, computer screen, and environment, different analysts have different thresholds at which GFP intensity is determined to be in a compartment or not. While it is encouraging that the same general patterns appear regardless of these different baselines, the large differences should be taken as a warning that manually scoring these kind

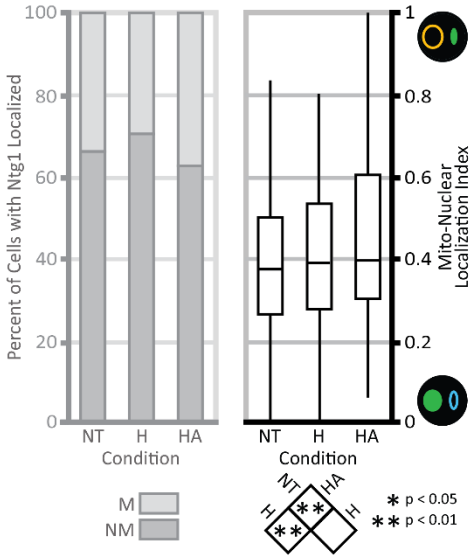

Figure 5-2. Q-SCAn measurement of Ntg1–GFP shows little change in localization. *S. cerevisiae* cells co-expressing integrated NLS–tdTomato and MTS–mCerulean and 2μ Ntg1–GFP were exposed to 20 mM hydrogen peroxide (H) or 20 mM hydrogen peroxide and 10 μg/mL antimycin A (HA), or were untreated (NT), and were subsequently imaged by fluorescence microscopy. Distribution of Ntg1–GFP was quantified by Q-SCAn (0 = mitochondrial to 1 = nuclear). Faded manual scoring of same data shown for comparison (see Figure 5-1). Horizontal black line indicates the sample median, white box the interquartile range, and vertical black line the adjacent range.

of results can introduce significant variability, and as such may not be easily reproducible. These results also highlight the need for objective, automated quantification techniques, such as Q-SCAn (4).

Ntg1 dynamic compartmentalization is not detected by Q-SCAn

The next concern was to find out what the Q-SCAn algorithm determines about the same cells that were scored manually. Q-SCAn results were qualitatively different from the manually scoring results. No significant difference in localization was observed for Ntg1–GFP between untreated and hydrogen peroxide–treated cells (Figure 5-2). And while the hydrogen peroxide plus antimycin A condition was significantly different from either of the others, the distribution was significantly more nuclear, the opposite of the trend that had been observed with manual scoring (Figure 5-2). Q-SCAn has been extensively validated against proteins with known localization and with Ntg1–GFP localization signal mutants (4), so it is unlikely that Q-SCAn is producing an erroneous indication of localization. On the same note, the fact that two different analysts working with the same data produced consistent patterns, and that these patterns were also consistent between biological replicates (not shown) and prior experiments (2,3) strongly suggests that manual scoring also measures a real effect. These results suggest that the two methods capture different aspects of protein localization.

Manual scoring simulated by Q-SCAn partially recapitulates manual scoring results

One possible explanation for why Q-SCAn and the manual scoring methods were

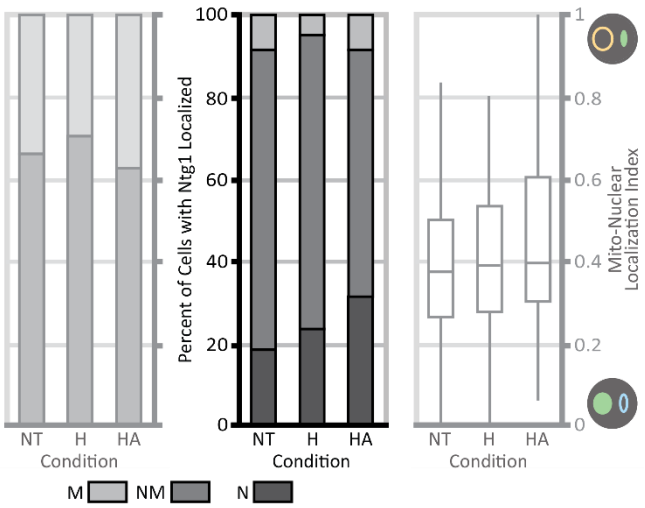

Figure 5-3. Simulated manual scoring method with Q-SCAn data partially recapitulates manual scoring results. *S. cerevisiae* cells co-expressing integrated NLS–tdTomato and MTS–mCerulean and 2µ Ntg1–GFP were exposed to 20 mM hydrogen peroxide (H) or 20 mM hydrogen peroxide and 10 µg/mL antimycin A (HA), or were untreated (NT), and were subsequently imaged by fluorescence microscopy. Cells were categorized as mitochondrial (M), nuclear (N), or nuclear and mitochondrial (NM) by thresholding the compartment mean GFP intensities by 1.5× the cytoplasmic mean GFP intensity. Faded manual scoring (left) and Q-SCAn results (right) of same data shown for comparison (see Figures 5-1 and 5-2).

producing different results is that they were measuring subtly different aspects of the cellular Ntg1 distribution. To test this idea, the compartment fluorescence intensity data which Q-SCAn generated was processed in a manner similar to the manual method, where the mitochondrial and nuclear GFP intensity within each compartment was compared to the cytoplasmic intensity and categorized as nuclear, mitochondrial, or nuclear and mitochondrial. This process revealed some very interesting information. The boundary between mitochondrial and nuclear/mitochondrial in the simulated method closely recapitulated the behavior of the manual scoring results, which show a nuclear shift in the hydrogen peroxide condition and a mitochondrial shift in the hydrogen peroxide plus antimycin A condition (Figure 5-3). However, the boundary between nuclear and nuclear/mitochondrial in the simulated method more closely reflected the behavior of the Q-SCAn results, in which the hydrogen peroxide plus antimycin A condition showed a nuclear shift (Figure 5-3). These results demonstrate that the two underlying quantification methodologies are measuring different aspects of the distribution of Ntg1 within cells.

Severe oxidative stress treatment reduces Ntg1–GFP fluorescence intensity

The finding that the two methodologies are measuring different aspects of the distribution of Ntg1 immediately raised the issue of what is actually being measured by these

techniques. The manual scoring technique is based on determining whether there is a detectable amount of GFP signal in either compartment. For this method to find a shift towards mitochondria, the nuclear signal would have to drop below the conceptual detection threshold in many cells. This would suggest that the nuclear signal was decreasing across the population in that condition.

To gain insight on this matter, the mean nuclear, mitochondrial, and cytoplasmic Ntg1–GFP intensities measured by Q-SCAn were examined, along with the intensities of the marker proteins. This analysis demonstrated that hydrogen peroxide treatment caused a slight reduction of both nuclear and mitochondrial GFP intensity (Figure 5-4A). However, a drastic reduction in GFP intensities accompanied antimycin A addition, and mitochondrial GFP intensity was affected more strongly than nuclear GFP (Figure 5-4A). However, the nuclear and mitochondrial marker proteins were not strongly affected (Figure 5-4B). Reactive oxygen itself may have a negative effect on GFP fluorescence, but it would seem that mCerulean, as a GFP derivative, would also be affected if that were the case. Oxidative stress suppresses classical nuclear import (7), which may explain the slight reduction observed with the nuclear marker protein tdTomato, but the effect is weak, and does not explain the loss of mitochondrial intensity. It would be worthwhile to compare Ntg1–GFP levels to those of an unrelated GFP fusion protein to determine whether loss of signal is a general effect for

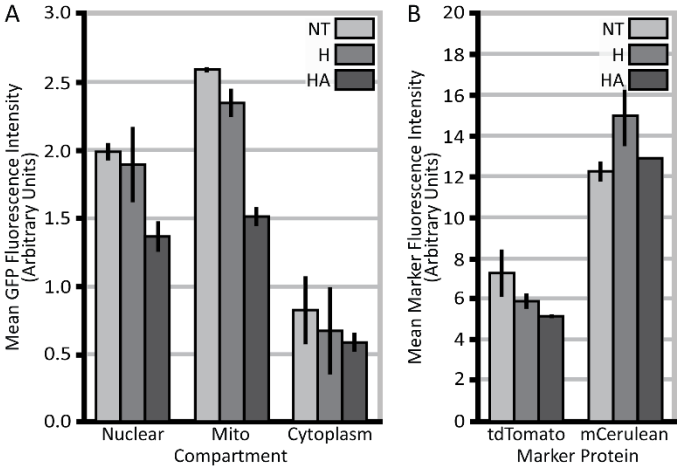

Figure 5-4. Oxidative stress reduces compartmental Ntg1–GFP mean fluorescence intensity. *S. cerevisiae* cells co-expressing integrated NLS–tdTomato and MTS–mCerulean and 2μ Ntg1–GFP were exposed to 20 mM hydrogen peroxide (H) or 20 mM hydrogen peroxide and 10 μg/mL antimycin A (HA), or were untreated (NT), and were subsequently imaged by fluorescence microscopy. Vertical bars indicate standard deviation over two experiments. A) Mean GFP fluorescence intensity in each cellular compartment averaged over the cell population. B) Mean marker protein (tdTomato, mCerulean) within their respective compartments (nucleus, mitochondria) averaged over the cell population.

all proteins (or at least GFP fusion proteins) or if the effect is specific to Ntg1. These results also indicate that the manual scoring technique may be highly sensitive to overall expression level, which Q-SCAn is largely unaffected by, as Q-SCAn was designed to be independent of protein expression level. These results also highlight the broader impacts of treating cells with hydrogen peroxide with antimycin A, and emphasize that these chemical treatments have nonspecific secondary effects.

