Role of human polynucleotide kinase/phosphatase (PNKP) in nuclear and mitochondrial DNA repair

by

Sudip Subedi

A thesis submitted in partial fulfillment of the requirements for the degree of

Master of Science

in

Experimental Oncology

Department of Oncology
University of Alberta

Abstract

DNA is continuously under stress from the insults inflicted by both endogenous and exogenous agents. These agents include endogenous reactive oxygen species (ROS), ionizing radiation (IR) and many common chemotherapeutic drugs, which can inflict a plethora of DNA lesions including abasic sites, DNA single-strand breaks (SSBs) and double-strand breaks (DSBs). Such damages can be mutagenic or lethal and are often linked to cancer and degenerative diseases. Cells are equipped with sophisticated repair machineries to prevent the mutational and cytotoxic consequences of DNA damage, which involve the orchestration of various repair proteins. Frequently, IR, ROS and topoisomerase 1 inhibitor-induced strand breaks bear termini with enzymatically incompatible ends such as 5'-hydroxyl and 3'-phosphate, which need to be processed to generate ligatable ends, i.e. 5'-phosphate and 3'-hydroxyl termini. The dual functioning enzyme polynucleotide kinase/phosphatase (PNKP) contains both a kinase domain to phosphorylate 5'-OH termini and a phosphatase domain to replace 3'-phosphate with 3'-OH termini. PNKP's role in nuclear repair pathways like base excision/single strand-break repair (BER/SSBR) and non-homologous end joining (NHEJ) for DSB repair has been well established.

Mitochondria of human cells contain a 16.5 kbp circular DNA molecule (mtDNA), which is even more under threat of endogenous ROS owing to its proximity to the electron transport chain and lack of chromatin-associated protection unlike nuclear DNA. Natural mutation frequencies in some regions of mtDNA are 20-100 fold higher than the nuclear DNA. The oxidative damages and breaks of the mitochondrial genome have been implicated in various human degenerative diseases, aging and cancer. Several DNA

repair pathways have been shown to be involved in the maintenance of mtDNA. Here we show that functionally active full-length PNKP localizes to human mitochondria.

Based on these findings, we hypothesized that mitochondrial PNKP (mtPNKP) must be critical for cell survival. In order to begin to examine the importance of mtPNKP in cell survival, we set out to generate cell lines that express PNKP exclusively either in the nucleus or mitochondria. To achieve that, we mapped the nuclear localization signal (NLS) and mitochondrial targeting sequence (MTS) of PNKP and used site-directed mutagenesis to mutate them in an attempt to impair PNKP trafficking into the organelle of choice, e.g. mutating the NLS to impair PNKP import into the nucleus. As an alternative approach, we used the pShooter organelle directing vector system to guide PNKP to the organelle of choice. Once the cell lines are generated, we will challenge them with various genotoxic and chemotherapeutic agents. As part of the targeting approach we generated human A549 PNKP knockout cells using CRISPR technology.

Preface

Some of the research conducted for this thesis has been published as Nasser Tahbaz, Sudip Subedi, and Michael Weinfeld. (2012) "Role of polynucleotide kinase/phosphatase in mitochondrial DNA repair". Nucleic Acid Res. 40:(8); 3484-95. I was responsible for the experimental design, performance and data collection for a part of the published research.

Acknowledgement

I would like to express my gratitude to my supervisor, Dr. Michael Weinfeld for his constant support and supervision. Without your motivation, achieving a master's degree would not have been possible.

Secondly, I would like to acknowledge my family for patiently waiting for my studies to finish and bearing with me when I was not present in some important days of their lives. I will certainly make up for them in years to follow.

Most importantly, I would like to thank Dr. Nasser Tahbaz for supervising me in my earlier days and training me in using lab equipment. Thank you for allowing me to carry on your 'mitochondrial DNA' legacy.

In addition, I would like to thank my best friends in the lab Mesfin Fanta, Bingcheng Jiang, Ismail Abdou, Raymond Roland and Paviga Limudomporn. You guys made my research experience joyous and lively.

Furthermore, I would like to thank Dr. Feridoun Karimi-Busheri, Dr. Aghdass Rasouli-Nia, Dr. Rajam Mani and Dr. Mohamed El-Gendy for their advice and support. I would also express my gratitude to imaging facility, especially Geraldine Barron and Dr. Xuejun Sun.Without their help this project would not have been possible.

Lastly, I would like to thank Dr. Mary Hitt, Dr. Gordon Chan and Dr. Andrew Shaw for allowing me to work on a rotation in their labs during my qualifying semester.

TABLE OF CONTENTS

LIST OF TABLES

PART: I

PART: II

APPENDIX-A

LIST OF FIGURES

PART: I

PART: II

LIST OF ABBREVIATIONS

Abbreviation	Full name
%	per cent
ºC	degree celsius
^{32}P	phosphate 32 radio-isotope
3'OH	3' hydroxyl group
3'P	3' phosphate group
5'OH	5' hydroxyl group
5'P	5' phosphate group
A	adenine
ADP	adenosine diphosphate
APE1	apurinic/apyrimidinic endonuclease 1
APLF	PNKP-APTX-like factor
AP site	apurinic/apyrimidinic site
APTX	aprataxin
ARM	armadillo
ATM	ataxia telangiectasia-mutated
ATP	adenosine triphosphate
BARD	BRCA1-associated RING domain
BER	base excision repair
BLAST	Basic Local Alignment Search Tool
bp	base pair
C	cytosine
COX	Cytochrome C oxidase
CO_2	carbon dioxide
CRISPR	Clusters of regularly inter-spaced short palindromic repeat
crRNA	CRISPR RNA)
CRY	cryptochrome circadian clock
Cyt	cytochrome
DAPI	4',6-diamidino-2-phenylindole
DMEM	Dulbecco's Modified Eagle Medium
DMEM/F12	Dulbecco's Modified Eagle Medium: nutrient mixture F-12
DNA	deoxyribonucleic acid
DNA-PK	deoxyribonucleic acid protein kinase
DNA-PKcs	deoxyribonucleic acid protein kinase catalytic subunit
DSB	double strand break
DSBR	double strand break repair
dsDNA	double-stranded DNA
EDTA	ethylenediaminetetraacetic acid
FACS	fluorescence activated cell sorting
FBS	fetal bovine serum
FEN	flap endonuclease
FG	phenylalanine-glycine

FHA	forkhead associated
FRAP	fluorescence recovery after photobleaching
G	guanine
g	gravity (units)
GAP	GTPase activating protein
GDP	guanosine diphosphate
GFP	green fluorescent protein
GTP	guanosine-5'-triphosphate
h	hour
H (1, 2B)	histone (1, 2B)
H101	lab-made mouse monoclonal antibody against PNKP
H_2O_2	hydrogen peroxide
HCl	hydrochloric acid
hPNKP	human polynucleotide kinase/phosphatase
HJ	holliday junctions
HR	homologous recombination (homology directed repair)
HRP	horseradish peroxidase
Hsc	heat shock cognate
Hsp	heat shock protein
IBB	Importin-β binding
IDT	Integrated DNA Technologies
IM	inner membrane
IMS	intermembrane space
$InsP_3R$	Inositol 1, 4, 5-triphosphate Receptor
IR	ionizing radiation
kb	kilo base pair
KD	knockdown
kDa	kilo-dalton
L	liter
LB	lysogeny broth
lpSSBR	long-path SSBR
MCSZ	Microcephaly, early-onset, intractable seizures and developmental delay disorder
mer	oligomer
mg	milligram
$MgCl_2$	magnesium chloride
MGMT	O^6-alkylguanine DNA alkyltransferase
Mia	mitochondrial IMS import and assembly
min	minutes
ml	milliliter
MLH	MutL homolog 1
mm	millimeter
mM	millimolar
MMP	mitochondrial processing peptidase
MMR	mismatch repair
MRN	Mre11-Rad50-Nbs1

mRNA	messenger RNA
mtBER	mitochondrial BER
mtDNA	mitochondrial DNA
MTH	mutT homologue
mtHsp	mitochondrial Hsp
mtMMR	mitochondrial MMR
mtPNKP	mitochondrial PNKP
MTS	mitochondrial target signal
NaCl	sodium chloride
NADH	nicotinamide adenine dinucleotide (reduced)
NaF	sodium fluoride
Na_3VO_4	sodium orthovanadate
NCBI	National Center for Biotechnology Information
NE	nuclear envelope
NER	nucleotide excision repair
NEIL	nei endonuclease VIII-like
ng	nanograms
NH_2	amino
NHEJ	non-homologous end joining
nM	nanomolar
nm	nanometer
NPC	nuclear pore complex
NTP	nucleoside triphosphate
NLS	nuclear localization signal
OGG1	8-oxoguanine DNA-glycosylase 1
OM	outer membrane
Oxa	oxidase assembly
PAM	pre-sequence translocase associated motor (mitochondria)
PAM	protospacer adjacent motif (CRISPR)
PAR	poly(ADP-ribose)
PARG	poly(ADP-ribose) glycohydrolase
PARP	poly(ADP-ribose) polymerase
PBS	phosphate-buffered saline
PBST	phosphate-buffered saline-tween 20
PCNA	proliferating cell nuclear antigen
PCR	polymerase chain reaction
PFA	paraformaldehyde
pH	potential hydrogen
PKC δ	protein kinase δ
Plk1	polo-like kinase 1
PMSF	phenylmethylsulfonyl fluoride
PNKP	human polynucleotide kinase/phosphatase
Pol β	DNA polymerase beta
PTEN	phosphatase and tensin homolog deleted on chromosome 10
RIPA	radioimmunoprecipitation buffer

RNA	ribonucleic acid
ROS	reactive oxygen species
RPA	replication protein A
rpm	revolutions per minute
rRNA	ribosomal RNA
RT-PCR	Reverse transcription - PCR
s	seconds
SDS	sodium dodecyl sulfate
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
sgRNA	single-guide RNA
SHP1	Src homology region 2 domain-containing phosphatase-1
SOD1	superoxide dismutase 1
SSB	single-strand break
spSSBR	short-patch SSBR
SSBR	single-strand break repair
ssDNA	single-strand DNA
T	thymine
TBE	Tris/Borate/EDTA
TDP	tyrosyl DNA-phosphodiesterase
TOB/SAM	Topogenesis of mitochondrial outer membrane β-barrel proteins/sorting and assemble-machinery
TOM	translocase of the outer membrane
TIM	translocase of the inner membrane
tracrRNA	transacting RNA
Tris-HCl	tris(hydroxymethyl)aminomethane-hydrochloric acid
tRNA	transfer RNA
Top1	topoisomerase I
T4 PNK	T4 Bacteriophage polynucleotide kinase/phosphatase polynucleotide kinase/phosphatase
UV	ultraviolet
V	volt
VDAC	voltage-dependent anion channels
WT	wildtype
µg	microgram
µg/ml	microgram/milliliter
µl	microliter
µM	micromolar
UNG	uracil-DNA glycosylase
UV	ultraviolet
XLF	XRCC4-like factor
XL-qPCR	Extra-large qPCR (XL-qPCR)
XRCC1	X-ray repair complementing defective repair in Chinese hamster cells 1
XRCC4	X-ray repair complementing defective repair in Chinese hamster cells 4

YB1	Y-box binding protein 1
α	alpha
β	beta
γ	gamma
δ	delta
λ	wavelength (cm)

Copyright licenses

Licensee: Sudip Subedi. License Date: July 12, 2015. License Number: 3666571111459. Publication: Cell. Title: DNA Single Strand Break Repair and Spinocerebellar Ataxia. Type of use: Reuse in a thesis/dissertation. Total: 0.00 USD.

Licensee: Sudip Subedi. License Date: July 15, 2015. License Number: 3670340219291. Publication: Oncogene. Title: Regulation and mechanisms of mammalian double-strand break repair. Type of use: Reuse in a thesis/dissertation. Total: 0.00 USD.

Licensee: Sudip Subedi. License Date: July 15, 2015. License Number: 3670321270408. Publication: DNA Repair. Title: DNA double-strand break repair: From mechanistic understanding to cancer treatment. Type of use: Reuse in a thesis/dissertation. Total: 0.00 USD.

Licensee: Sudip Subedi. License Date: July 15, 2015. License Number: 3670340489163. Publication: Trends in Biochemical Sciences. Title: Mitochondrial protein import: from transport pathways to an integrated network. Type of use: Reuse in a thesis/dissertation. Total: 0.00 USD.

Licensee: Sudip Subedi. License Date: July 15, 2015. License Number: 3670330742062. Publication: The EMBO Journal. Title: Transport into and out of the cell nucleus. Type of use: Reuse in a thesis/dissertation. Total: 0.00 USD.

Licensee: Sudip Subedi. License Date: July 12, 2015. License Number: 3666790685740. Publication: Molecular Cell. Title: The Molecular Architecture of the Mammalian DNA Repair Enzyme, Polynucleotide Kinase. Type of use: Reuse in a thesis/dissertation. Total: 0.00 USD.

Licensee: Sudip Subedi. License Date: July 12, 2015. License Number: 3666791364085. Publication: Trends in Biochemical Sciences. Title: Tidying up loose ends: the role of polynucleotide kinase/phosphatase in DNA strand break repair. Type of use: Reuse in a thesis/dissertation. Total: 0.00 USD.

Licensee: Sudip Subedi. License Date: July 12, 2015. License Number: 3666570712029. Publication: Nucleic Acids Research. Title: Role of polynucleotide kinase/phosphatase in mitochondrial DNA repair. Type of use: Reuse in a thesis/dissertation. Total: 0.00 USD.