These data broadly indicate that there are confounding factors introduced by using chemical oxidative stress–inducers to study DNA repair protein localization. Thus, there is a need for a new approach to introduce targeted and specific base damage in vivo. A novel derivative of a red fluorescent protein, KillerRed, has been employed to specifically target and carefully modulate ROS generation within the cell by controlling the amount of green light that the cells are exposed to (8,9). This protein would be a useful tool to introduce controlled, compartment-targeted ROS, reducing the potential side effects from whole-cell exposures.

Acknowledgements

I would like to thank the Emory University Integrated Cellular Imaging Core of the Winship Cancer Institute (National Cancer Institute Cancer Center Support Grant P30CA138292), especially Debbie Martinson and Dr. Adam Marcus. I would also like to thank the members of the Doetsch and Corbett laboratories, especially Dr. Dan Swartzlander, for his constructive feedback in designing these experiments and for acting as one of the analysts for the manual scoring comparison.

References

1. Potter VR, Reif AE. Inhibition of an electron transport component by antimycin A. J Biol Chem. 1952;194(1):287-297.

2. 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. doi:10.1128/mcb.01357-08.

3. 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. doi:10.1093/nar/gkq108.

4. Bauer NC, Corbett AH, Doetsch PW. Automated quantification of the subcellular localization of multicompartment proteins via Q-SCAn. Traffic. 2013;14(12):1200-1208. doi:10.1111/tra.12118.

5. Gietz RD, Woods RA. Transformation of yeast by lithium acetate/single-stranded carrier DNA/polyethylene glycol method. Methods Enzymol. 2002;350:87-96. doi:10.1016/S0076-6879(02)50957-5.

6. Zar JH. Biostatistical Analysis. 4th edition. Upper Saddle River, NJ: Prentice Hall; 1999. 195-199, 223-225.

7. Kodiha M, Tran D, Qian C, Morogan A, Presley JF, Brown CM, Stochaj U. Oxidative stress mislocalizes and retains transport factor importin-alpha and nucleoporins Nup153 and Nup88 in nuclei where they generate high molecular mass complexes. Biochim Biophys Acta. 2008;1783(3):405-418. doi:10.1016/j.bbamcr.2007.10.022.

8. Bulina ME, Chudakov DM, Britanova OV, Yanushevich YG, Staroverov DB, Chepurnykh TV, Merzlyak EM, Shkrob MA, Lukyanov S, Lukyanov KA. A genetically encoded photosensitizer. Nat Biotechnol. 2006;24(1):95-99. doi:10.1038/nbt1175.

9. Roy A, Carpentier P, Bourgeois D, Field M. Diffusion pathways of oxygen species in the phototoxic fluorescent protein KillerRed. Photochem Photobiol Sci. 2010;9(10):1342-1350. doi:10.1039/c0pp00141d.

Chapter 6
DISCUSSION

The work described in this dissertation examines a novel, rapid mode of regulation for the critical base excision and strand incision repair (BESIR) pathway whereby nuclear or mitochondrial base lesions are responsible for recruiting BESIR proteins to nuclei or mitochondria, respectively (Figure 1-13). This work contributes to our understanding of the role of protein localization in regulating BESIR and the mechanisms involved in generating different cellular distributions of BESIR proteins. A major accomplishment was the development of Q-SCAn (1) (Chapter 3), an automated and unbiased method to examine the distribution of proteins between subcellular compartments, a replacement for the laborious and subjective manual scoring method used previously (2,3). Q-SCAn provided previously inaccessible information about the distribution of the *Saccharomyces cerevisiae* transcription factor Yap1, and also made feasible the measurement of the relative contributions of the predicted localization signals of two *S. cerevisiae* BESIR glycosylases, Ntg1 and Ung1 (1). With respect to defining DNA lesions that trigger repair protein localization, evidence supporting the hypothesis that abasic sites are responsible for base damage–induced reactive oxygen species (ROS) signaling is presented in Chapter 4. Sequence analysis of the full range of human and *S. cerevisiae* BESIR proteins, described in Chapter 2, provided insight into these potential modes of regulation, including localization, contributing an important resource for future work. Other studies (Chapters 4 and 5) suggest a need for more specific damage-inducing agents to push this work forward as well as approaches to resolve differences between the prior subjective approach and Q-SCAn. These advances are summarized in Figure 6-1. The work described in this dissertation has laid the necessary and important groundwork for future investigations into the role of localization and signaling in the regulation of a critical DNA repair pathway.

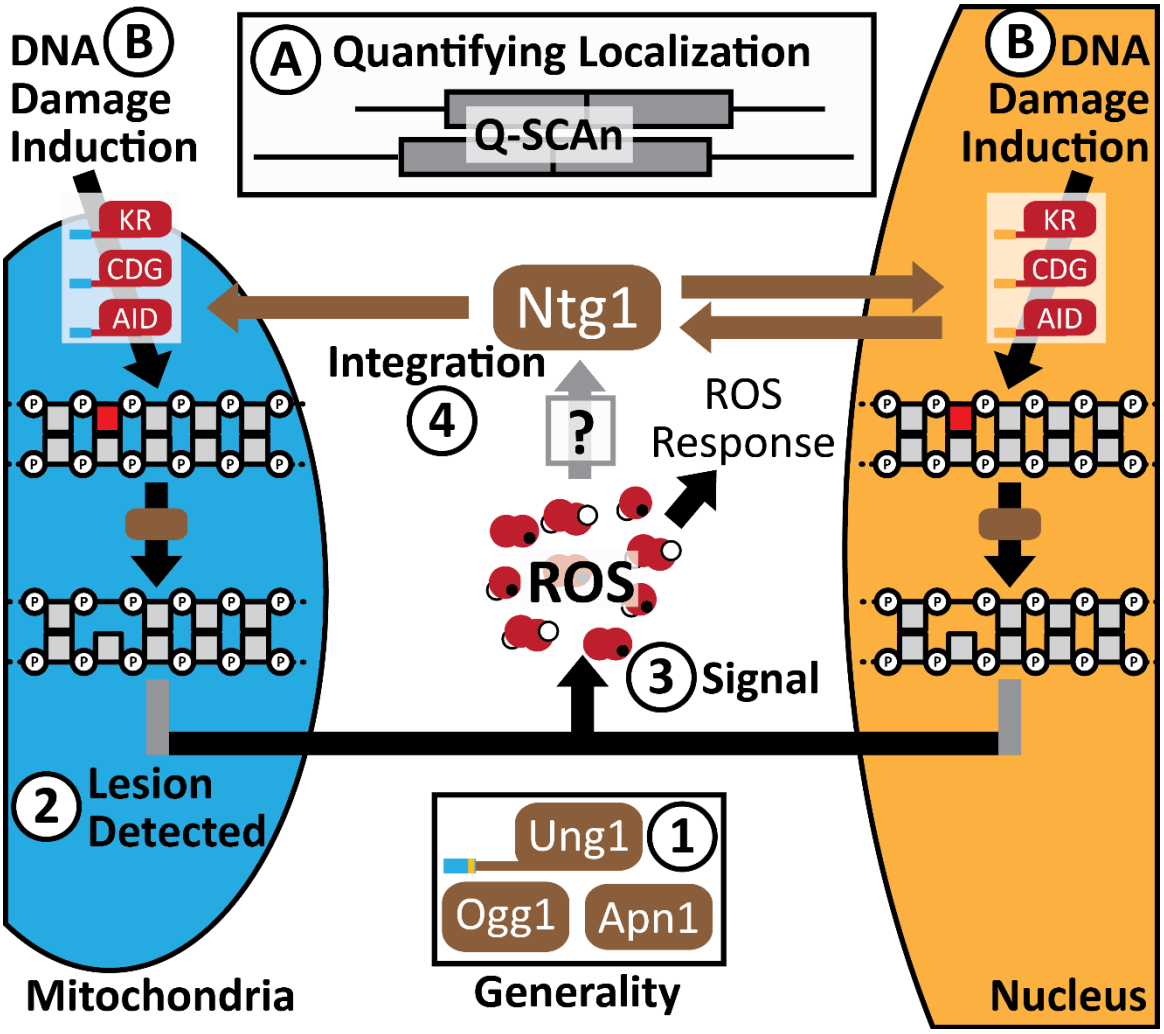

Figure 6-1. Dynamic compartmentalization model and summary of findings. 1) Ung1 localization sequences have been characterized, demonstrating an overlapping MTS (blue) and NLS (orange). 2 and 3) Evidence supports the hypothesis that abasic sites, and not base lesions, are responsible for base damage–induced ROS signaling, though any connection between ROS and dynamic compartmentalization remains speculation. 4) The mechanism directly modulating Ntg1 localization remains unknown. Supporting future studies, A) Q-SCAn was developed to replace the prior manual scoring method, and B) enzymes were identified which can be used to produce compartment specific DNA damage, KillerRed (KR), cytosine–DNA glycosylase (CDG), and activation-induced deaminase (AID).