Licensee: Sudip Subedi. License Date: July 12, 2015. License Number: 3666800552561. Publication: Nature Methods. Title: Cas9 as a versatile tool for engineering biology. Type of use: Reuse in a thesis/dissertation. Total: 0.00 USD.

Licensee: Sudip Subedi. License Date: July 12, 2015. License Number: 3666801415098. Publication: Science. Title: Genome-scale CRISPR-Cas9 knockout screening in human cells. Type of use: Reuse in a thesis/dissertation. Total: 0.00 USD

Licensee: Sudip Subedi. License Date: July 12, 2015. License Number: 3666810100142. Publication: Nature Protocols. Title: CRISPR interference (CRISPRi) for sequence-specific control of gene expression. Type of use: Reuse in a thesis/dissertation. Total: 0.00 USD.

Part I

Chapter I

Introduction

1.1 DNA damage and repair

DNA is continuously under stress from the insults inflicted by both endogenously produced and exogenous agents. These agents may include agents that are produced during natural cellular metabolism like endogenous reactive oxygen species (ROS), or external agents like ionizing radiation (IR) or ultraviolet radiation (UV) and many common chemotherapeutic drugs like topoisomerase poisons, and DNA cross-linking chemicals[12]. These agents can inflict a plethora of DNA lesions including abasic sites, DNA single-strand breaks (SSBs) and double-strand breaks (DSBs). DSBs are considered to be the most deleterious of all because an unrepaired DSB can cause mutagenic translocations or deletions or destine a cell for cell cycle arrest or apoptosis[16, 17]. Although less deleterious, SSBs can become serious lesions for the cell as inefficiency in repairing them can result in formation of DSBs due to replication fork collapse, which occurs when a replication fork runs into the site of SSB lesions[16].

Spontaneously occurring events such as loss of purines and cytosine deamination and ROS-mediated base damage are commonplace in cells. If unrepaired, they may lead to mutations or SSBs and DSBs. It is believed that under normal cellular processes a cell sustains approximately 10,000 SSBs and 10 DSBs per day. These naturally transpiring events if coupled with effects of exogenous environmental factors can dramatically raise the incidence of aforementioned genotoxic lesions. Such damages can lead to genomic instability and are often linked with neurodegenerative diseases and various types of cancers[12, 18].

Fortunately, cells are equipped with sophisticated DNA repair machineries to

prevent the mutational and cytotoxic consequences of DNA damage, which involve orchestrations of various DNA repair proteins. Various DNA repair pathways operate in a coordinated fashion to rectify those damages in order to maintain genomic integrity. These pathways include nucleotide excision repair (NER), which handles bulky adducts such as pyrimidine dimers generated by UV radiation[19]; base excision repair (BER), which deals with smaller base lesions and sugar damage and SSBs[20] and; DSB repair pathways like non-homologous end joining (NHEJ) and homology directed repair (HR).

1.1.1 Base Excision Repair (BER)

Base lesions arise due to base hydrolysis, oxidation, alkylation, deamination or other form of alterations usually inflicted on DNA molecules by several endogenous and/or exogenous genotoxic agents[18, 21]. If left unrepaired, these lesions get converted into mutations potentially affecting genetic stability and cellular metabolism. The BER pathway can eliminate the aforementioned base damages. BER initiates by recognition of the DNA lesion by a DNA glycosylase, which cleaves the N-glycosylic bond that connects the DNA base with the sugar phosphate backbone[22]. There are 11 different damage-specific mammalian DNA glycosylases discovered so far[23]. However, they are grouped into two broad categories: monofunctional DNA glycosylases cleave the N-glycosylic bond by flipping the damaged base out of the double helix generating an apurinic/apyrimidinic (AP site) in the DNA. The AP site is further processed by an AP endonuclease (APE1 in mammals) resulting in a nick or SSB[24]. Bifunctional DNA glycosylases on the other hand, possess both N-glycosylic bond cleaving and AP lyase

activity, thus it does not require an extra AP endonuclease action to produce an SSB[24]. The action of AP endonucleases and bifunctional DNA glycosylases can produce obstructive DNA termini such as 5'-deoxyribose phosphate[25] or 3'-α,β-unsaturated aldehyde[26, 27]that are incompatible for DNA polymerases to replace missing nucleotides or DNA ligases to rejoin broken strands. The unligatable DNA termini must be cleaned up by specialized enzymes in order to generate the correct 3'-OH and 5'-P ends. After correct termini are generated, BER-mediated DNA repair is succeeded either by a short-patch (spSSBR) or long-patch (lpSSBR) DNA repair pathway.

spSSBR is a major pathway for repairing SSBs and is orchestrated by several key repair enzymes. Most BER is carried out by spSSBR, which involves elimination and replacement of only one nucleotide. SSBR includes four main steps: damage recognition, DNA end processing, DNA gap filling and DNA ligation (Fig 1)[1]. The process initiates with poly(ADP-ribose) polymerase I (PARP1), which recognizes and binds to the site of damage[28]. Upon SSB binding, PARP1 undergoes conformational change and activation[29]. Activated PARP1 then catalyzes the formation of long poly(ADP-ribose) (PAR) chains on itself and various other substrates including histones H1 and H2B and transcription factors[29-31]. The accumulation of negatively charged PAR chains on the proteins enables their dissociation from the chromosome leading to chromatin decondensation, thereby allowing downstream repair proteins access to the damage site[32, 33]. Additionally, PAR polymers flag the site for recruitment of diverse repair proteins like X-ray complementing group 1 (XRCC1), polynucleotide kinase/phosphatase (PNKP) and DNA ligase III[34, 35]. PAR moieties have a very short lifespan at SSBs and are quickly degraded by poly(ADP-ribose) glycohydrolase (PARG) to facilitate DNA

repair[36]. XRCC1 is important in its role as a scaffold protein that recruits other repair proteins involved. XRCC1 has also been shown to stimulate the activity and turnover rate of PNKP[37]. At the site of damage, PNKP via its 3'-phosphatase and 5'-OH kinase functions restores the incompatible ends on SSB to ligatable 3' hydroxyl (3'OH) and 5' phosphate (5'P) moieties[9, 38]. Apart from PNKP, other SSB end-processing enzymes known to be involved in BER are DNA polymerase β[25], APE1[39], aprataxin (APTX) [40] and tyrosyl-DNA phosphodiesterase I (TDP)[41]. Once the ends are restored, DNA polβ replaces the missing nucleotides and XRCC1 – DNA ligase III complex seals the resulting gap to accomplish spSSBR[42].

lpSSBR is an alternative SSBR pathway, which occurs only if the 5'-ends are blocked and cannot be restored by the spSSBR end-processing enzymes[43]. In lpSSBR pathway, DNA polβ incorporates a single nucleotide onto the restored 3'-OH end of the nick. Replicative DNA pol δ, in the presence of PCNA, continues elongating the DNA in a strand displacement fashion. As a result, a 2-12nt long-flap is formed, which is cleaved off by flap endonuclease 1 (FEN1) leaving behind a compatible 5'-P end on the break. The final nick sealing by DNA ligase I completes the repair[31, 44] (Fig.1).

SSBR is considered an error-free pathway[42]. The importance of SSBR pathway is underscored by the fact that DNA can sustain tens of thousands SSBs per cell per day by the action of reactive oxygen species (ROS), intracellular metabolites, and spontaneous DNA decay[42]. As mentioned earlier, unrepaired SSBs can lead to replication fork collapse thereby forming noxious lesions such as DSBs[45].

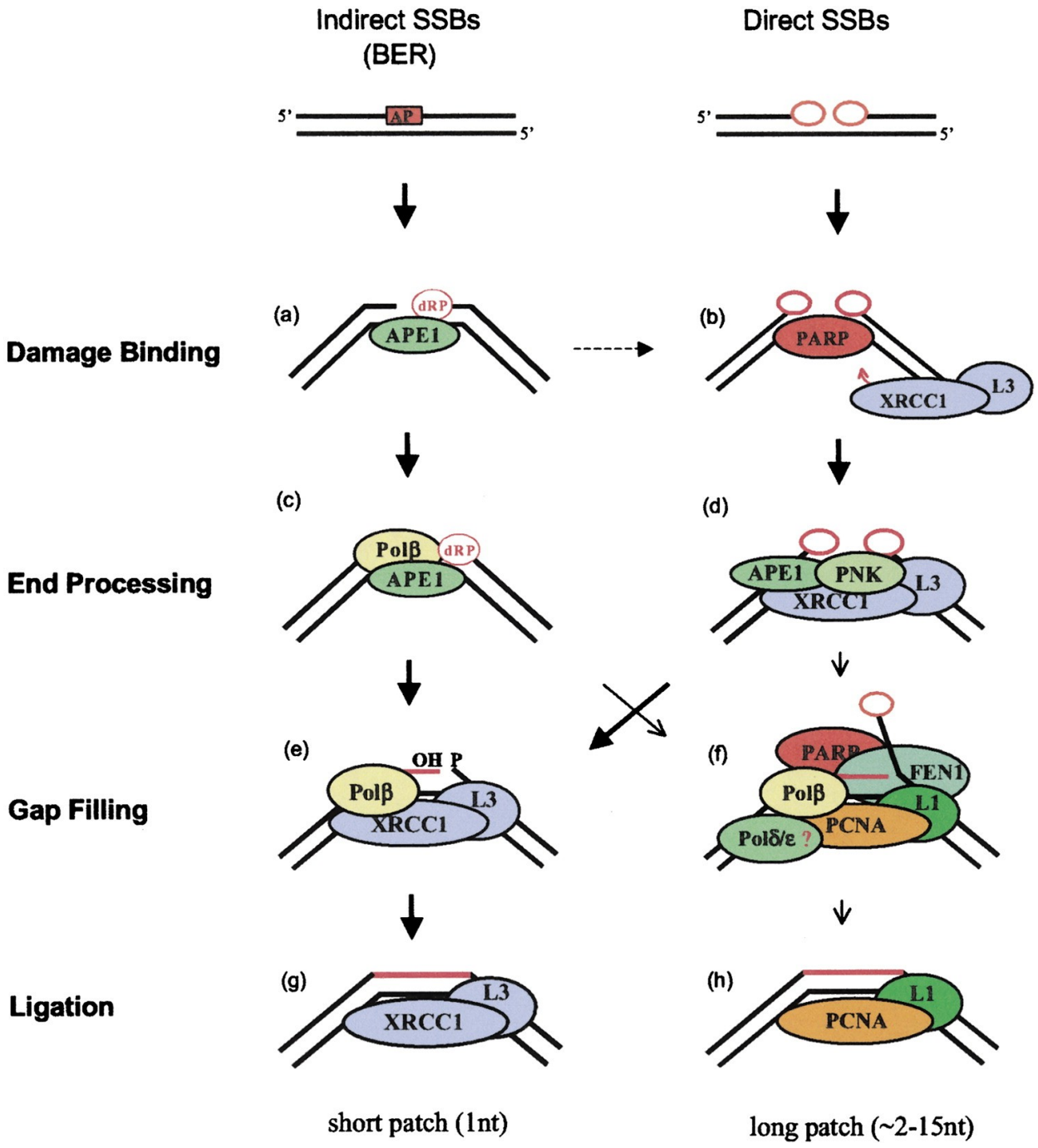


Figure 1. A schematic model of short patch SSBR and long patch SSBR. Steps are outlined in Section 1.1.1. L1: DNA ligase I, L3: DNA ligase III[1]. Reproduced by permission from Elsevier: Cell. Caldecott K.W., Copyright © 2003.

1.1.2 Non-Homologous End Joining (NHEJ)

While the HR-pathway is active only during the late S and G2 phases of cell cycle, NHEJ is active throughout the cell cycle. Unlike HR, NHEJ pathway does not require a sister chromatid as a template to repair DSBs[46]. NHEJ is therefore recognized as the major DSB repair pathway. Meanwhile, NHEJ is also considered an error-prone pathway because instead of restoring the damage back to its original form, it simply cleans up the damaged ends and ligates them[47]. Therefore there is an elevated incidence of loss (or sometimes gain) of nucleotides post-repair.

NHEJ begins with recognition of a DSB lesion by the Ku70/Ku80 heterodimer, which binds to the free ends of DSBs[48, 49]. The heterodimer recruits the DNA-dependent protein kinase catalytic subunit (DNA-PKcs) to the DNA ends, which functions to keep the two free ends of the DSB tethered to each other[47]. The Ku-heterodimer is also responsible for stimulating the serine/threonine kinase activity of DNA-PKcs[49]. The phosphorylation cascade begins with auto-phosphorylation of DNA-PKcs followed by phosphorylation of several participating proteins including Ku70, Ku80, XRCC4, XLF, Artemis and DNA Ligase IV[46]. Auto-phosphorylation of DNA-PKcs is believed to induce a conformational change that allows end-processing enzymes to gain access on the ends of DNA[50]. Artemis, an end-processing protein, forms a complex with DNA-PKcs and becomes activated. The complex formed by DNA-PKcs-Artemis has several of nuclease activities, including 5' endonuclease activity, 3' endonuclease activity and hairpin opening activity. Besides that, Artemis also possesses a 5' exonuclease activity of its own[47]. The action of the complex can clean up 5’ or 3’ overhangs forming blunt or near-blunt ends on DSBs. Although Artemis is the major end

processor of this pathway, there are additional sets of end-processing enzymes like PNKP, APTX and PNKP-APTX-like factor (APLF) that act on ‘dirty’ DNA ends and render them DNA ligase-compatible[46]. Once clean DNA ends are produced, the X-family of DNA polymerases is responsible for filling the gaps[51]. DNA ligation is accomplished by a complex composed of the proteins X-ray repair cross complementing protein 4 (XRCC4), XRCC4-like factor (XLF) and DNA Lig IV to complete the NHEJ repair process. The NHEJ steps are summarized in Figure 2[47].

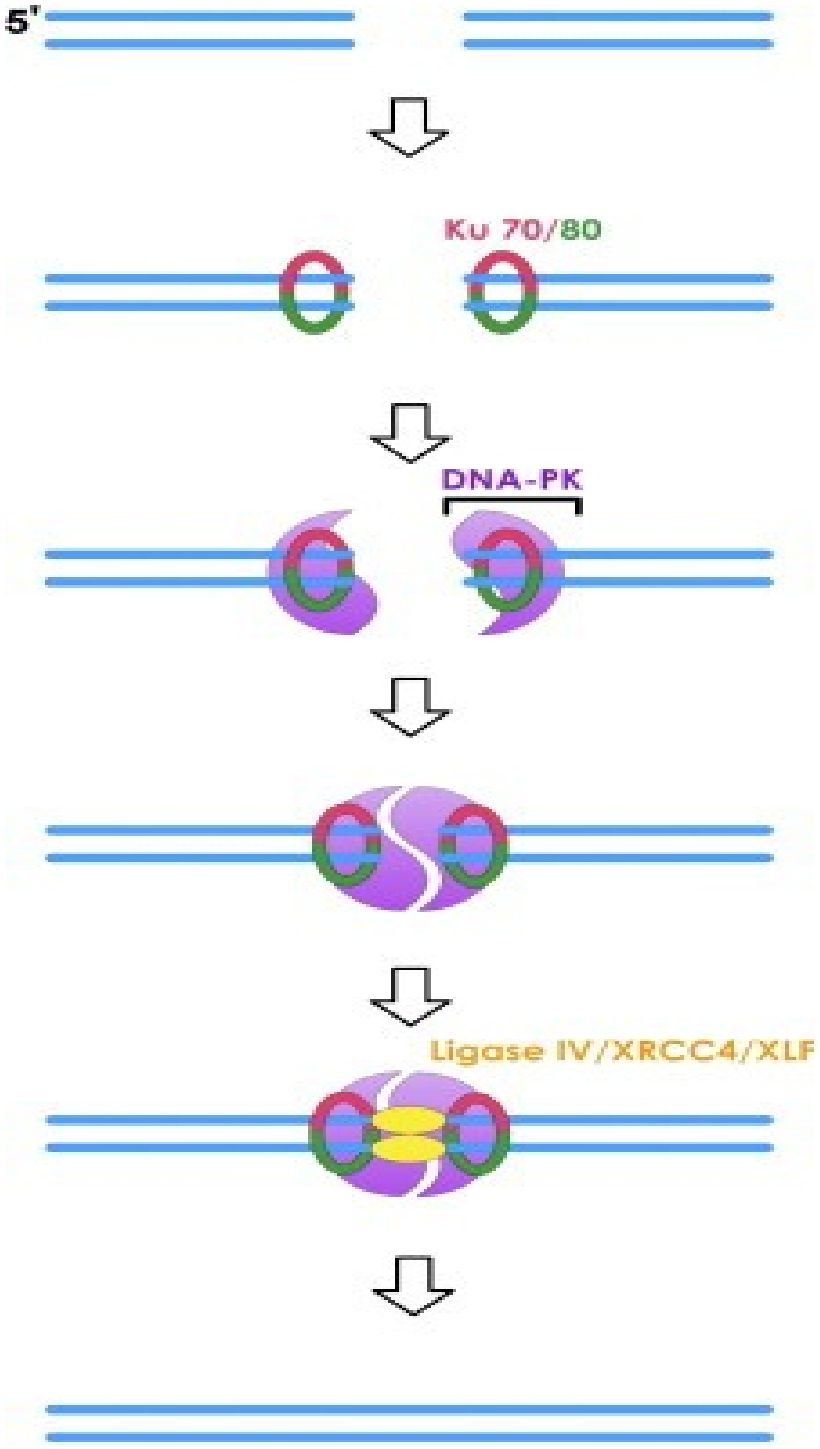


Figure 2. A representative model of non-homologous end joining (NHEJ) DNA repair pathway. Steps are outlined in Section 1.1.2. L4: DNA ligase IV[7]. Reproduced by permission from Elsevier Inc.: DNA Repair. Helleday et al., Copyright © **2**007.

1.1.3 Homology-directed (HR) Repair

HR occurs during late S/G2 phase of cell cycle and utilizes the sister chromatid as a template to replicate the damaged chromosome[52]. Therefore it is considered to be an error-free pathway. As all other DNA repair pathways, HR begins with damage recognition. This is achieved by binding of the Mre11-Rad50-Nbs1 (MRN) complex to the DSB ends[53]. Studies suggest a potential role of PARP1 in the recognition process[54]. The MRN complex additionally is responsible to hold the DNA ends in close proximity.[55] Upon DSB recognition, Nbs1 activates ATM kinase, which phosphorylates a number of substrate proteins, thereby initiating a full DNA damage response[56]. ATM phosphorylation allows for recruitment of CtIP nuclease, which in turn resects DNA end in a 5' to 3' direction to form single-stranded DNA (ssDNA) tails[57]. Next, the exposed ssDNA becomes coated with DNA replication protein A (RPA), which denatures the secondary structures present on the DNA[58]. Subsequently, RAD51 recombinase filament is assembled on the ssDNA replacing RPA proteins. The recruitment of Rad51 protein is mediated by the action of BRCA1/BARD1 and BRCA2/DSS1 complexes[58]. Rad51 promotes searching for the homologous sequence on the sister chromatid, and mediates strand invasion once the homology has been established[59]. DNA polymerase η copies the genetic information from the homologous chromosome. In this process, a DNA joint molecule is formed bearing double Holliday Junctions (HJs)[47]. In the next step, HJ resolution occurs to complete the repair process. Different protein complexes have been implicated for HJ resolution forming a different nature of repair product (cross-over or non cross-over) as a result[57]. The steps of HR-pathway are summarized in Figure 3.

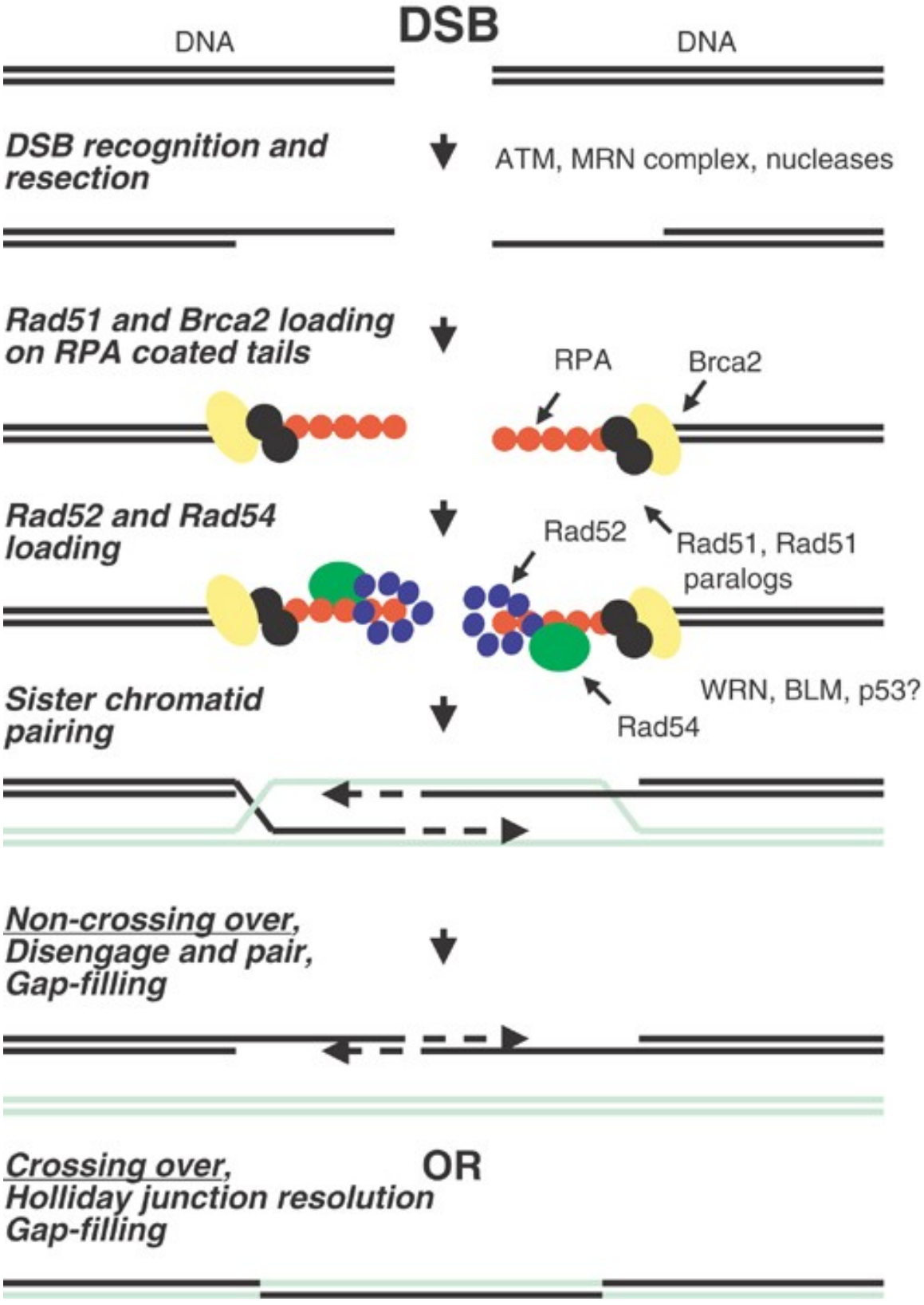


Figure 3. A representative model of homology-directed (HR) DNA repair pathway. Steps are outlined in Section 1.1.4. HJ: Holliday junction[5]. Reproduced by permission from Nature Publishing Group: Oncogene. Valerie K. and Povirk L.F., Copyright © 2003.

1.2 Mitochondrial DNA (mtDNA)

The mammalian mtDNA is a 16.6 kb closed-circular, double stranded DNA molecule (Fig 4). In mammals, mtDNA is strictly inherited from maternal lineage. Unlike its nuclear counterpart, it can replicate during all phases of cell cycle and may be present in thousands of copies per cell[60]. In human mitochondria, mtDNA is found in a structure called the nucleoid. Each nucleoid structure houses 8-10 mtDNA molecules packaged with a number of proteins. Many such structures are found throughout the mitochondrial network localizing close to the inner surface of the inner mitochondrial membrane[61, 62].

Based on the GC content of the strands, mtDNA is distinguished into heavy and light strands with the heavy strand harboring most of the encoding information. mtDNA is a massively coding DNA molecule with overlapping genes, no introns and very short intergenic sequences[63]. The genes encoded in mtDNA are all essential for cellular ATP production by oxidative phosphorylation. Genes of the mtDNA encode for 2 rRNAs, 22 tRNAs and 13 polypeptides of the electron transport chain[64]. Out of the 13 proteins, 7 are the subunits of complex I (NADH dehydrogenase), 3 are the subunits of complex IV (cytochrome c oxidase) and two are the subunits of complex V (ATP synthase) and cytochrome b (a subunit of complex III)[62, 64]. The genes are transcribed and translated within the mitochondrial matrix.

In contrast to just two copies of nuclear DNA, a somatic mammalian cell can contain up to 1000-10,000 copies of mtDNA[6]. Whereas nuclear genetic material present in each cell needs to be mandatorily repaired to ensure cell survival, loss of even a substantial fraction of mtDNA molecules can be tolerated by cells without any detrimental effect.

Therefore a pathogenic mutation has to be present in all copies (termed homoplasmy) and affect a large number of mitochondria simultaneously in order to be pathologically visible[61].