BESIR: Critical Pathway with Many Unknowns

DNA is a critically important macromolecule, providing the instructions to build and maintain cells, tissues, organ systems, and organisms. However, the fidelity of this molecule is under continuous threat from both exogenous and endogenous sources. Base lesions are the most common and mutagenic type of damage that can be inflicted on DNA (Figure 1-2). Repairing such lesions is thus a very important process to prevent deleterious changes from accumulating in an organism's genome. Several alkyl lesions and UV photoproducts can be

directly reversed without additional harm to the DNA. Most base lesions, however, are processed by BESIR, a proposed merging of three interrelated repair pathways: base excision repair (BER), nucleotide incision repair (NIR), and ribonucleotide excision repair (RER) (Figure 1-4).

While the biochemical mechanism of the BESIR pathway has been well characterized, the regulation of the pathway has remained unclear. As reviewed in Chapter 1, some pathway components have had positive and negative interaction partners described, but precisely what functional interactions underlie these effects is still largely unknown. For example, human XRCC1 interacts with most BESIR proteins and coordinates their action (4-7), but no solid evidence has yet been found on the mechanism of action of this coordination. Additionally, while most BESIR proteins are localized to both the nucleus and mitochondria, very little work has gone into understanding the role of protein localization in regulating the BESIR pathway. The human BESIR glycosylase NTHL1 becomes mislocalized in a subset of cancers (8-10), suggesting a potential role in oncogenesis resulting from oxidative stress–induced mutagenesis. There is still much to learn about how these pathways are regulated, especially at the level of localization.

Dynamic Compartmentalization: A Novel Mode of DNA Repair Regulation

The current best-characterized example of regulating DNA repair activity through localization is the *S. cerevisiae* BESIR *N*-glycosylase/AP lyase Ntg1. As reviewed in Chapter 1, Ntg1 excises oxidized pyrimidines and is localized to both the nucleus and mitochondria. Ntg1 undergoes a shift in its cellular distribution upon different reactive oxygen species (ROS) induction conditions, accumulating more in the nucleus under hydrogen peroxide (increased extramitochondrial ROS), while accumulating more in mitochondria under hydrogen peroxide plus antimycin A treatment (increased mitochondrial and extramitochondrial ROS) (2). This behavior also appears to be dependent on the presence of DNA damage. This novel

mode of regulation has been termed dynamic compartmentalization.

Maintaining tight control over DNA repair protein levels is important to avoid the creation of adventitious lesions from nonspecific activity or accumulation of potentially lethal intermediates. As most BESIR proteins are shared by both nuclei and mitochondria, and both compartments have generally different repair needs due to their differing environments and replication cycles, regulation of BESIR activity by shifting the localization of the pathway components would be an efficient solution. Regulating localization allows a common pool of protein to be rapidly mobilized, saving the energetic costs of newly synthesizing separate proteins with dedicated destinations. For example, importing or exporting a protein to the nucleus converts just one GTP to GDP, while synthesis of a new protein the length of Ntg1 requires the hydrolysis of approximately 1600 high-energy phosphate bonds. As reviewed in Chapter 1, there are many ways by which cells modulate protein localization. Several human DNA repair proteins have since been found to shift localization as well (11,12). Investigating the localization of repair proteins is an important area of inquiry for understanding how DNA repair is regulated and how defects in this regulation can lead to human disease.

Four key questions were raised by the dynamic compartmentalization discovery (Figure 6-1[1-4]): 1) Is dynamic compartmentalization general to BESIR, or is it specific to one protein? 2) Which DNA lesions are responsible for inducing dynamic compartmentalization? 3) How is information about the lesion transmitted to the protein? 4) How is this information integrated at the protein to modulate its distribution?

Question 1: Generality of dynamic compartmentalization in BESIR

Previous studies showed relocalization of Ntg1 in response to DNA damage, but the question was raised as to whether it was specific to Ntg1 or if it was general to other repair proteins (Figure 6-1[1]). To address this question, an additional BESIR protein was selected for analysis. In support of this effort, I analyzed all *S. cerevisiae* BESIR proteins for predicted

localization sequences (Chapter 2). One protein that emerged as an optimal candidate for analysis was the uracil–DNA glycosylase Ung1 (13). Like Ntg1, Ung1 is localized to both the nucleus and mitochondria (14). Consensus targeting signals for both compartments are located at the N-terminus of the protein in a similar pattern to Ntg1 (1) (Figure 3-5). This arrangement of targeting signals could suggest a competition between localization signal receptors, and could indicate that Ntg1 and Ung1 share a similar mode of regulation. Ung1 differs from Ntg1 in several important respects. Ung1 is a monofunctional glycosylase, lacking AP lyase activity (13,15-17). In addition, Ung1 has a more nuclear steady-state localization compared with Ntg1 (1). Finally, Ung1 is 10-fold more abundant than Ntg1 (18,19). For these reasons, Ung1 was selected as the next protein to study for dynamic compartmentalization.

Two major challenges arose in pursuing Ung1 that blocked further investigation. First, the heavy nuclear localization of Ung1 rendered it unlikely that the manual scoring method could pick up a nuclear shift in localization. Second, chemical agents to induce cytosine deamination in vivo was determined to be impractical due to their high non-mutagenic toxicity. These challenges and the work to overcome them are discussed in depth below.

Questions 2 and 3: Detected lesions and signal transduction

For the localization of a protein to shift in response to DNA damage, a lesion must be detected (Figure 6-1[2]) and then a signal capable of reaching protein pools in other compartments must be generated (Figure 6-1[3]). The nature of this lesion and the details of signal transduction are the next critical questions for dynamic compartmentalization. There are a finite number of indicators a cell could use to detect DNA base damage: the lesions themselves, repair intermediate lesions, or released damage products. Detection of strand breaks, the most dangerous lesions to a cell, operates by sensing the lesions directly: double-strand breaks are detected by the MRN complex, which leads to ATM kinase pathway signaling; extended single-strand gaps are recognized by RPA, leading to ATR kinase pathway signaling

(20). Direct sensing of DNA lesions is a common theme in DNA damage signaling, so if a signaling mechanism for BESIR-repairable damage exists, as is suggested by the dynamic compartmentalization of Ntg1, the base lesions themselves or released base lesions are candidates for the proximate DNA damage signal. There is a precedent for glycosylases allosterically binding to the released lesion: The human 8-oxoguanine glycosylase/AP lyase OGG1 binds to free 8-oxoguanine, leading to the activation of small GTPases including Rho (21), Rac (22), and Ras (23). However, a signaling system specific to each type of lesion would be complicated. Due to overlapping specificities of BESIR glycosylases, and the need to recruit common downstream processing enzymes as well as the lesion-specific upstream enzymes, one possibility is that an abasic site, the common BESIR intermediate (24), is responsible for initiating the signal. In addition to being a common intermediate, each compartment maintains a basal level of its complement of glycosylases, ensuring that any lesions present will be quickly converted to abasic sites (25). Each abasic site sensed by the cell would lead to the generation of a signal which can propagate to a distant protein pool. The competing levels of signals from each compartment would regulate the localization of the BESIR proteins (26).

Identifying and characterizing the signals generated by base damage is another key question for understanding dynamic compartmentalization of Ntg1. Ideally, the signal should be rapidly generated and transduced. Double-strand breaks and single-strand gaps are signaled by kinase cascades through ATM and ATR, respectively (20), and single-strand breaks in human cells are signaled by PARP activity (27). Base damage signaling is still being elucidated. ROS levels increase in BESIR- and BESIR/NER-deficient *S. cerevisiae* cells, as well as in cells treated with the non-redox-active DNA alkylating agent MMS (2,28), suggesting that ROS is generated in response to base damage. Reactive oxygen signals can be propagated in solution, although for certain species like singlet oxygen, distance is limited in cytosol to less than 1 μm (29), and hydrogen peroxide—which other, poorly diffusible ROS are readily

converted to—can diffuse across membranes via aquaporins (30). Thus, base damage–induced ROS signaling is a good candidate for the dynamic compartmentalization signal. While there are indications that ROS may be involved in dynamic compartmentalization signaling, the evidence is not yet sufficient to demonstrate a causal link (2).

Base damage–induced ROS may be a specific consequence of cellular abasic site levels (28). To test this hypothesis, I collaborated with an Emory undergraduate student. Taking advantage of powerful yeast genetics, we generated *S. cerevisiae* strains with deficiencies in abasic site–processing and in combination with a knockout of the Ung1 glycosylase. This approach resulted in an array of strains containing differing steady state levels of abasic sites. Spontaneous mutation frequencies and ROS levels of these cells were measured (Figure 4-3). These results demonstrated that ROS levels in the repair-deficient strains are highly dependent on the generation of abasic sites, which strongly supports the hypothesis that a common product of DNA damage is resulting in the production of ROS. However, the mechanism sensing abasic sites and generating ROS remains unknown.

The next obvious question along this line of inquiry is to evaluate whether ROS is generated in response to nuclear abasic sites, mitochondrial abasic sites, or both. However, this experiment requires a reliable way to generate compartment-specific abasic sites. One way to achieve this goal would be to harness a mutant of uracil–DNA glycosylase which excises cytosine from DNA (31). When this protein was expressed in *S. cerevisiae*, many adventitious abasic sites were generated (32). This enzyme could be fused to a specific localization signal sequence, thereby targeting the damage to a specific compartment. Taken a step further, these experiments could be connected back to dynamic compartmentalization by expressing CDG or an oxidase in each compartment and examining any shift in localization. These experiments provide a direct test for the hypothesis that compartmental abasic sites and/or ROS are responsible for dynamic compartmentalization. If only one compartment

produces ROS and the other does not, the question of how the other compartment signals the presence of base damage remains open. However, it is possible to invoke dynamic compartmentalization even if only one compartment signals for damage, where the default is to favor the other compartment. If both compartments produce ROS, this raises the question of how compartment specificity is conferred. Given the short range of ROS generation, it could be a proximity-driven event, acting as a local concentration gradient to recruit proteins from near the compartment.

Question 4: Signal integration and localization shift

After the base damage signals have been generated, they must be integrated at the protein to invoke a change in its localization (Figure 6-1[4]). The types of mechanisms that cause proteins to change localization, reviewed in Chapter 1, provide some potential candidates. One possibility is that the base damage signal could cause the post-translational modification of the relocalizing protein to modulate the strength or accessibility of localization signals. Alternatively, another protein could be the target of the signal, which then binds to and either masks or contributes localization signals to the relocalizing protein.

Work on Ntg1 demonstrated that it is modified by the small ubiquitin-like modifier, SUMO, at a C-terminal lysine while in the nucleus in oxidative stress conditions (3). This finding suggested that sumoylation could play a role in dynamic compartmentalization. Four additional semi-redundant sumoylation sites have since been identified, and up to 5% of Ntg1 becomes sumoylated by 1–3 molecules of SUMO under oxidative stress (19). Although knocking out the five lysines eliminates sumoylation of Ntg1, no change was observed in dynamic compartmentalization, demonstrating that SUMO has no apparent role in regulating Ntg1 localization (19). Some preliminary data with MMS exposure and a SUMO-null Ntg1 mutant have indicated a role for Ntg1 sumoylation in the DNA damage checkpoint (unpublished data).

The results of BESIR protein sequence analysis have highlighted some potential sites in or near localization signals that could modulate localization if modified, but no particular pattern is apparent (Chapter 2). There also appears to be an enrichment of putative palmitoylation sites in the mitochondrial BESIR proteins, and it is tempting to speculate that these sites could target the proteins to the mitochondrial surface, which for some proteins is necessary before import (33-35). Sensitive mass spectrometry of BESIR proteins in various cellular conditions to determine the presence of post-translational modifications would be able to provide much of this information. Modifications in or near localization signals should take priority in a candidate-based approach. Not knowing which signaling pathways are involved will limit the ability to address this question for the time being, though obvious candidates to initially investigate would include the canonical ATM/ATR DNA damage response pathways.

Challenges to Studying BESIR Dynamic Compartmentalization

Two critical challenges must be overcome to make progress in addressing all of the questions surrounding dynamic compartmentalization (Figure 6-1[A,B]): quantifying protein localization and inducing compartment-specific DNA base damage. A third challenge impacts all scientific research and was encountered during this work: the state of scientific software.

Quantifying protein localization

One of the primary challenges to studying dynamic compartmentalization was the manual scoring method used to quantify the localization of Ntg1. In this method, confocal fluorescence micrographs were collected of cells expressing GFP-tagged proteins and stained for nuclear and mitochondrial DNA with DAPI, and for mitochondria with MitoTracker (2). Images were manually analyzed by a researcher for GFP signal overlapping with DAPI and/or MitoTracker, and thereby scored as nuclear, mitochondrial, or nuclear and mitochondrial (2) (Figure 1-13[A]). In essence, this method assesses whether either compartment in a cell has

GFP signal above some threshold of detection. There were several specific aspects of this method that were problematic. First, it is a subjective technique, and the theoretical threshold depends on the experience, expertise, and visual acuity of the experimentalist, the software and display hardware used, and the background lighting. Second, the time required to score sufficient cells (at least 200/sample) was prohibitive. Third, the binning approach is relatively insensitive to subtle shifts in protein distribution, and would have particular difficulty detecting a nuclear shift in a protein that was already highly concentrated within the nucleus, such as Ung1. Fourth, the marker dyes may stain nonspecific structures or whole cells, rendering a variable number of cells, and occasionally whole samples, unscorable. These issues prevented rapid progress in studying dynamic compartmentalization.

In light of these problems, I conceived and developed the automated image analysis method, quantitative subcellular compartmentalization analysis (Q-SCAn) (1) (Chapter 3; Figure 6-1[A]). In brief, Q-SCAn relies on the malleable genetic system of *S. cerevisiae* to express nuclear and mitochondrial fluorescent marker proteins (NLS–tdTomato and MTS–mCerulean, respectively) along with the GFP-tagged protein of interest. Confocal fluorescence micrographs of these cells are analyzed by an algorithm that identifies cells from the brightfield image and the compartments within each cell from the fluorescent marker proteins, then quantifies the mean GFP fluorescence within each compartment. Finally, Q-SCAn can be used to calculate a "localization index" with a value of 0 for proteins entirely in the mitochondria and a value of 1 for proteins entirely in the nucleus (Figure 3-1). An important aspect of Q-SCAn is that the methodology is implemented within software designed for high-throughput image analysis and which has an active support community, CellProfiler. Furthermore the algorithm is readily modifiable. Q-SCAn addresses each of the weaknesses of the manual method: it is objective, automated, can detect subtle shifts in distribution, and relies on consistent markers with a high signal-to-noise ratio.

To validate Q-SCAn, I examined relocalization of the oxidative stress response transcription factor Yap1. Results of this analysis not only reproduced data which was collected by a manual scoring method but also revealed information about the population distribution that was not accessible through manual scoring (Figure 3-3). In addition, Q-SCAn allowed a detailed characterization of the contributions of localization signals to the intracellular distribution of Ntg1 (Figure 3-4). Q-SCAn was then employed to define functional targeting signals in Ung1 (Figures 3-5, 6-1[1]). This proof of principle study demonstrates the power of automated quantification methods to extract useful information from microscopy images.

Q-SCAn has potential impacts beyond quantifying the distribution of proteins among several samples. Along with continued developments in the area of high-throughput experimental apparatuses, Q-SCAn could make it possible to run a full screen of a gene library or knockout collection to identify genes that modulate localization of a target protein, an approach that was not previously amenable to high-throughput screening. Q-SCAn can also be readily expanded to other model systems, different sets of compartments, other compartment markers, or to 3D imaging. Q-SCAn could also be applied to provide clinical diagnostics, as mislocalization of proteins is associated with disease processes including cancer, autoimmune disorders, and degenerative disorders.

Induction of compartment-targeted, specific base lesions

An idealized DNA damage and repair experiment would include the ability to introduce defined lesions in a controlled manner. In vitro, this can be easily accomplished by synthesizing DNA with a modified base, or by treating DNA with an agent producing a known range of lesions. Certain kinds of in vivo experiments can also take advantage of synthesized lesions, transfected or transformed into cells (36). No such precision tools have generally been available for genomes in vivo. Instead, the field relies on exogenous chemical treatments

to induce the desired damage.