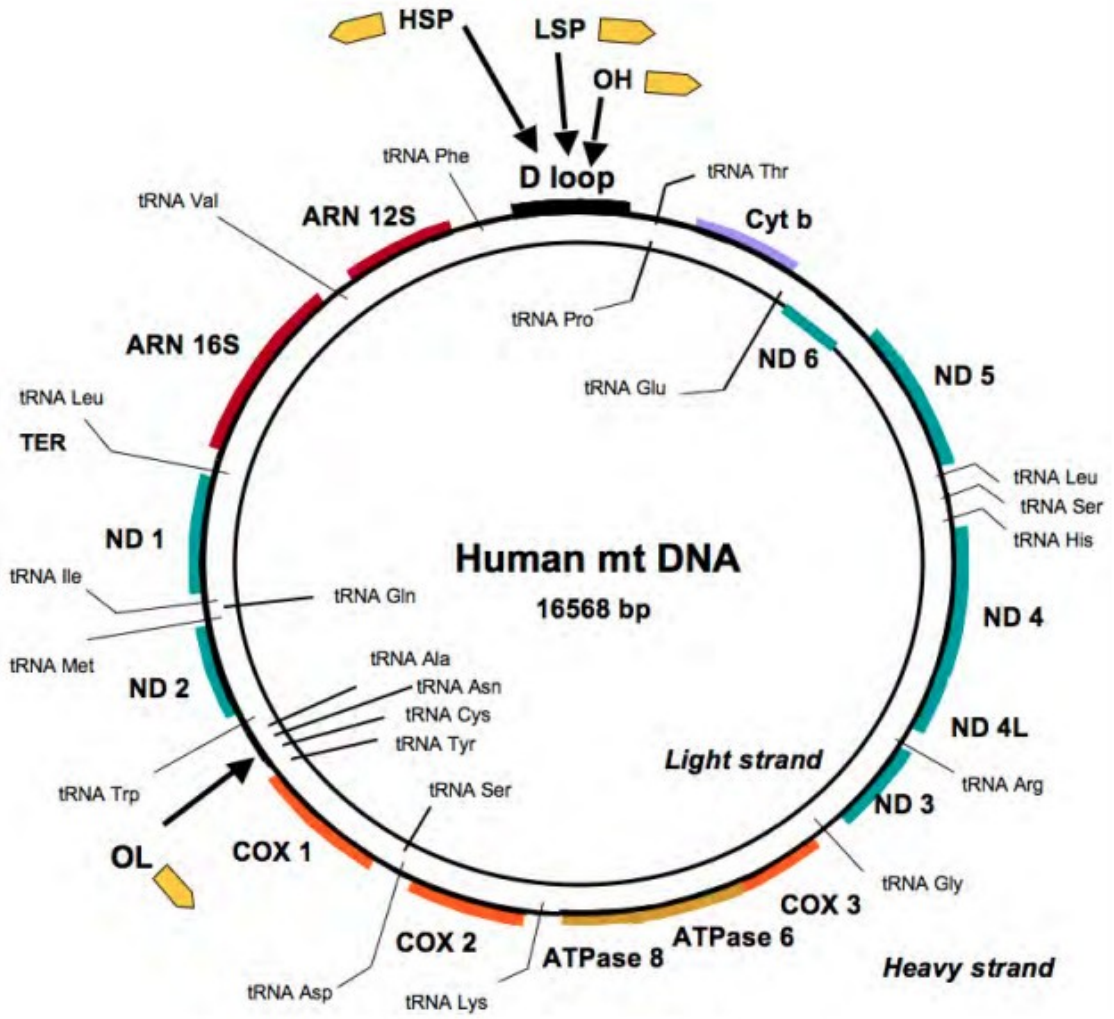


Figure 4. Representation of human mitochondrial DNA (mtDNA). The features of mtDNA are described in Section 1.2. D-loop: ~600-bp triple strand regulatory region, HSP: heavy strand, LSP: light strand, OH: origin of replication, ND1-6: NADH dehydrogenase subunits, COX1-3: Cytochrome c oxidative subunits, ATP6 and ATP8: subunits of ATP synthase, Cyt b: Cytochrome b[6]. Reproduced by permission from Frontier in Bioscience. Nadege B., Patrick L. and Rossignol R., Copyright © **2**009.

1.3 DNA damage and repair in mitochondria

Due to the close proximity to the electron transport chain and lack of chromatin-based insulation, it is believed that mtDNA is more susceptible to oxidative attack from ROS compared to nuclear DNA[65]. Studies indicate that mtDNA accumulates mutations at 10 times the rate of nuclear DNA[66].

Because mitochondria are a major source of ROS within a cell, mtDNA is relatively more at the peril of oxidative damage compared to nuclear DNA. In addition to ROS-induced base lesions, mtDNA is known to sustain various other damages such as strand breaks (SSB and DSB) in the same manner as nuclear DNA[67]. Furthermore, incorporation of mismatched bases against modified bases and errors in replication have also been reported[68]. Moreover, mtDNA has been shown to sustain damage caused by alkylation, creation of abasic sites as a result of glycosidic bond hydrolysis, spontaneous deamination of cytosines and DNA adducts (reviewed in [62]). In light of these reports, it is probable that identical or similar pathways that operate in the nucleus should also function in the mitochondria in order to preserve mtDNA integrity. Indeed, many nuclear encoded DNA repair proteins localize to the mitochondria.

1.3.1 Direct Reversal

Interestingly, CRY1 and CRY2, homologs of photolyase enzymes that are responsible for direct reversal of UV-induced pyrimidine dimers in yeast and plants but not mammals, were shown to be present in mammalian mitochondria and are involved in regulation of circadian rhythms[69, 70]. Another study suggested that MGMT, another

direct repair protein known to reverse alkylation-based DNA damage in the nucleus, might be operating in mitochondria[62, 71].

1.3.2 Mismatch Repair (MMR)

In the nucleus, the MMR pathway repairs base-base mismatches arising from spontaneous conversion of bases and errors incurred due to replicative DNA polymerase slippage particularly on repetitive genomic elements[72]. Intriguingly, although some MMR proteins (e.g., MLH1) localize in mitochondria[73], correction of mismatches sustained by mtDNA were found to be occurring in a manner independent of the conventional nuclear MMR pathway[68]. Instead, a recent finding showed that mtMMR relies on Y-box binding protein 1 (YB-1), a known transcription factor operating in the nucleus. In mitochondria, YB-1 is believed to trigger the mtMMR machinery and function in a way similar to the MutSα-complex in the nucleus[74]. Another nuclear MMR protein MTH1 operates in mitochondria as a sanitation enzyme that hydrolyses 8-oxo-NTPs into their monophosphates, which are not the normal substrates of DNA polymerase, thereby lowering the incidence of their incorporation into DNA[75]. Given the evidence so far, it appears that the nuclear and mitochondrial MMRs operate through two distinct pathways recruiting distinct sets of proteins.

1.3.3 DSB repair in mitochondria

DSBs are also common threats to mtDNA integrity. Caldecott et al. reported that mtDNA sustains 1000-fold more SSBs compared to nuclear DNA when exposed to H_2O_2[42]. DSBs in nuclear DNA are corrected by HR-mediated pathway and NHEJ

pathway. Although the HR-mediated pathway was shown to be present in plant and yeast mitochondria, it is considered to be an infrequent event in mammalian counterparts. Experiments conducted by mixing wild-type mtDNA with deletion-harboring mtDNA showed no presence of recombinant DNA molecules[76]. Some other studies registered only a low level of recombination occurring in mtDNA. A study conducted by Kajander et al on human heart discovered recombination junctions and catenations in mtDNA, the hallmarks of HR-mediated pathway, but concluded that recombination was occurring to initiate a replication process instead of repair[77]. It has been suggested that although there is a presence of recombination-mediating proteins in mitochondria, limited contact between different mtDNA molecules is a restricting factor for the process to occur[78]. However, HR proteins known to operate in the nucleus such as Rad51 and DNA2 have been shown to localize in mitochondria[79, 80].

Although collective agreement is missing, multiple studies suggest the presence of the NHEJ pathway in mitochondria. In an experiment conducted in mice, DSBs were induced in mtDNA by targeting *ScaI* restriction enzyme to mitochondria. The loss of DNA around the rejoined restriction site was in accord with the typical footprint observed with the NHEJ pathway[81]. Studies have shown KU70 and KU86 proteins to cross-react with mitochondrial extracts possessing DNA-end binding activity and depletion of DNA-PKcs from the cells was detrimental to mtDNA integrity indicating the presence of NHEJ as a repair pathway in mitochondria[78].

1.3.4 Base Excision Repair (BER)

BER is the best-understood primary pathway known for maintenance of mtDNA integrity. BER is known to correct DNA modifications caused by alkylation, deamination, oxidation or spontaneous loss of bases. Identical to nuclear BER, mtBER initiates with glycosylase enzyme-mediated recognition of modified bases and cleavage of the N-glycosidic bond to create an abasic site. The major mitochondrial glycosylases are 8-oxoguanine DNA-glycosylase1 (OGG1) and uracil-DNA glycosylase (UNG) both encoded on nuclear DNA[82]. UNG is a monofunctional glycosylase that removes uracils on DNA produced due to cytosine deamination. Once abasic sites are formed, the dual localized (nuclear and mitochondria) APE1 cleaves the phosphodiester bond 5' to the AP site. The resulting gap is then filled up by DNA polymerase γ. The final step of mtBER is sealing up of the nick by DNA ligase III[78]. OGG1 is a bifunctional glycosylase that recognizes and cleaves 8-oxoG from DNA. This enzyme cleaves DNA at the abasic site to generate a 3' α,β-unsaturated aldehyde, which is removed by APE1[82]. Additionally, several NEIL glycosylases (NEIL1, NEIL2 and NEIL3) have also been shown to be a part of mtBER repairing Fapy lesions as bifunctional glycosylases[82-84]. NEIL glycosylases usually form 3'P ends on DNA, which requires the action of PNKP to render it compatible for gap tailoring[85].

1.4 Mitochondrial Protein Import

Mitochondria are double membranous organelles comprising four sub-compartments: the outer membrane (OM), inner membrane (IM) and two aqueous

compartments, the matrix and the intermembrane space (IMS). Mammalian mitochondria contain approximately 1500 different proteins[86]. Only 1% out of the total proteins found in mitochondria are synthesized by mitochondrial ribosomes present in the matrix. That is to say, the large remainder of mitochondrial proteins are encoded by nuclear DNA and synthesized on cytosolic ribosomes in the same manner as nuclear proteins[87]. The only difference being that a majority of mitochondrial proteins are synthesized as precursor proteins (pre-proteins) that are directed into the organelle by specialized import machinery responsible for their recognition, translocation and membrane-insertion[88, 89]. The import of proteins relies on cytosolic soluble chaperones such as heat shock protein 90 (Hsp90) or heat shock cognate 70 (Hsc70) that keep the cargo soluble in the cytosol and are also responsible for their efficient assemblage and trafficking towards mitochondria[90]. Subsequently, protein sorting into respective sub-compartments is mediated by complex protein machineries, termed translocators, present on the OM and IM of the organelle[90].

1.4.1 Mitochondrial Transport Signal (MTS) and Import Machinery

The various pathways of mitochondrial import of the proteins synthesized in cytosolic ribosomes are summarized in Figure 5[89]. Proteins destined for the matrix and a majority that are destined for the IMS and IM bear a cleavable N-terminal amino acid extension, termed a pre-sequence and also considered as the classical mitochondrial target signal (MTS). In general, the N-terminal pre-sequences are 15-50 amino acid residues in length that form positively charged amphipathic α-helices (having both

hydrophilic and hydrophobic subdomains). Most pre-sequences are proteolytically cleaved, typically 5-20 residues, by mitochondrial processing peptidase (MMP) once the cargo reaches the matrix[87, 90, 91]. Amphipathic α-helices are critical for protein import because they promote the interaction of cargo with the multi-protein TOM (translocase of the outer membrane) complex, which is the entry point of mitochondria for nuclear coded proteins. The hydrophobic surface present on one half of the pre-sequence α-helix is recognized by TOM20 and positively charged surface on the other half by TOM22[87]. TOM20, TOM22 and TOM70 act as pre-sequence receptors by recognizing the MTS on pre-proteins via their cytosolic domains and translocate the cargo through a channel formed by β-barrel membrane protein TOM40[89, 92].

Once through the OM, the cargos destined for the matrix are transferred to TIM23 multi-protein complex (translocase of the inner membrane). TIM23-complex is composed of TIM17, TIM21, TIM23 and TIM50 as the components of membrane integrated translocation channel and TIM14, TIM16, TIM44, Mge1 and mtHsp70 as the components of an import motor known as pre-sequence translocase associated motor complex (PAM-complex)[88]. The pre-sequence first comes in contact with TIM50 once the cargo is partially translocated through the TOM40-channel of the OM. The IMS domains of TIM50 and TIM23 interact and together coordinate the direction of pre-proteins into the channel formed by TIM23. The membrane potential across the IM drives the insertion of pre-protein though the channel into the matrix, a process assisted by mtHsp40 and TIM44[88, 89].

Alternatively, the proteins that do not possess any cleavable pre-sequence, such as polytopic inner membrane proteins and ATP/ADP carrier proteins, possess internal

targeting signals that are recognized by TOM70 of the TOM-complex[89]. These proteins, upon ATP hydrolysis, are released by TOM70 subsequently entering into the TOM40 channel. Once in the IMS, they are bound by TIM9-TIM10 complex or TIM8-TIM13 complex that direct the cargo to the TIM22-complex translocase, which in turn delivers the cargo into the IM via a membrane potential driven process[89, 91, 93].

There are other pathways of mitochondria protein import. Proteins destined for the OM possess a β-signal on their C-terminus distinct from α-helical pre-sequences. The β-signal is recognized by the Topogenesis of mitochondrial outer membrane β-barrel proteins/sorting and assemble-machinery (TOB/SAM) complex[94]. Similarly, proteins destined for the IMS are imported by the TIM40/mitochondrial IMS import and assembly (Mia) complex[95]. Intriguingly, proteins (both mitochondrial and nuclear encoded) present in the matrix, such as COX2, can be directed to the IM by an oxidase assembly (Oxa) pathway[96].

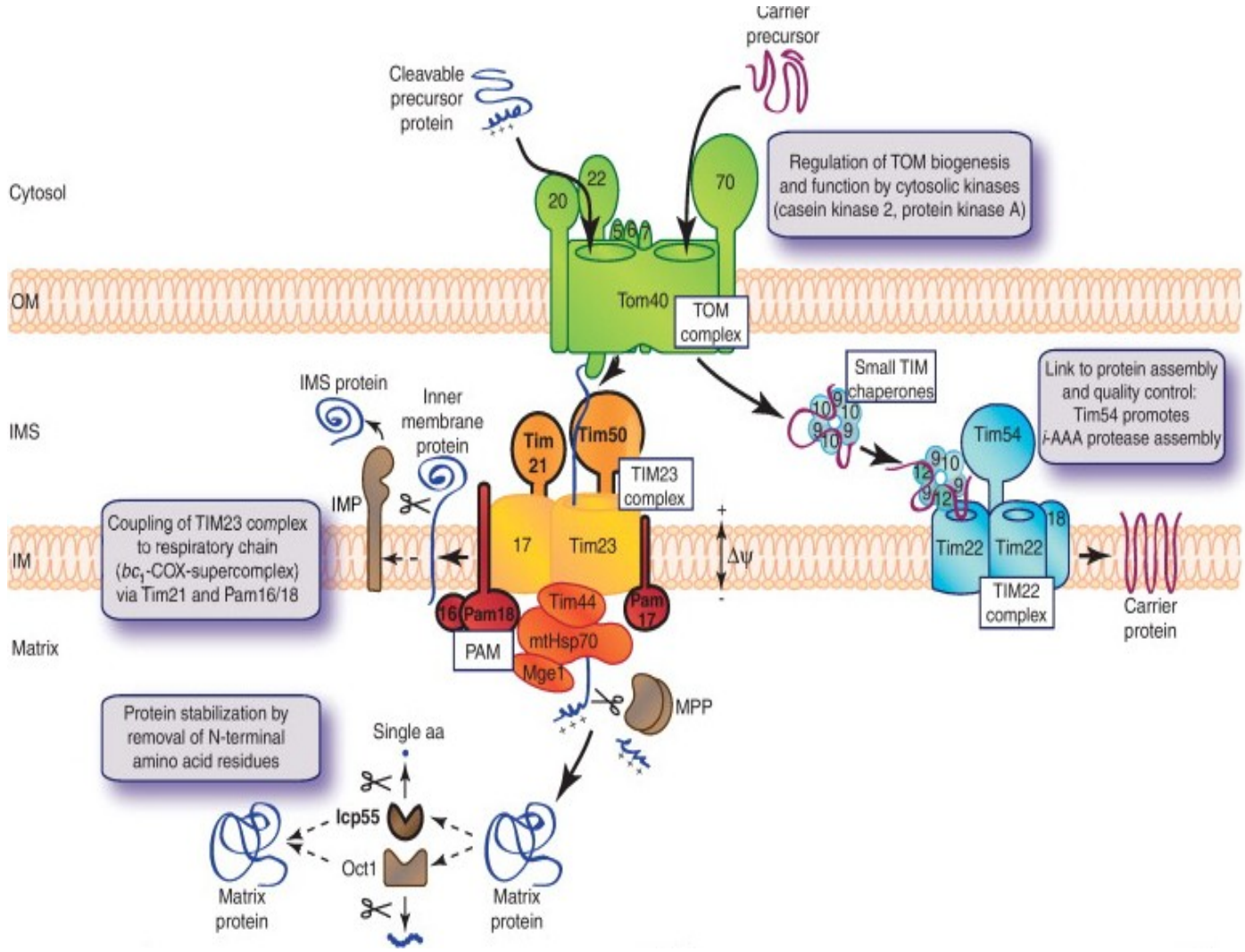


Figure 5. Classical pathways of mitochondrial protein import and proteins involved. The TOM40 complex is represented in light blue, TIM23 complex and TIM22 complex in light green and pink respectively, TOB/SAM complex in orange, MMc/PAM proteins in yellow, TIM 40/MIA proteins in brown and OXA proteins in purple. The numbers 1-5 in red represent the number of pathways involved in mitochondrial protein import. The pathways are described in Section 1.4.1[13]. Reproduced by permission from Elsevier: Trends in Biochemical Sciences. Becker T., Bottinger L. and Pfanner N., Copyright © **2**011.

1.5 Nuclear protein import

1.5.1 Import machinery: Nuclear Pore Complex (NPC), Nuclear Localization Signal (NLS) and Importins

In eukaryotic cells, the nucleus is physically separated from the cytoplasm by a double membranous nuclear envelope. Nuclear pore complexes (NPCs) are multi-protein structures present on the envelope, which regulate the bidirectional traffic of molecules across the envelope between the nucleus and cytoplasm[97]. NPCs are a cylindrical tripartite structure 90 nm long and 50 nm wide anchored between the inner and outer layers of the nuclear envelope[98]. In general, molecules including proteins of molecular mass up to 40 kDa in size can transfer through NPCs by passive diffusion whereas larger cargoes require an active energy consuming trafficking machinery to guide them into the nucleus[99, 100]. Although believed to be as many as 100 at one time, proteomic analysis experiments have identified 30 different proteins termed nucleoporins that constitute NPCs[101]. A subset of nucleoporins contains characteristic phenylalanine-glycine (FG) repeats (FG-domain) lining the central transport channel of the NPC[101].

Specialized soluble receptor proteins known as karyopherins mediate the energy dependent translocation pathways through NPCs. Karyopherins are responsible for both nuclear export and import of proteins through NPCs. In the case of nuclear import, these receptor proteins are collectively termed Importins[102].

The protein import machinery is composed of importins that can selectively bind to specific signal sequences present on the substrate proteins. Proteins that are destined for the nucleus contain a special amino acid sequence called the Nuclear Localization

Signal (NLS). The best-described and well-characterized transport signal is the classical NLS made up of basic amino acid (mostly lysine and arginine) stretches[103]. The residues can be in a single stretch, monopartite NLS, like PKKKRRV of SV40 large T-antigen, or in two stretches, bipartite NLS like KRPAATKKAGQAKKKK of *Xenopus* nucleoplasmin[97, 103].

1.5.2 Nuclear Import Process

The steps of nuclear protein import are represented in Figure 6. An adaptor protein Importin-α functioning as an NLS receptor recognizes NLS present on the substrates. Following recognition, Importin-α interacts with Importin-β forming an import complex (Importin-α/β heterodimer complex: cargo)[104]. Structural studies of Importin-α:cargo:Importin-β interaction have revealed two structural and functional domains on Importin-α vital for establishing the interactions: a C-terminal domain that consists of 10 tandem armadillo (ARM) repeats important for making contacts with the NLS within the substrate and an N-terminal Importin-β binding domain (IBB)[105-107]. The import complex docks on the NPC and crosses over towards the nucleus, a process facilitated by interaction between Importin-β with FG-domain on nucleoporins[108]. Once in the inner face of the NPC, Ran-GTP, a low molecular weight Ras-family GTPase, binds to Importin-β. The binding induces a conformational change of Importin-β causing dissociation of the import complex[109]. This results in the deposition of Importin-α and cargo in the nucleoplasm leaving behind Importin-β and Ran-GTP accumulated at the NPC. Eventually, Importin-β and Ran-GTP translocate into the

cytoplasm where Ran-GTP is hydrolyzed into Ran-GDP by action of a GTPase activating protein (GAP) freeing Importin-β[110].

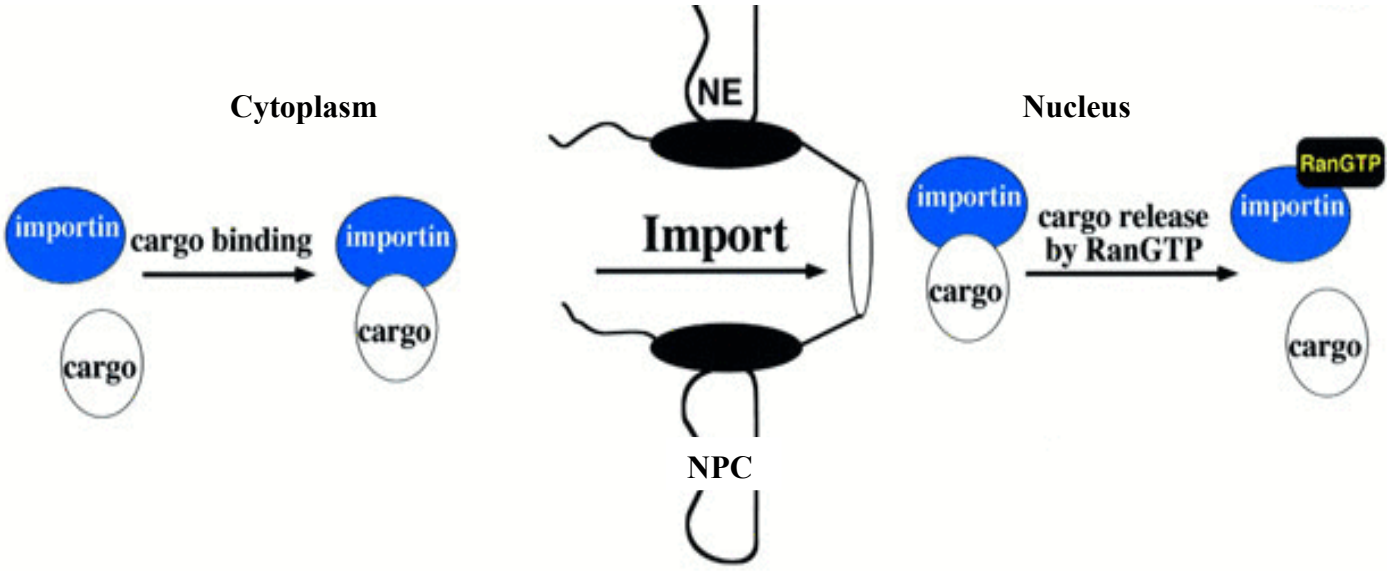


Figure 6. Classical nuclear protein import pathway via Importin proteins. The steps of the pathway are described in Section 1.5.2[4]. NE: Nuclear Envelope, NPC: nuclear pore channel. Adapted by permission from John Wiley and Sons: The EMBO Journal. Gorlich D., Copyright © 1998.

1.6 PNKP as a DNA Repair Enzyme

Often the ends generated at DNA breaks induced by endogenous and exogenous agents are required to be processed before making them compatible for gap tailoring and strand rejoining carried out by DNA polymerases and ligases. Frequently, IR, ROS and topoisomerase 1 inhibitor-induced strand breaks bear termini with enzymatically incompatible ends such as 5'-hydroxyl and 3'-phosphate, which need to be processed to generate ligatable ends, i.e. 5'-phosphate and 3'-hydroxyl termini. The dual functioning enzyme PNKP is a 57-kDa protein that contains both a kinase domain to phosphorylate 5'-OH to form 5'-P termini and a phosphatase domain to replace 3'-P with 3'-OH termini[111, 112]. PNKP is composed of three domains, an N-terminal forkhead-associated (FHA) domain and catalytically active phosphatase and kinase domains (Fig.7)[3]. The FHA domain is important for its role to interact with two critical scaffold proteins XRCC1 and XRCC4 in SSBR and NHEJ pathway, respectively[9]. Multiple studies conducted by our laboratory demonstrated that this interaction of the FHA domain of PNKP with XRCC1 and XRCC4 increases the turnover rate of PNKP at the DNA breaks speeding up the rate of end processing[9].

PNKP possesses a continuous domain containing catalytic phosphatase and kinase domain similar to T4 polynucleotide kinase (Fig 8 represents the molecular structure of murine PNKP which bears ~80% similarity with human PNKP)[3]. The phosphatase domain of PNKP belongs to the haloacid dehalogenase (HAD) superfamily with a conserved DxDGT motif [3, 113]. The first aspartate residue of the DxDGT motif forms a covalent bond to the substrate to generate a phospho-aspartate intermediate. Asp170 within the protein binds the substrate to facilitate reaction progression to facilitate

3'phosphatase activity of PNKP.[114].

The kinase domain of PNKP consists of a 5-stranded parallel β-sheet belonging to the adenylate family of kinases. ATP binding is required for kinase function of PNKP and is characterized by two Walker motifs, A and B[3]. The Walker motif A interacts with the β- and γ-phosphates of ATP and Asp 421 and Ser 378 residues of Walker motif B form a hydrogen bond aiding the positioning of Mg^2.[3]. PNKP kinase action is mediated by Asp 396 where it activates the 5'-OH for nucleophilic attack[114].

The phosphatase activity of PNKP has been shown to be more active than the kinase activity. Therefore, the phosphatase activity takes precedence in case of breaks where both 5'-OH and 3'-P termini are present[115].

PNKP's role in nuclear repair pathways like base excision/single strand-break repair (BER/SSBR) and non-homologous end joining (NHEJ) for DSB repair has been well established (reviewed in [9]). Figure 10 displays all the DNA repair pathways where PNKP has been shown to participate as an end-processing enzyme. It has been proposed that PNKP operates at a majority of breaks caused by the ROS attack on DNA bases since 50-70 per cent of those break harbor 3'P termini[116, 117]. Additionally, similar termini were reported to be generated due to abortive Top1 activity and actions of bifunctional glycosylases during BER[118, 119].

Shen et al in their report described the linkage of certain mutations on PNKP with severe neurological autosomal recessive disease (MCSZ) characterized by microcephaly and seizures[120]. A later study conducted by Reynolds et al investigated the impacts of these mutations on the specific activities of PNKP. They concluded that the mutations

critically reduced the stability and cellular levels of PNKP lowering the rates of DNA repair[117].

DNA repair proteins have been one of the favorite targets in cancer therapy. DNA damage and repair is a critical aspect when it comes to sensitization of tumor cells to radiation and chemotherapy. Recently, we have reported success with a small molecule inhibitor targeting the phosphatase activity of PNKP. The inhibitor enhanced the sensitivity of cells when exposed to IR and the topoisomerase I poison, camptothecin, which is the parent compound of irinotecan and topotecan, two clinically used topoisomerase I poisons used to treat colon and ovarian cancers, respectively[121, 122].