Some agents are highly specific, such as the topoisomerase 2 inhibitor etoposide, which prevents topoisomerase from ligating the DNA after cleaving it (37). Others have varying degrees of side reactions, like hydrogen peroxide. While hydrogen peroxide produces other ROS that can react with almost any biological molecule, and thus can have effects other than DNA damage, the toxicity of hydrogen peroxide is relatively low compared with its mutagenicity. For many purposes, the side effects can be safely ignored because they would be fairly broad-spectrum and would not otherwise be expected to greatly affect the DNA damage–relevant endpoints, such as mutagenesis. However, oxidative treatment with hydrogen peroxide may not ultimately be suitable for DNA damage–induced localization studies, as it appears to have previously uncharacterized confounding effects on either the GFP tag or the whole protein (Figure 5-4A). There are also some agents that can introduce specific damage in vitro, but have overwhelmingly toxic effects in vivo, such as the cytosine deamination agent bisulfite. An attempt to use bisulfite to induce cytosine deamination to generate Ung1-repaired lesions was not successful (Figure 4-1). Another challenge with using chemical agents like hydrogen peroxide and MMS is that they can produce a range of lesions (38,39). Lesions produced by these agents can be recognized by multiple repair proteins, confounding attempts to study a single substrate or enzyme. While the lesions induced by these agents have been defined in vitro, there is little knowledge of the precise range of lesions produced inside the cell. To further complicate matters, some of the oxidative reactions on DNA, for instance, depend on the redox state of the surrounding medium (38). Thus, methods to induce and study the consequences of specific lesions in vivo are not currently available.

In addition to the challenge of introducing specific lesions, the situation is even more complicated if the goal is to create different damage conditions in different parts of the cell. Such compartment-specific damage is particularly desirable in order to study dynamic

compartmentalization (Figure 6-1B). When Ntg1 dynamic compartmentalization was identified, cellular ROS measurements indicated that moderate hydrogen peroxide treatment elevated whole-cell, but not mitochondrial, ROS levels, suggesting that hydrogen peroxide treatment alone would primarily impact the nucleus (2). When hydrogen peroxide was supplemented with the electron transport chain inhibitor antimycin A (40), however, ROS measurements increased in both mitochondria and in the whole cell, indicating that oxidative stress was primarily increasing in mitochondria (2). Increased frequency of mitochondrial lesions has been observed with this treatment (41). While this evidence is strongly suggestive, no direct measure had been made of the specific spectrum of lesions generated in nuclei or mitochondria, as no method existed to do so. To address this issue, a collaborative effort to employ glycosylase-mediated base release followed by liquid or gas chromatography and isotope dilution tandem mass spectrometry (42) has been initiated. This method quantifies specific base lesions within genomic DNA with high sensitivity (42). Preliminary results from fractionated *S. cerevisiae* indicated that hydrogen peroxide treatment damages DNA in both nuclei and mitochondria, but in slightly different ways. Continued work on this front will reveal much about the mechanisms of these agents, and will provide a test of the assumption of organellar preferences for these agents.

The challenges to applying chemical agents to achieve compartment-specific damage strongly indicate a need for a different approach to introduce specific lesions to specific compartments in vivo without major confounding effects. For targeting, very few approaches are better than an internally expressed protein containing strong localization signals and binding domains because they operate using the same highly efficient machinery that cells evolved to localize proteins and keep distinct pools separate from one another. Targeting of small molecules is often based on some general characteristic of the targeted compartment, but such characteristics typically are not entirely unique. For example, MitoTracker dye

depends on the oxidizing environment of the mitochondria to react with mitochondrial proteins and thus accumulate there. However, MitoTracker can also label the endoplasmic reticulum and other lipid-bound organelles, as is described in the specifications and observed in my studies (1). A preference for one compartment over another is often the best that can be done for such compounds. Proteins, however, can be very specifically targeted by localization signals and binding domains, a fact exploited in Q-SCAn (1).

One approach is to target specific DNA-modifying enzymes to specific cellular compartments. Several candidate enzymes are available to introduce targeted DNA damage. The most common ones used today are the homing endonucleases, zinc finger endonucleases, TALENs, and CRISPR, which introduce very specific double-strand breaks in vivo (43-45). However, the goal of this work is to study DNA base damage. Three proteins that are well suited for this task were identified: cytosine–DNA glycosylase, KillerRed, and activation-induced deaminase (Figure 6-1[B]).

The first candidate protein is a mutant of the human uracil–DNA glycosylase, which converted it into a cytosine–DNA glycosylase (31). When expressed in *S. cerevisiae*, this CDG induced the formation of many adventitious abasic sites in DNA (32). This enzyme therefore allows a controlled, direct increase of abasic site levels within the nucleus or mitochondria very specifically. Introducing abasic sites directly into DNA is not an approach that was previously available.

A second protein, KillerRed, is a novel derivative of a red fluorescent protein which produces singlet oxygen when excited by green light (46,47). While the lesions induced by KillerRed are still indirect, via generated ROS, the ROS will be targeted to a specific compartment, and the amount of ROS KillerRed produces can be carefully tuned.

A hypermutagenic activation-induced deaminase has been identified in sea lamprey and characterized in *S. cerevisiae* (48), allowing cytosine deamination damage (e.g. uracil) to

be introduced in vivo. As with the other enzymes, this protein can be targeted to nuclei or mitochondria, replacing bisulfite.

The only lesion type not accounted for by this set of enzymes is alkylation damage. No obvious enzymatic chemistry exists that would directly generate alkylation damage. Creative targeting of KillerRed may be able to generate reactive lipid peroxidation products, which generate exocyclic etheno adducts (49). With the three available proteins, it will be possible to introduce fairly specific base lesions, to target them to one organelle, and to precisely control the amount of damage in vivo.

For all their challenges, chemical agents have flexibility that molecular biological constructs do not have. Chemical agents can be applied to any cell without the additional up-front preparation of genetically modifying each cell line. Chemical agents also have an advantage in evenly affecting the cell population and allows direct control over exposure levels, as they do not depend on the inherent heterogeneity of gene expression. However, these costs are well worth the benefits of enhanced specificity and reduced secondary effects. Targeted damage enzymes have the broad potential to enable more specific and controlled DNA damage experiments in vivo than could be accomplished with chemical agents.

One hurdle to implementing these protein DNA damage tools will be refining the conditions for using KillerRed, which has never been used in *S. cerevisiae*. Optimizing the amount of light exposure will take some effort, though fortunately some reports of KillerRed use a simple visible light lamp, suggesting that this will not be technically challenging. Detailed protocols and guidelines are available for mammalian cells (50), which can be adapted. Some studies have reported fusing KillerRed to a DNA-binding protein to enhance proximity to the DNA in human cells (51). Such an approach could be helpful here, although yeast nuclei are much smaller than human nuclei, so the benefit of enhanced targeting may be reduced. The other challenge is that KillerRed is a fluorescent protein which is primarily red but has

some minor green fluorescence. These fluorescence properties may interfere with both detection of the Q-SCAn marker protein tdTomato and the measurement of a GFP-fused target protein (52). Potential ways to deal with this interference include swapping the marker colors or even dropping the tdTomato and relying only on KillerRed to provide compartment information, and crosstalk analysis and linear unmixing could be applied to remove erroneous green signal from the GFP channel. While KillerRed will require some optimization, this process should not be especially difficult to accomplish, and the result will be a versatile tool to control the generation of ROS in vivo.

State of scientific software

A key component of Q-SCAn was linking data collection to data analysis. Software tools are steadily becoming a bigger and more important part of everyday biological research, from online databases to commercial software packages to published scripts. However, much scientific software suffers from poor user interfaces, few standards, and lack of support or regular maintenance. Even a large commercial scientific software package like LASERGENE has an ancient user interface which, while powerful, is not particularly easy to use. These problems are even more apparent with specialty software created to implement custom algorithms for published research. As discussed in Chapter 2, many of the published sequence analysis algorithms are scattered throughout the literature, and those which are supposed to be available online as public resources may end up abandoned or broken within just a few years. Those which do work require manual entry of a text sequence, and may not tolerate the helpful breaks and line numbers introduced for readability in common sequence displays. The output is often not standardized either. Transcribing the relevant results from the predictors used in Chapter 2 required many hours of manual work per protein. Even ImageJ (53), one of the standard image analysis programs developed by the NIH, has been relatively stagnant for years (though work on a modern update is underway), and attempting to

automate ImageJ using its macro scripting language is a challenge because the documentation is spotty and the editor is little more than a text editor. These problems form an impediment to the broad i of the underlying useful algorithms.

These general issues are not entirely unexpected. The scientists developing the small specialized scripts and programs have, as a primary goal, obtaining results for their own research projects and developing new algorithms. Once the research is made available, the demands on their time incentivize working on the next thing rather than improving or maintaining the previous thing. Writing software for a broad audience and supporting it takes time and effort which does not directly contribute to one's own research, and very little funding is available for ongoing support of such software. Additionally, sometimes these programs are written by biologists with little experience in software design, or by computational biologists with little experience of the bench scientists who will be using it. These systematic issues will need to be addressed in order for the situation to improve.

One of the highlights of good scientific programming is CellProfiler, the image analysis program created and maintained by the Broad Institute (54). They have made a steady effort to improve the software and make it more user-friendly. Support is actively available to researchers wishing to use the program. These efforts were one of the major reasons that CellProfiler was selected for Q-SCAn, and this is also one of the advantages Q-SCAn has over a competing software, Cell-ID (55), which is a custom-built program in a low-level programming language and with relatively little documentation or support available. Fortunately, the NIH recognizes the need for improving scientific software and has limited funding opportunities available specifically for supporting core software infrastructure, which CellProfiler has taken advantage of. Small laboratories primarily doing bench research do not often have this luxury. This situation is unfortunate, because the impact of their work will be much smaller than it should be, slowing down the overall pace of biological research.

Revisiting Dynamic Compartmentalization

The observation that Ntg1 shifted localization towards a compartment undergoing oxidative DNA damage opened up base damage–induced dynamic compartmentalization as a mechanism regulating BESIR. The goal of this dissertation was to expand our understanding of this novel regulatory mechanism. As the manual scoring method employed with Ntg1 had several limitations, Q-SCAn was developed to provide a more robust, objective measure of the distribution of a protein between the nucleus and mitochondria. However, in the first experiment to directly compare the results from the manual scoring method and Q-SCAn (Chapter 5), the two methods did not agree with each other: In the hydrogen peroxide plus antimycin A condition, the manual scoring method recorded a shift towards mitochondria while Q-SCAn recorded a significant shift towards the nucleus (Figure 5-2). These mutually incompatible results required an explanation.

An important clue to explain the inconsistency between methods came from simulating manual scoring using the Q-SCAn data (Figure 5-3). This simulation suggested that the discrepancy was related to a reduced ability to manually detect nuclear-localized GFP compared with mitochondrial-localized GFP in the hydrogen peroxide plus antimycin A condition. An analysis of the compartment GFP intensities in the same condition demonstrated further that GFP intensities were drastically reduced in both nuclei and mitochondria (Figure 5-4). These results suggest that the mitochondrial shift observed in hydrogen peroxide plus antimycin A in the previous studies was caused by overall loss of GFP fluorescence intensity, rather than by a shift in protein distribution. Note that the nuclear shift consistently observed under hydrogen peroxide treatment would not be affected by this problem. While Q-SCAn results were not significant for this condition, the inherent variability may simply require more cells to be collected before the shift is detectable by Q-SCAn. Fortunately, because Q-SCAn enables rapid, automated analysis of images, an increase in the number of

cells analyzed should not be a challenge.

The above findings lead to the question of why GFP intensity is so severely depressed by hydrogen peroxide plus antimycin A treatment. This condition could be directly affecting GFP fluorescence, or it may indicate suppression of Ntg1 levels. The former is unlikely because the fluorescent marker proteins were largely unaffected. The effect could also be due to other biological impacts of antimycin A–driven inhibition of the electron transport chain. These impacts include disruption of the mitochondrial proton gradient necessary for mitochondrial protein import (56-59), energy starvation (60), and disruption of iron-sulfur cluster biosynthesis (59). However, prior work suggests that antimycin A alone has little effect on oxidative stress or Ntg1 (2). Therefore oxidative stress may synergistically interact with one or more of the other effects of antimycin A treatment.

Another aspect of the dynamic compartmentalization model which deserves closer examination is the dependence of localization shifts on compartmental DNA damage. The evidence supporting this part of the model is that the mitochondrial Ntg1–GFP intensity was reduced in respiration-deficient cells lacking mitochondrial DNA (ρ^0), and that hydrogen peroxide with antimycin A did not enhance localization to mitochondria (2). As noted above and reviewed in Chapter 1, mitochondrial protein import heavily depends on the electromotive force provided by the proton gradient across the inner mitochondrial membrane which is generated by the mitochondrial electron transport chain (56-58). Mitochondrial DNA encodes gene products critical to the proper function of the electron transport chain (61,62). Without the expression of these gene products the proton gradient collapses, and thus reduces import of proteins into the mitochondrial matrix (59). Untreated ρ^0 cells in the previous study showed a highly nuclear steady-state localization compared with untreated ρ^+ cells, and the Ntg1 distribution in ρ^0 cells did not change in response to either treatment (2). These results are consistent with loss of mitochondrial protein import capacity.

An elegant experiment to test this hypothesis is possible using a hyperactive mutant of the F_1F_0-ATPase, *ATP1-111*, which restores the proton gradient to high levels in ρ^0 cells (59,63) and respiratory-deficient cells containing defective copies of mitochondrial DNA (ρ^-) (64). Examining the localization of Ntg1 in these strains provides a direct test of both the DNA damage–dependence of dynamic compartmentalization and of whether the observed loss of dynamic compartmentalization is due to defective mitochondrial protein import.

Despite these issues, evidence is growing that the localization of Ntg1 is highly regulated. A recent report demonstrated that mitochondrial localization of Ntg1 is dependent on active Tel1 and/or Mec1, the yeast orthologs of ATM and ATR, respectively (65). These researchers proposed that sumoylated Ntg1 is localized to the nucleus and that the DNA damage response kinases maintain a pool of non-sumoylated Ntg1 (65). However, as previously discussed, unpublished results suggest that sumoylation has no effect on Ntg1 localization (19). Intriguingly, two strongly predicted phosphorylation sites, S^9 and T^{23}, straddle the Ntg1 bipartite NLS (Figure 2-1), both of which are likely substrates for the Chk1 kinase (PHOSIDA motif matcher (66)). Chk1 is a major downstream effector of Mec1 (67). Thus, I propose that Mec1 activation results in the phosphorylation of Ntg1 within its NLS, reducing the binding affinity of the import receptor importin α for the NLS, thereby reducing nuclear import of Ntg1 (Figure 1-7[1]). (Note that this information would not have been available without combining the results of multiple sequence analysis tools, a prime example of how important it is that these analysis tools are easy to use, readily available, and continuously maintained.) Importantly, the human homolog of NTHL1 has a strongly predicted phosphorylation site at S^{22} (Figure 2-1), which is also a predicted substrate for human CHK1 (PHOSIDA motif matcher (66)). This residue is near one of the predicted NLS sequences, suggesting that the mechanism is conserved. However, analyzing the phosphomimetic mutant sequence suggests that mitochondrial localization would also be affected by reducing the net charge of the MTS (iPSORT

(68)). Mec1/ATR signaling is invoked when extensive single-stranded DNA is present as well as during homologous recombination (20), indicating the presence of severe DNA damage. Thus, this proposed mechanism may be part of a broader regulatory strategy to suppress base lesion repair in favor of resolving the more severe lesions first. This model can be immediately tested by creating phosphomimetic and phosphodeficient mutants of Ntg1. Overall, These findings demonstrate that Ntg1 does undergo dynamic compartmentalization, though it may take a different form than previously thought.

Conclusion

My dissertation work has developed the methods necessary to vigorously pursue the questions surrounding dynamic compartmentalization and has provided important insights into base damage signaling and repair protein localization. I have developed a novel image analysis technique to measure protein distribution (Q-SCAn) to accurately measure shifts in localization. I performed and recorded an in-depth sequence analysis of the BESIR proteins, which has produced interesting insights and pointed to clues about BESIR protein regulation in several instances so far. I have characterized the Ung1 localization signals and contributed to the understanding of the localization signals of Ntg1 and the regulation of Yap1 in response to oxidative stress. I have surveyed the literature to identify and organize the major pathways responsible for modulating relocalization. I have also provided evidence that abasic sites are responsible for the generation of base damage–induced ROS. This work has yielded important insights, but much more work remains. The most crucial next step is to employ targeted DNA enzymes to generate compartment-specific DNA damage, and Q-SCAn should be used to determine whether dynamic compartmentalization is occurring for all of the nucleomitochondrial BESIR proteins: Ntg1, Ung1, Ogg1 and Apn1. Obtaining these results will provide a foundation from which the role of localization in the regulation of BESIR can be fully investigated.

References

1. Bauer NC, Corbett AH, Doetsch PW. Automated quantification of the subcellular localization of multicompartment proteins via Q-SCAn. Traffic. 2013;14(12):1200-1208. doi:10.1111/tra.12118.

2. 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. doi:10.1128/mcb.01357-08.

3. 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. doi:10.1093/nar/gkq108.

4. Hanssen-Bauer A, Solvang-Garten K, Akbari M, Otterlei M. X-ray repair cross complementing protein 1 in base excision repair. Int J Mol Sci. 2012;13(12):17210-17229. doi:10.3390/ijms131217210.