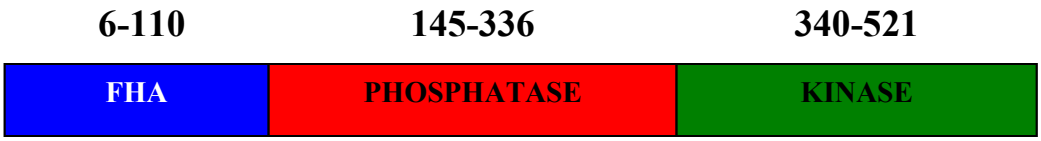


Figure 7: Subdomains of PNKP. Described are FHA (blue), phosphatase (red) and kinase (green) subdomains. The amino acid designation of each subdomain is represented by numbers mentioned on top of them.

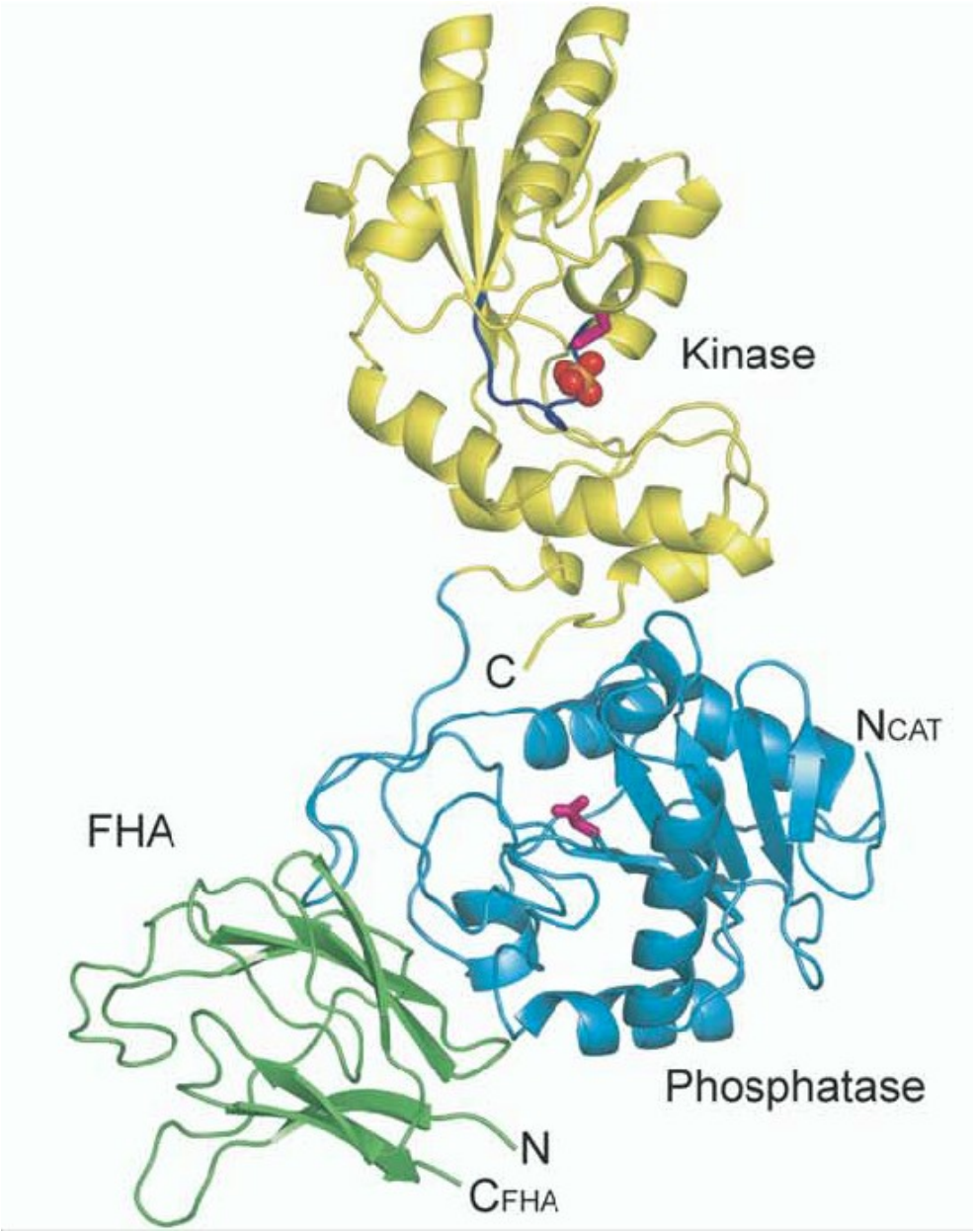


Figure 8: Molecular structure of murine PNKP. Murine PNKP shares ~80% similarity with human PNKP. FHA subdomain is depicted in green and catalytic kinase and phosphatase subdomains are depicted in blue and yellow, respectively[3]. Reproduced by permission from Elsevier: Molecular Cell. Bernstein N.K. et al., Copyright © **2**005

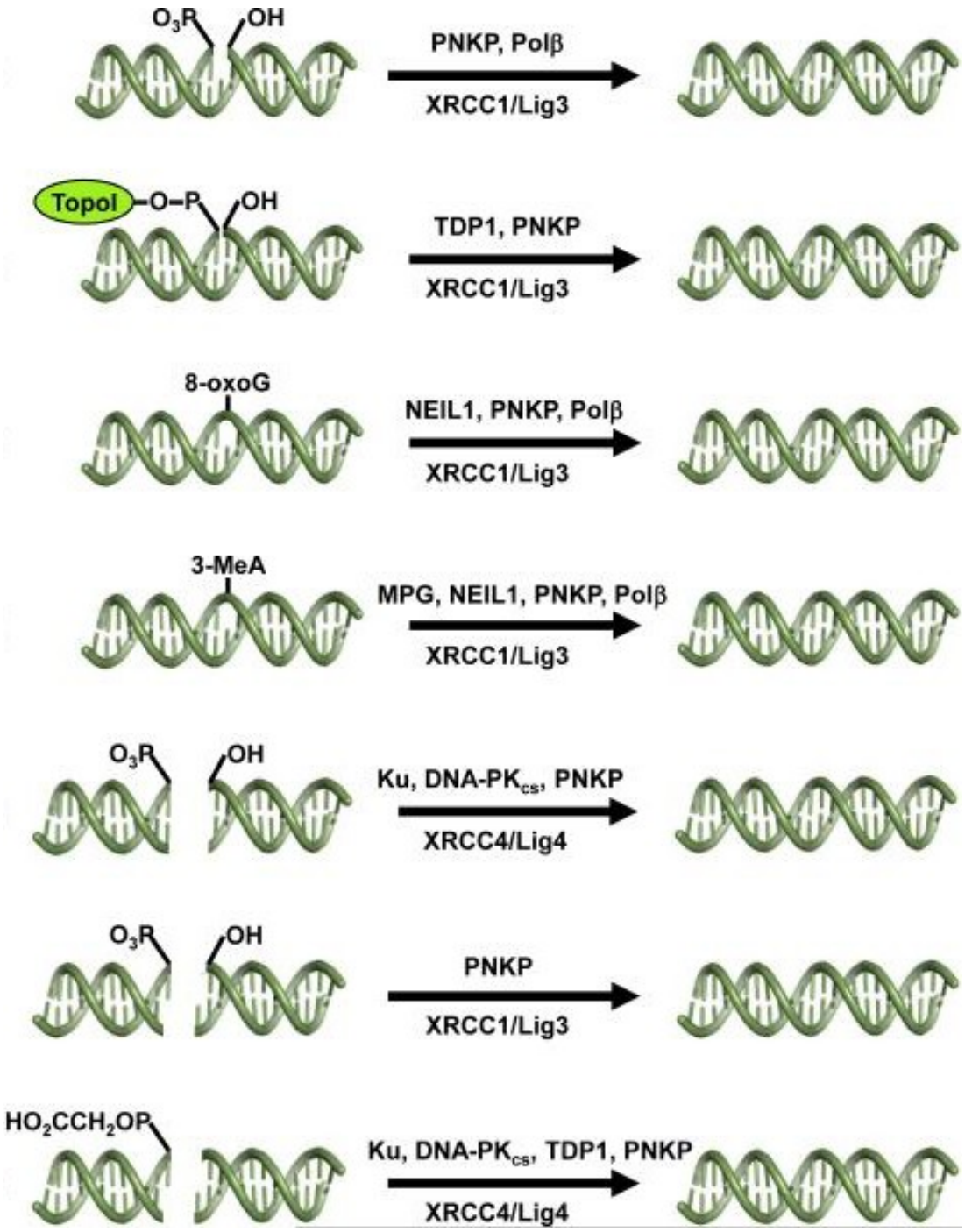


Figure 9: Various DNA repair pathways where PNKP is involved as an end-processing enzyme[9]. Reproduced by permission from Elsevier: Trends in Biochemical Sciences. Weinfeld M. et al., Copyright © **2**011

1.7 Synthetic Lethality

Although synthetic lethality has been gaining interest in the field of cancer therapy in recent years, the concept exists since Dobzhansky described it in his genetic studies carried out in *Drosophila* in 1946[123]. He described synthetic lethality as a lethality induced as a result of disruption of two non-allelic, non-essential genes or their proteins in the same cell. In other words, when loss of function of two different genes in a cell is not lethal on their own but when combined together, induces lethality that is termed as Synthetic lethality. Usually, a cell tolerates loss of one gene function by relying on a redundant pathway. However in cancer cells, loss of housekeeping or caretaker genes is common. The concept has been adopted in cancer therapy in the light of the fact that human malignancies are genetically inept in one or more pathways, which make them particularly vulnerable to synthetic lethal approaches. Therefore in principle, a drug can be used to induce specific lethality in cancer cells by targeting a synthetic lethal partner protein of a protein that is already disrupted during tumor progression. It is an attractive therapy approach since the targeted drug selectively kills cancer cells while sparing the normal ones.

Recent successes with synthetic lethality approach have been reported with breast and ovarian cancer cells harboring defective copies of breast cancer susceptibly genes, BRCA1 and BRCA2[124, 125]. In the absence of functioning BRCA proteins, cells are unable to execute HR pathway to repair DSBs. The teams led by Ashworth and Helleday used inhibitors against PARP1, an important protein involved in SSB repair, to achieve lethality selectively of BRCA 1/2 - defective cancer cells. The proposed mode of synthetic lethality in this case was by generating DSBs through PARP inhibition. As

SSBs persist in the absence of PARP1, replication forks collapse during S-phase converting SSBs into DSBs. In BRCA-less (and thus HR-less) cells, DSBs would accumulate causing cells to die through apoptosis or necrosis. Normal cells possessing wild type BRCA proteins would still be able to repair their DSBs and therefore are more likely to survive PARP-inhibitors[124, 125]. Multiple clinical trials that have been carried out established the clinical significance of PARP1-inhibitors in cancer therapy[126, 127].

The concept has been extended to other DNA repair pathway proteins. Our lab has recently published two reports identifying synthetic lethal partners of PNKP. A transfection-based genetic screen using a library of siRNAs targeting 6961 genes was conducted that identified 425 potential synthetically lethal partners to PNKP[128, 129]. One of them was SHP-1 (protein tyrosine kinase or PTPN6), a tumor suppressor protein known to be frequently mutated in malignant lymphomas, leukemias and prostate cancer[130]. SHP-1 depletion in cells has been associated with elevation in the levels of ROS. It was proposed that lethality was achieved due to overproduction of ROS-induced single strand DNA breaks in the absence of SHP-1. Since repair of ROS-induced DNA breaks often requires PNKP[9], PNKP depletion results in failure of repair of those breaks and conversion into DSBs inducing apoptosis[129].

A second synthetic lethal partner to PNKP identified was the tumor suppressor phosphatase and tensin homolog deleted on chromosome 10 (PTEN)[128]. PTEN is known to be the second most frequently disrupted gene after p53 in malignancies. PTEN is found to be compromised in various hereditary and sporadic cancers including glioblastoma, endometrial cancer and prostate cancer[131]. It was proposed that the co-

depletion/inhibition of PNKP and PTEN inflicts 'synthetic sickness' in cells enhancing their radiosensitivity. The study suggested potential success of radiation combined with PNKP-inhibitors in treatment of PTEN-deficient cancer cells[128].

1.8 Thesis Focus

The focus of part 1 of the thesis will encompass the strategy we developed in order to examine the relative importance of nuclear and mitochondrial PNKP on cell survival.

Studies implicate ROS to be a major agent to generate DNA SSBs (~10^4 SSBs per cell per day) during normal metabolism[132]. Since many of these breaks harbour unligatable termini such as 3'-P and 5'-OH, it is logical to assume that PNKP is one of the critical enzymes for maintenance of DNA integrity and cell survival. Because of the fact that mitochondria are a major source of ROS in mammalian cells, mtDNA molecules are considered to be under immense risk of sustaining the ROS-induced SSBs. Furthermore, since a full-length dually functional PNKP has been found to localize in mitochondria[8], we hypothesized that PNKP indeed plays a critical role in repairing ROS-induced SSBs in mtDNA and loss of mtPNKP therefore should severely impact the survivability of the cells. To pursue this, we decided to generate two different cell lines exclusively expressing PNKP either in the nucleus or mitochondria only. Once the cell lines are generated, the next step would be to challenge these cell lines with different ROS inducing agents such as IR and H_2O_2 and analyze the impact on cell survival (Fig. 11).