5. Sokhansanj BA, Rodrigue GR, Fitch JP, Wilson DM, 3rd. A quantitative model of human DNA base excision repair. I. Mechanistic insights. Nucleic Acids Res. 2002;30(8):1817-1825. doi:10.1093/nar/30.8.1817.

6. Campalans A, Marsin S, Nakabeppu Y, O'Connor T R, Boiteux S, Radicella JP. XRCC1 interactions with multiple DNA glycosylases: a model for its recruitment to base excision repair. DNA Repair. 2005;4(7):826-835. doi:10.1016/j.dnarep.2005.04.014.

7. Vidal AE, Boiteux S, Hickson ID, Radicella JP. XRCC1 coordinates the initial and late stages of DNA abasic site repair through protein-protein interactions. EMBO J. 2001;20(22):6530-6539. doi:10.1093/emboj/20.22.6530.

8. Goto M, Shinmura K, Igarashi H, Kobayashi M, Konno H, Yamada H, Iwaizumi M, Kageyama S, Tsuneyoshi T, Tsugane S, Sugimura H. Altered expression of the human base excision repair gene *NTH1* in gastric cancer. Carcinogenesis. 2009;30(8):1345-1352. doi:10.1093/carcin/bgp108.

9. Koketsu S, Watanabe T, Nagawa H. Expression of DNA repair protein: MYH, NTH1, and MTH1 in colorectal cancer. Hepatogastroenterology. 2004;51(57):638-642.

10. Karahalil B, Bohr VA, De Souza-Pinto NC. Base excision repair activities differ in human lung cancer cells and corresponding normal controls. Anticancer Res. 2010;30(12):4963-4971.

11. da Costa NM, Hautefeuille A, Cros MP, Melendez ME, Waters T, Swann P, Hainaut P, Pinto LF. Transcriptional regulation of thymine DNA glycosylase (TDG) by the tumor suppressor protein p53. Cell Cycle. 2012;11(24):4570-4578. doi:10.4161/cc.22843.

12. van Loon B, Samson LD. Alkyladenine DNA glycosylase (AAG) localizes to mitochondria and interacts with mitochondrial single-stranded binding protein (mtSSB). DNA Repair. 2013;12(3):177-187. doi:10.1016/j.dnarep.2012.11.009.

13. Percival KJ, Klein MB, Burgers PM. Molecular cloning and primary structure of the uracil-DNA-glycosylase gene from *Saccharomyces cerevisiae*. J Biol Chem. 1989;264(5):2593-2598.

14. Chatterjee A, Singh KK. Uracil-DNA glycosylase–deficient yeast exhibit a mitochondrial mutator phenotype. Nucleic Acids Res. 2001;29(24):4935-4940. doi:10.1093/nar/29.24.4935.

15. Crosby B, Prakash L, Davis H, Hinkle DC. Purification and characterization of a uracil-DNA glycosylase from the yeast, *Saccharomyces cerevisiae*. Nucleic Acids Res. 1981;9(21):5797-5809. doi:10.1093/nar/9.21.5797.

16. Impellizzeri KJ, Anderson B, Burgers PM. The spectrum of spontaneous mutations in a *Saccharomyces cerevisiae* uracil-DNA-glycosylase mutant limits the function of this enzyme to cytosine deamination repair. J Bacteriol. 1991;173(21):6807-6810.

17. Burgers PM, Klein MB. Selection by genetic transformation of a *Saccharomyces cerevisiae* mutant defective for the nuclear uracil-DNA-glycosylase. J Bacteriol. 1986;166(3):905-913.

18. Ghaemmaghami S, Huh WK, Bower K, Howson RW, Belle A, Dephoure N, O'Shea EK, Weissman JS. Global analysis of protein expression in yeast. Nature. 2003;425(6959):737-741. doi:10.1038/nature02046.

19. Swartzlander DB. Regulation of base excision repair in response to genotoxic stress [Dissertation]. Atlanta, GA: Emory University; 2012.

20. Goodarzi AA, Jeggo PA. The repair and signaling responses to DNA double-strand breaks. Adv Genet. 2013;82:1-45. doi:10.1016/b978-0-12-407676-1.00001-9.

21. Luo J, Hosoki K, Bacsi A, Radak Z, Hegde ML, Sur S, Hazra TK, Brasier AR, Ba X, Boldogh I. 8-Oxoguanine DNA glycosylase-1-mediated DNA repair is associated with Rho GTPase activation and alpha-smooth muscle actin polymerization. Free radical biology & medicine. 2014;73:430-438. doi:10.1016/j.freeradbiomed.2014.03.030.

22. Hajas G, Bacsi A, Aguilera-Aguirre L, Hegde ML, Tapas KH, Sur S, Radak Z, Ba X, Boldogh I. 8-Oxoguanine DNA glycosylase-1 links DNA repair to cellular signaling via the activation of the small GTPase Rac1. Free radical biology & medicine. 2013;61c:384-394. doi:10.1016/j.freeradbiomed.2013.04.011.

23. German P, Szaniszlo P, Hajas G, Radak Z, Bacsi A, Hazra TK, Hegde ML, Ba X, Boldogh I. Activation of cellular signaling by 8-oxoguanine DNA glycosylase-1–initiated DNA base excision repair. DNA Repair. 2013;12(10):856-863. doi:10.1016/j.dnarep.2013.06.006.

24. 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. doi:10.1016/j.biochi.2003.11.003.

25. Huh W-K, Falvo JV, Gerke LC, Carroll AS, Howson RW, Weissman JS, O'Shea EK. Global analysis of protein localization in budding yeast. Nature. 2003;425(6959):686-691. doi:10.1038/nature02026.

26. 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. Prog Mol Biol Trans Sci: Mechanisms of DNA Repair. Oxford, UK: Academic Press; 2012. 93-121.

27. Woodhouse BC, Dianov GL. Poly ADP-ribose polymerase-1: an international molecule of mystery. DNA Repair. 2008;7(7):1077-1086. doi:10.1016/j.dnarep.2008.03.009.

28. Rowe LA, Degtyareva N, Doetsch PW. DNA damage-induced reactive oxygen species (ROS) stress response in *Saccharomyces cerevisiae*. Free Radical Biol Med. 2008;45(8):1167-1177. doi:10.1016/j.freeradbiomed.2008.07.018.

29. Kim S, Tachikawa T, Fujitsuka M, Majima T. Far-red fluorescence probe for monitoring singlet oxygen during photodynamic therapy. J Am Chem Soc. 2014;136(33):11707-11715. doi:10.1021/ja504279r.

30. Bienert GP, Chaumont F. Aquaporin-facilitated transmembrane diffusion of hydrogen peroxide. Biochim Biophys Acta. 2014;1840(5):1596-1604. doi:10.1016/j.bbagen.2013.09.017.

31. Kavli B, Slupphaug G, Mol CD, Arvai AS, Peterson SB, Tainer JA, Krokan HE. Excision of cytosine and thymine from DNA by mutants of human uracil-DNA glycosylase. EMBO J. 1996;15(13):3442-3447.

32. Auerbach P, Bennett RA, Bailey EA, Krokan HE, Demple B. Mutagenic specificity of endogenously generated abasic sites in *Saccharomyces cerevisiae* chromosomal DNA. Proc Natl Acad Sci USA. 2005;102(49):17711-17716. doi:10.1073/pnas.0504643102.

33. Darshi M, Trinh KN, Murphy AN, Taylor SS. Targeting and import mechanism of coiled-coil helix coiled-coil helix domain-containing protein 3 (ChChd3) into the mitochondrial intermembrane space. J Biol Chem. 2012;287(47):39480-39491. doi:10.1074/jbc.M112.387696.

34. Lynes EM, Bui M, Yap MC, Benson MD, Schneider B, Ellgaard L, Berthiaume LG, Simmen T. Palmitoylated TMX and calnexin target to the mitochondria-associated membrane. EMBO J. 2012;31(2):457-470. doi:10.1038/emboj.2011.384.

35. Merrick BA, Dhungana S, Williams JG, Aloor JJ, Peddada S, Tomer KB, Fessler MB. Proteomic profiling of *S*-acylated macrophage proteins identifies a role for palmitoylation in mitochondrial targeting of phospholipid scramblase 3. Mol Cell Proteomics. 2011;10(10):M110.006007. doi:10.1074/mcp.M110.006007.

36. Saxowsky TT, Meadows KL, Klungland A, Doetsch PW. 8-Oxoguanine-mediated transcriptional mutagenesis causes Ras activation in mammalian cells. Proc Natl Acad Sci USA. 2008;105(48):18877-18882. doi:10.1073/pnas.0806464105.

37. Pommier Y, Leo E, Zhang H, Marchand C. DNA topoisomerases and their poisoning by anticancer and antibacterial drugs. Chem Biol. 2010;17(5):421-433. doi:10.1016/j.chembiol.2010.04.012.