In the next two chapters, I will describe the experiments to identify the mitochondrial transport signal (MTS) and nuclear localization signal (NLS) of PNKP. Then follows the description of the strategy we undertook using these signals to generate cell lines with mtPNKP or nucPNKP. I will then explain the use of the pShooter vector system as an alternative approach to target PNKP to nuclei and mitochondria.

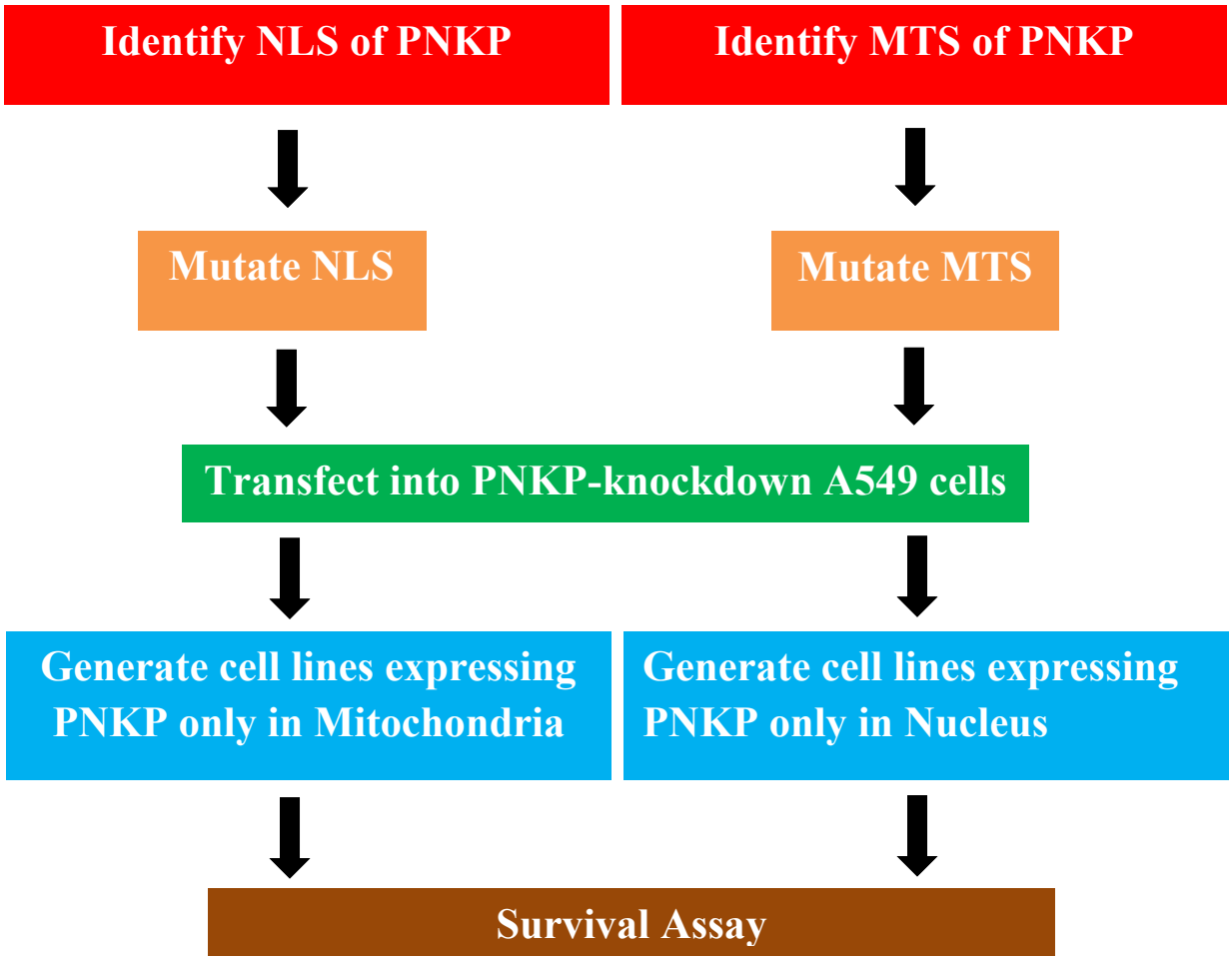


Figure 10. Thesis Focus.

Chapter II

Materials and Methods

2.1 Mitochondrial PNKP Experiments

2.1.1 Construct design and site directed mutagenesis

The constructs composed of the C-terminal region of PNKP, with and without the putative mitochondrial-targeting sequence, fused to GFP (CmtsPNKP + GFP and the CPNKP + GFP, respectively) were prepared using the pmEGFPN1 vector from Clontech.

To prepare CmtsPNKP + GFP, primer pair 1 (Table 1) was used. To prepare CPNKP + GFP, we used the reverse primer from CmtsPNKP + GFP and the forward primer no. 2 (Table 1). To mutate MTS in CmtsPNKP + GFP (mutCmtsPNKP + GFP) we used CmtsPNKP + GFP as a template. The site directed mutagenesis was accomplished using primer pair 3 (Table 1).

	Mutant	Forward primer	Reverse primer
1	**CmtsPNKP+GFP**	TGACTAAGCTTGCACCCAGGATGGCCAGGTACGTCCAGTGTG	GCTGATGGATCCCGGCCCTCGGAGAACTGGCAG
2	**CPNKP+GFP**	TGACTAAGCTTGCACCCAGGATGGGCGTCCCCTGCCGCTG	GCTGATGGATCCCGGCCCTCGGAGAACTGGCAG
3	**mutCmtsPNKP+GFP**	GCTCAAGCTTGCACCCAGGATGGACGGGGACGTCCAGTGTGCCCGAGCC	GGCTCGGGCACACTGGACGTCCCCGTCCATCCTGGGTGCAAGCTTGAGC

Table 1: Primers designed for MTS experiments.

2.1.2 MtDNA repair assay using XL-qPCR

H_2O_2 was added for 1h to the media of PNKP knock down (KD) A549 cells transfected with different constructs or vector only. Whole-cell (genomic and mitochondrial) DNAs were extracted using a miniprep kit according to the manufacturer's instructions (Qiagen) from untreated (control) and H_2O_2-treated cells after 0, 0.5, 2 or 4 h of repair. Extra-large qPCR (XL-qPCR) was performed using a Gene Amp kit (Applied Biosystems) following conditions and primers described[133]. Fluorescence quantification of XL-qPCR products was achieved using the Quant-iT Pico Green dsDNA assay kit (Invitrogen).

2.2 Nuclear Localization Signal (NLS) Experiments

2.2.1 Construct design and site directed mutagenesis

We generated the PNKP+GFP construct by cloning a full-length PNKP cDNA into pCMV–AC-mGFP vector (Origene, USA) using AsiSI and MluI restriction sites. The restriction enzymes were purchased from New England Biolabs.

We generated NLS-mutants of PNKP by site-directed PCR mutagenesis using QuickChange® II XL Site Directed Mutagenesis Kit (Agilent Technologies, USA) following the manufacturer’s protocol. The mutants and primer pairs to generate them are listed in Table 2.

	Random Mutations		
	Mutant	**Forward primer**	**Reverse primer**
1	**NLS1 (R)**	CAACTGGGCCCCGGGGTGGATGATGAAAGACTTCTCCTGCGCCG	CGGCGCAGGAGAAGTCTTTCATCATCCACCCCGGGGCCCAGTTG
2	**NLS2 (R)**	GAGAGATGCTGAGCTGCCGGAGGAGGGTATGCGGAAGTCAAACCCCG	CGGGGTTTGACTTCCGCATACCCTCCTCCGGCAGCTCAGCATCTCTC
3	**NLS1/2 (R)**	1+2 (forward)	1+2 (reverse)
4	$NLS2_{down}$ **(R)**	GCCGGAGGAGGGTATGGCGGCGTCAAACCCCGGCTGG	CCAGCCGGGGTTTGACGCCGCCATACCCTCCTCCGGC
	Alanine-substituted Mutations		
	Mutant	**Forward primer**	**Reverse primer**
5	**NLS1**	CGGCCAACTGGGCCCCGGGGGCGGCGGCGGCAGACTTCTCCTGCGCCGATC	GATCGGCGCAGGAGAAGTCTGCCGCCGCCGCCCCCGGGGCCCAGTTGGCCG
6	**NLS2**	GAAGAGAGATGCTGAGCTGCCGGCGGCGGCTATGCGGAAGTCAAACCCCGGC	GCCGGGGTTTGACTTCCGCATAGCCGCCGCCGGCAGCTCAGCATCTCTCTT
7	$NLS2_{down}$	CCGGCGGCGGCTATGGCGGCGTCAAACCCCGGCTG	CAGCCGGGGTTTGACGCCGCCATAGCCGCCGCCGG
8	$NLS2_{up/down}$	CTGGTGTCCCAAGATGAGGCGGCAGATGCTGAGCTGCCG	CGGCAGCTCAGCATCTGCCGCCTCATCTTGGGACACCAG
9	**NLSΔ135-147**	CAACTTCTCCAAGTTCTCCTCAGCATCTCTCTTCTC	GAGAAGAGAGATGCTGAGGAGAACTTGGAGAAGTTG
10	**NLSΔ135-142**	GAAGAGAGATGCTGAGTCAAACCCCGGCTGGG	CCCAGCCGGGGTTTGACTCAGCATCTCTCTTC

Table 2: Primers designed for NLS experiments.

2.2.2. Fluorescence recovery after photobleaching (FRAP)

Cells were cultured on 35-mm glass bottom culture dishes (MatTek Corporation, USA) 24 h before the experiment. The following day, cells were transfected with DNA constructs PNKP+GFP and mutant-NLS2+GFP using Lipofectamine2000 (Invitrogen). 24 h post transfection, cells were placed on a heated stage (37°C) of a Zeiss LSM 510 multi-photon microscope (Carl Zeiss, Inc.). GFP fluorescence was bleached selecting the entire nucleus of the cell in focus using a 488-nm argon laser set to 100% output and then GFP recovery in the nucleus from the cytoplasm was recorded. GFP fluorescence imaging was recorded after excitation with a 488-nm argon laser using a 515–540 nm band-pass filter.

2.3 pShooter Experiments

2.3.1 Construct design

The pShooter vectors pCMV-myc-mito and pCMV-myc-nuc were purchased from Invitrogen. PNKP tagged with GFP was rescued from pCMV-AC-PNKP-mGFP construct and sub-cloned into the SalI site of the pShooter vectors. The mitochondria-targeted vector with an MTS was designated pShooter myc-Mito-PNKP-GFP and the nucleus-targeted vector containing an NLS was designated pShooter myc-Nuc-PNKP-GFP. The pShooter-myc-mito vector harboring PNKP with NLS2 mutations was designated pShooter myc-Mito-NLS2-GFP. DNA sequencing confirmed proper ligation and insert directionality.

2.3.2 Fluorescence assisted cell sorting (FACS)

HeLa cells were pre-plated in 6-well dishes (Corning Inc., USA) to ~80% confluency a day prior to transfection. Cells were transfected with 500 ng of pShooter myc-Mito-NLS2-GFP and 5 µl of Lipofectamine2000 per well. 24 h post transfection, cells were trypsinized, pelleted and re-suspended in Basic Sorting Buffer (1X Ca/Mg^{2+} free PBS, 1 mM EDTA, 25 mM HEPES pH 7.0, 1% FBS) and single cells were sorted into each well of a 96-well plate using a BD Influx cell sorter (BD BioSciences, USA). GFP signal was used as a fluorescent marker to sort cells. Two different 96-well plates were prepared containing cells either with high or medium GFP expression levels. The plates were incubated and cells were expanded to obtain individual clones.

2.4 Fixed Immunofluorescence

A549 and HeLa cells were grown to 20% confluency on coverslips 24 h before fixing. Cells were fixed with 4% paraformaldehyde (PFA) in PBS, pH 7.5, for 5 min at room temperature. Cells were then permeabilized with PBS containing 0.5% Triton X-100 for 5 min, and then washed with PBS three times. Alternatively, cells were fixed and permeabilized using methanol:acetone (1:1) organic solvent mixture. Following permeabilization, cells were incubated with primary antibody (anti-VDAC, 5 µg/ml, Abcam; anti-mitofilin, 2 µg/ml, Calbiochem; H101[134], 5 µg/ml, Weinfeld lab) for 1 h at room temperature. Cells were then rinsed with PBS containing 0.1% Triton X-100 and then washed three times with PBS. Cells were subsequently incubated with a secondary antibody (Alexa Fluor® 488 goat anti-mouse IgG (H+L), 2 µg/ml, Invitrogen; Alexa

Fluor® 555 goat anti-rabbit IgG (H+L), 2 μg/ml, Invitrogen) for 1 h at room temperature. Next, cells were rinsed with PBS containing 0.1% Triton X-100 and washed three times with PBS. Finally, cells were mounted onto slides with a 90% glycerol/PBS-based medium containing 0.5 μg/mL DAPI. Cells were observed using a Zeiss LSM 710 laser scanning confocal microscope (Carl Zeiss, Inc.).