38. Dizdaroglu M, Jaruga P. Mechanisms of free radical–induced damage to DNA. Free Radical Res. 2012;46(4):382-419. doi:10.3109/10715762.2011.653969.

39. Wyatt MD, Pittman DL. Methylating agents and DNA repair responses: methylated bases and sources of strand breaks. Chem Res Toxicol. 2006;19(12):1580-1594. doi:10.1021/tx060164e.

40. Potter VR, Reif AE. Inhibition of an electron transport component by antimycin A. J Biol Chem. 1952;194(1):287-297.

41. Doudican NA, Song B, Shadel GS, Doetsch PW. Oxidative DNA damage causes mitochondrial genomic instability in *Saccharomyces cerevisiae*. Mol Cell Biol. 2005;25(12):5196-5204. doi:10.1128/mcb.25.12.5196-5204.2005.

42. Jaruga P, Kirkali G, Dizdaroglu M. Measurement of formamidopyrimidines in DNA. Free Radical Biol Med. 2008;45(12):1601-1609. doi:10.1016/j.freeradbiomed.2008.09.019.

43. Joshi R, Ho KK, Tenney K, Chen JH, Golden BL, Gimble FS. Evolution of I-SceI homing endonucleases with increased DNA recognition site specificity. J Mol Biol. 2011;405(1):185-200. doi:10.1016/j.jmb.2010.10.029.

44. Stuckey S, Storici F. Gene knockouts, *in vivo* site-directed mutagenesis and other modifications using the *delitto perfetto* system in *Saccharomyces cerevisiae*. Methods Enzymol. 2013;Volume 533:103-131. doi:10.1016/B978-0-12-420067-8.00008-8.

45. Wijshake T, Baker DJ, van de Sluis B. Endonucleases: new tools to edit the mouse genome. Biochim Biophys Acta. 2014. doi:10.1016/j.bbadis.2014.04.020.

46. Bulina ME, Chudakov DM, Britanova OV, Yanushevich YG, Staroverov DB, Chepurnykh TV, Merzlyak EM, Shkrob MA, Lukyanov S, Lukyanov KA. A genetically encoded photosensitizer. Nat Biotechnol. 2006;24(1):95-99. doi:10.1038/nbt1175.

47. Roy A, Carpentier P, Bourgeois D, Field M. Diffusion pathways of oxygen species in the phototoxic fluorescent protein KillerRed. Photochem Photobiol Sci. 2010;9(10):1342-1350. doi:10.1039/c0pp00141d.

48. Lada AG, Dhar A, Boissy RJ, Hirano M, Rubel AA, Rogozin IB, Pavlov YI. AID/APOBEC cytosine deaminase induces genome-wide kataegis. Biol Direct. 2012;7:47. doi:10.1186/1745-6150-7-47.

49. Winczura A, Zdzalik D, Tudek B. Damage of DNA and proteins by major lipid peroxidation products in genome stability. Free Radical Res. 2012;46(4):442-459. doi:10.3109/10715762.2012.658516.

50. Jarvela TS, Linstedt AD. The application of KillerRed for acute protein inactivation in living cells. Curr Prot Cytometry. 2014;69:12.35.11-12.35.10. doi:10.1002/0471142956.cy1235s69.

51. Waldeck W, Mueller G, Glatting KH, Hotz-Wagenblatt A, Diessl N, Chotewutmonti S, Langowski J, Semmler W, Wiessler M, Braun K. Spatial localization of genes determined by intranuclear DNA fragmentation with the fusion proteins lamin KRED and histone KRED und visible light. Int J Med Sci. 2013;10(9):1136-1148. doi:10.7150/ijms.6121.

52. Nordgren M, Wang B, Apanasets O, Brees C, Veldhoven PP, Fransen M. Potential limitations in the use of KillerRed for fluorescence microscopy. J Microsc. 2012;245(3):229-235. doi:10.1111/j.1365-2818.2011.03564.x.

53. Schneider CA, Rasband WS, Eliceiri KW. NIH Image to ImageJ: 25 years of image analysis. Nat Methods. 2012;9(7):671-675. doi:10.1038/nmeth.2089.

54. Kamentsky L, Jones TR, Fraser A, Bray M-A, Logan DJ, Madden KL, Ljosa V, Rueden C, Eliceiri KW, Carpenter AE. Improved structure, function and compatibility for CellProfiler: modular high-throughput image analysis software. Bioinformatics. 2011;27(8):1179-1180. doi:10.1093/bioinformatics/btr095.

55. Bush A, Colman-Lerner A. Quantitative measurement of protein relocalization in live cells. Biophys J. 2013;104(3):727-736. doi:10.1016/j.bpj.2012.12.030.

56. Martin J, Mahlke K, Pfanner N. Role of an energized inner membrane in mitochondrial protein import. Δψ drives the movement of presequences. J Biol Chem. 1991;266(27):18051-18057.

57. Kulawiak B, Hopker J, Gebert M, Guiard B, Wiedemann N, Gebert N. The mitochondrial protein import machinery has multiple connections to the respiratory chain. Biochim Biophys Acta. 2013;1827(5):612-626. doi:10.1016/j.bbabio.2012.12.004.

58. Schleyer M, Schmidt B, Neupert W. Requirement of a Membrane Potential for the Posttranslational Transfer of Proteins into Mitochondsria. Eur J Biochem. 1982;125(1):109-116. doi:10.1111/j.1432-1033.1982.tb06657.x.

59. Veatch JR, McMurray MA, Nelson ZW, Gottschling DE. Mitochondrial dysfunction leads to nuclear genome instability via an iron-sulfur cluster defect. Cell. 2009;137(7):1247-1258.

60. Giloh H, Mager J. Inhibition of peptide chain initiation in lysates from ATP-depleted cells.I. Stages in the evolution of the lesion and its reversal by thiol compounds, cyclic AMP or purine derivatives and phosphorylated sugars. Biochim Biophys Acta. 1975;414(3):293-308.

61. Friedman JR, Nunnari J. Mitochondrial form and function. Nature. 2014;505(7483):335-343. doi:10.1038/nature12985.

62. Lill R, Hoffmann B, Molik S, Pierik AJ, Rietzschel N, Stehling O, Uzarska MA, Webert H, Wilbrecht C, Muhlenhoff U. The role of mitochondria in cellular iron-sulfur protein biogenesis and iron metabolism. Biochim Biophys Acta. 2012;1823(9):1491-1508. doi:10.1016/j.bbamcr.2012.05.009.

63. Francis BR, White KH, Thorsness PE. Mutations in the Atp1p and Atp3p subunits of yeast ATP synthase differentially affect respiration and fermentation in *Saccharomyces cerevisiae*. J Bioenerg Biomembr. 2007;39(2):127-144. doi:10.1007/s10863-007-9071-4.

64. Ferguson LR, von Borstel RC. Induction of the cytoplasmic 'petite' mutation by chemical and physical agents in *Saccharomyces cerevisiae*. Mutat Res. 1992;265(1):103-148. doi:10.1016/0027-5107(92)90042-Z.

65. Schroeder EA, Shadel GS. Crosstalk between mitochondrial stress signals regulates yeast chronological lifespan. Mech Age Dev. 2014;135(0):41-49. doi:10.1016/j.mad.2013.12.002.

66. Gnad F, Gunawardena J, Mann M. PHOSIDA 2011: the posttranslational modification database. Nucleic Acids Res. 2011;39(Database issue):D253-260. doi:10.1093/nar/gkq1159.

67. Chen Y, Sanchez Y. Chk1 in the DNA damage response: conserved roles from yeasts to mammals. DNA Repair. 2004;3(8–9):1025-1032. doi:10.1016/j.dnarep.2004.03.003.

68. Bannai H, Tamada Y, Maruyama O, Nakai K, Miyano S. Extensive feature detection of N-terminal protein sorting signals. Bioinformatics. 2002;18(2):298-305. doi:10.1093/bioinformatics/18.2.298.

Appendix
List of Supplemental Materials

Accompanying this dissertation are several supplemental computer files. These files are listed here, organized by chapter. All files are available in the Emory University Electronic Theses and Dissertations repository.

Chapter 2

BESIR Protein Predicted Sites - Yeast.xlsx

Contains the *Saccharomyces cerevisiae* base excision and strand incision repair (BESIR) protein sequence analysis data and calculations used to generate Figure 2-1.

BESIR Protein Predicted Sites - Human.xlsx

Contains the human BESIR protein sequence analysis data and calculations used to generate Figure 2-1.

Chapter 3

Q-SCAn Package.zip

Contains the files and additional documentation necessary to implement Quantitative Subcellular Compartmentalization Analysis (Q-SCAn). See the included README.txt file for further details.