2.5 Live immunofluorescence

HeLa or A549 cells were grown in DMEM/F12 (1:1) plus 10% fetal bovine serum on 35-mm glass bottom culture dishes (MatTek Corporation, USA). The following day, cells were transfected with DNA constructs using Lipofectamine2000. 24 h after transfection, cells were incubated with Hoechst 33258 for nuclear staining and/or MitoTracker Red (Life Technologies, USA) for mitochondrial staining to a final concentration of 2 μg/ml for 20 min and replaced with fresh medium for 10 min prior to imaging. Fluorescently stained cells were viewed on a Zeiss LSM 710 laser scanning confocal microscope mounted on an AxioObserver Z1 inverted microscope (Carl Zeiss, Jena, Germany) with a plan Apochromat 40 X (NA 1.3) oil immersion lens. Individual immunofluorescence channels were collected sequentially to avoid signal bleeding through. GFP signal was detected using 488-nm laser excitation and collected with a bandpass filter of 493–552 nm.

2.6 Western Blotting

Approximately 10^6 cells were washed twice with ice cold PBS and resuspended in RIPA buffer (50 mM Tris-HCl pH 7.6, 150 mM NaCl, 1% deoxycholate, 1% Triton X-100, 1 mM Na_3VO_4, 50 mM NaF, 1 mM phenylmethylsulfonyl fluoride, 1X Biotool Protease Inhibitor cocktail). Cells were then sonicated briefly and cell debris was spun down at 14,000 rpm for 20 min at 4°C. Determination of whole cell lysate protein concentration was then performed using the Bradford Assay.

50 μg of protein was added to 1X sample buffer (50 mM Tris-Hcl pH 6.8, 2% SDS, 10% glycerol, 1% β-mercaptoethanol, 12.5 mM EDTA, 0.02% bromophenol blue) and boiled for 5min. Samples were then separated by 10% SDS-PAGE (200 V for 55 minutes at room temperature) and transferred to a nitrocellulose membrane by wet transfer (100 V for 90 min at room temperature or 30 V overnight at 4°C). Non-specific binding was blocked with 5% skim milk in PBST (phosphate buffered saline Tween 20 solution) for 1 h at room temperature. Membranes were then immunostained with the following primary antibodies: mouse monoclonal anti-PNKP (home grown, diluted 1:300, overnight at 4°C), and goat anti-Actin (Santa Cruz Biotech, USA, 200 μg/ml diluted 1:2500, 1 h). Membranes were then subjected to 3 x 10 min washes in 5% milk before being incubated with the appropriate HRP-conjugated secondary antibody (Goat Anti-Mouse IgG, 28.4 mg/ml and Goat Anti-Rabbit IgG, 50 mg/ml, Jackson Immunoresearch Laboratory, PA) at a 1:5000 dilution in 5% milk for 45 min at room temperature. Membranes were then washed 3 x 10 min in 5% milk followed by a 30 min 1X PBS wash. Then the blots were incubated with 2 ml Lumi-Light Western Blotting substrate (Roche, Mississauga, ON) for 5 min before autoradiography.

2.7 DNA Sequencing

DNA sequencing was carried out using a BigDye terminator v3.1 Cycle Sequencing Kit (Applied Biosciences, UK). Approximately 500 ng of plasmid DNA was combined with 1.6 pmol of sequencing primer and added to the mix of 0.5 µl Big Dye Terminator and 1X Big Dye Buffer. On a thermocycler, the sequencing reaction mixture was kept at 96ºC for 1 min followed by 35 cycles of: 96ºC for 10sec, 50ºC for 5sec and 60ºC for 4 min, and finally held at 4ºC.

To precipitate DNA, 2.5 µl of 125 mM EDTA was added to each sample and mixed well. Then, 25 µl of absolute ethanol was added to the sample and held at room temperature for 15 min followed by centrifugation at 18000 *g* for 10 min at 4ºC. Ethanol was removed carefully from the sample, which was then washed with 70% ethanol. Ethanol was carefully drained and the samples were dried at 95ºC for 1 min on a heat-block. Precipitated DNA samples were resuspended in HiDye formamide injection buffer (Life Technologies, USA) by heating at 95ºC for 5 min. The samples were cooled on ice and loaded for sequencing on an ABI 310 genetic analyzer. The sequences obtained were collected using ABI PRISM® 310 DNA collection software and analyzed using DNA Strider software.

Chapter III

Results

3.1 Distribution pattern of PNKP under different fixation techniques

It has been well established by means of immunoblotting and immunofluorescence imaging that PNKP resides both in the nucleus and mitochondria[8, 111]. Different fixing and staining techniques affect the sub-cellular localization of proteins differentially. Therefore we employed two of the most commonly used fixation techniques, methanol: acetone (1:1) and 4% paraformaldehyde (4% PFA) to fix and stain A549 cells. We further diversified the analysis by using either mouse-monoclonal (homegrown) or rabbit-polyclonal (Sigma) antibodies raised against PNKP to observe the localization pattern of endogenous PNKP. In the experiment, Mitofilin and VDAC were used as mitochondrial markers and DAPI was used to stain the nucleus. As seen in Fig 12, we observed PNKP to localize predominantly in mitochondria in case of both monoclonal and polyclonal antibodies under methanol:acetone fixation although nuclear signal was also visible in lower intensity. Conversely, 4% PFA fixation showed a predominant PNKP localization in the nucleus. The results are summarized in Table 3.

Fig. 11

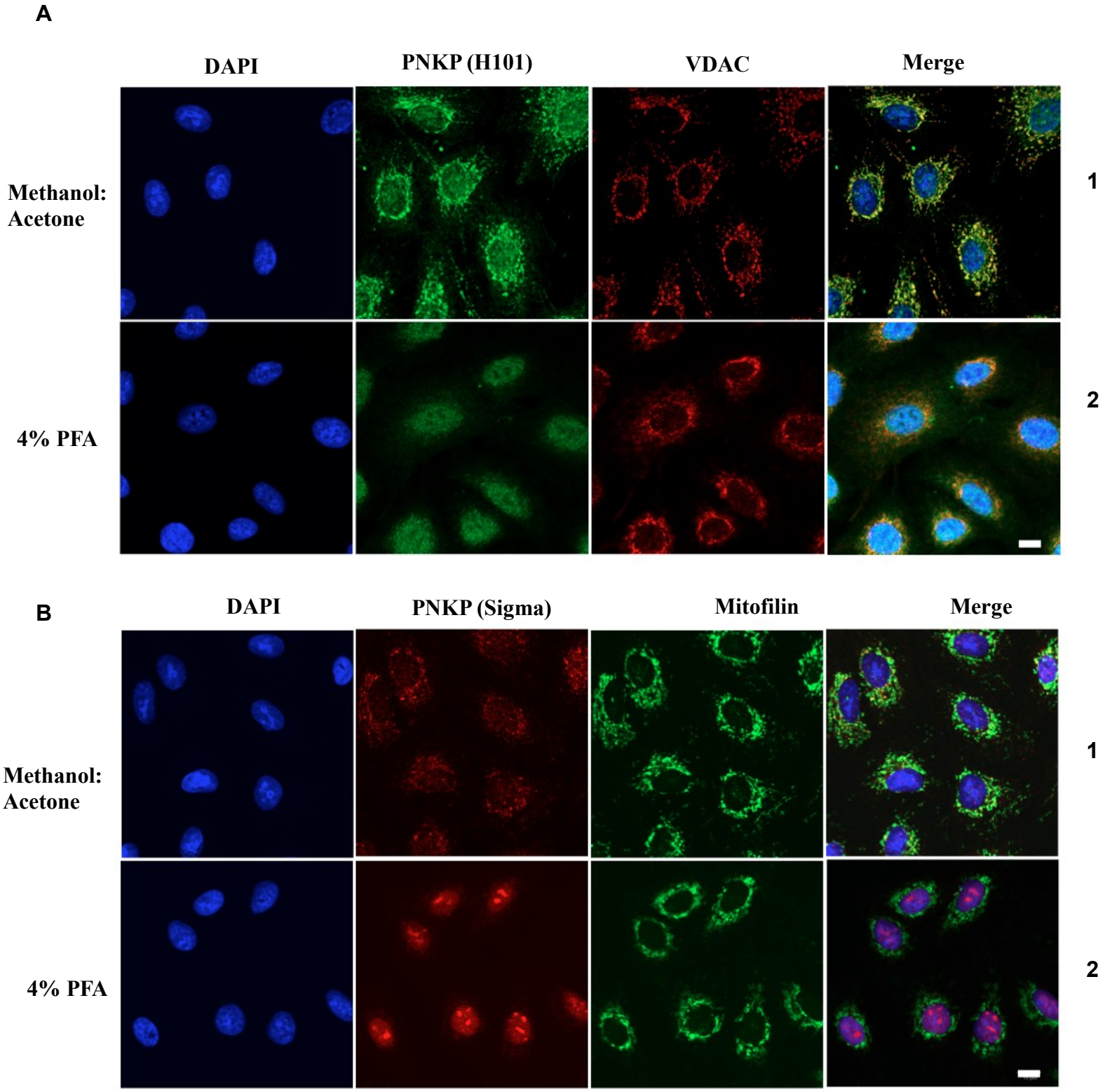

A
DAPI
PNKP (H101)
VDAC
Merge
Methanol: Acetone
1
4% PFA
2
B
DAPI
PNKP (Sigma)
Mitofilin
Merge
Methanol: Acetone
1
4% PFA
2

Figure 11. Localization pattern of endogenous PNKP under different fixation-staining approaches: A549 cells were grown on coverslips overnight and fixed either with (A) methanol:acetone (1:1) or (B) 4% paraformaldehyde (4% PFA). Either (1) H101, monoclonal or (2) Sigma, polyclonal anti-PNKP was used for staining. VDAC and Mitofilin were used as mitochondrial markers and DAPI was used to stain the nucleus.

Methanol:acetone fixation: Panel 1A shows PNKP localizing predominantly in the mitochondria with a relatively weaker signal in the nucleus, when anti-PNKP (mono) is used. This result is comparable with the results obtained with anti-PNKP (poly), as seen in Panel 1B. However, the nuclear signal in this case was relatively not as intense.

4% PFA fixation: Panel 2A shows a predominant nuclear signal of PNKP in the nucleus with a minimal mitochondrial signal, when anti-PNKP (mono) is used. This result is almost identical with the result obtained with anti-PNKP (poly), as shown in Panel 2B. Scale bar: 10 μm.

Antibody	Antibody type	Manufacturer	Fixation technique	Predominant PNKP localization
H101	Mouse monoclonal	Weinfeld Lab	4% PFA	Nuclear
H101	Mouse monoclonal	Weinfeld Lab	Methanol:Acetone (1:1)	Mitochondrial
αPNKP	Rabbit polyclonal	Sigma	4% PFA	Nuclear
αPNKP	Rabbit polyclonal	Sigma	Methanol:Acetone (1:1)	Mitochondrial

Table 3: Summary of localization pattern of endogenous PNKP under different fixation-staining approaches

3.2 Mitochondrial PNKP Experiments

3.2.1 PNKP contains a functional MTS close to its C-terminus

Mitochondrial-targeting signals (MTS) are most frequently found at the N-termini of proteins, but some of the mtDNA repair proteins reported to date do not contain amino-terminal MTS[8]. A computer-based analysis using multiple programs (including Mitoprot II and Predotar) showed that PNKP does not contain a canonical (N-terminal) MTS. However, a closer inspection of the sequence of PNKP revealed the presence of a 'cryptic' MTS close to the carboxy-terminus of the protein. This putative MTS consisted of amino acids ^{432}ARYVQCARAA^{441} and was identified by several computer programs, as mentioned. To determine if this MTS is functional, we fused the carboxy-terminus of PNKP with (CmtsPNKP) or without the MTS (CPNKP) to GFP, generating CmtsPNKP + GFP and CPNKP + GFP (Fig 13A). The constructs were then transfected into A549 (Fig. 13B). In all cases an additional methionine was added to the amino terminus of the fusion protein. The carboxy-terminal region of PNKP incorporating the MTS was functional as a mitochondrial-localization signal and transferred most of the expressed CmtsPNKP + GFP into the mitochondria (Fig. 13B, panel 1). However, when the MTS was not included with the carboxy-terminal region of PNKP (CPNKP + GFP), the GFP construct was not detected in the mitochondria (Fig 13B, panel 2), a result similar to the situation seen with GFP alone (Fig. 13B, panel 4). Further confirmation was provided by mutating the first three amino acids of the MTS as follows: A432D, R433G and Y434D. Computer analysis indicated that these mutations would dramatically decrease the capacity of the identified MTS to function as a true mitochondrial-

localization signal. We observed that the protein expressed by this mutated construct (mutCmtsPNKP + GFP) failed to localize to the mitochondria (Fig. 13B, panel 3).

To test if the MTS is functional in the context of full-length PNKP, XL-qPCR was used to compare the DNA repair functionality of HA-tagged-PNKP-mts or HAPNKP-mts (MTS-mutated form of HAPNKP incorporating the same mutations to the MTS as described above) to HAPNKP in mtDNA repair. XL-qPCR assay is based on the principle that many DNA lesions can slow down or block the progression of DNA polymerase. So a DNA sample that has sustained lesser amount of lesions will be amplified via PCR in greater amounts relative to the DNA with more lesions. That is to say, a DNA sample extracted from H_2O_2-treated cells will produce lesser PCR products than the DNA obtained from untreated cells[133]. HAPNKP (positive control), vector only and HAPNKP-mts were transfected into PNKP KD cell lines prepared from A549 cells. As shown in Fig. 14, the loss of the MTS resulted in a clear decrease in the activity of PNKP during mtDNA repair, similar to transfection with the vector only. The control western blot showed that the level of expression of HAPNKP-mts was the same as HAPNKP and thus the observed effect was not due to a lower level of protein expression of the mutated form of PNKP. The data shown is based on three independent experiments.

Fig. 12

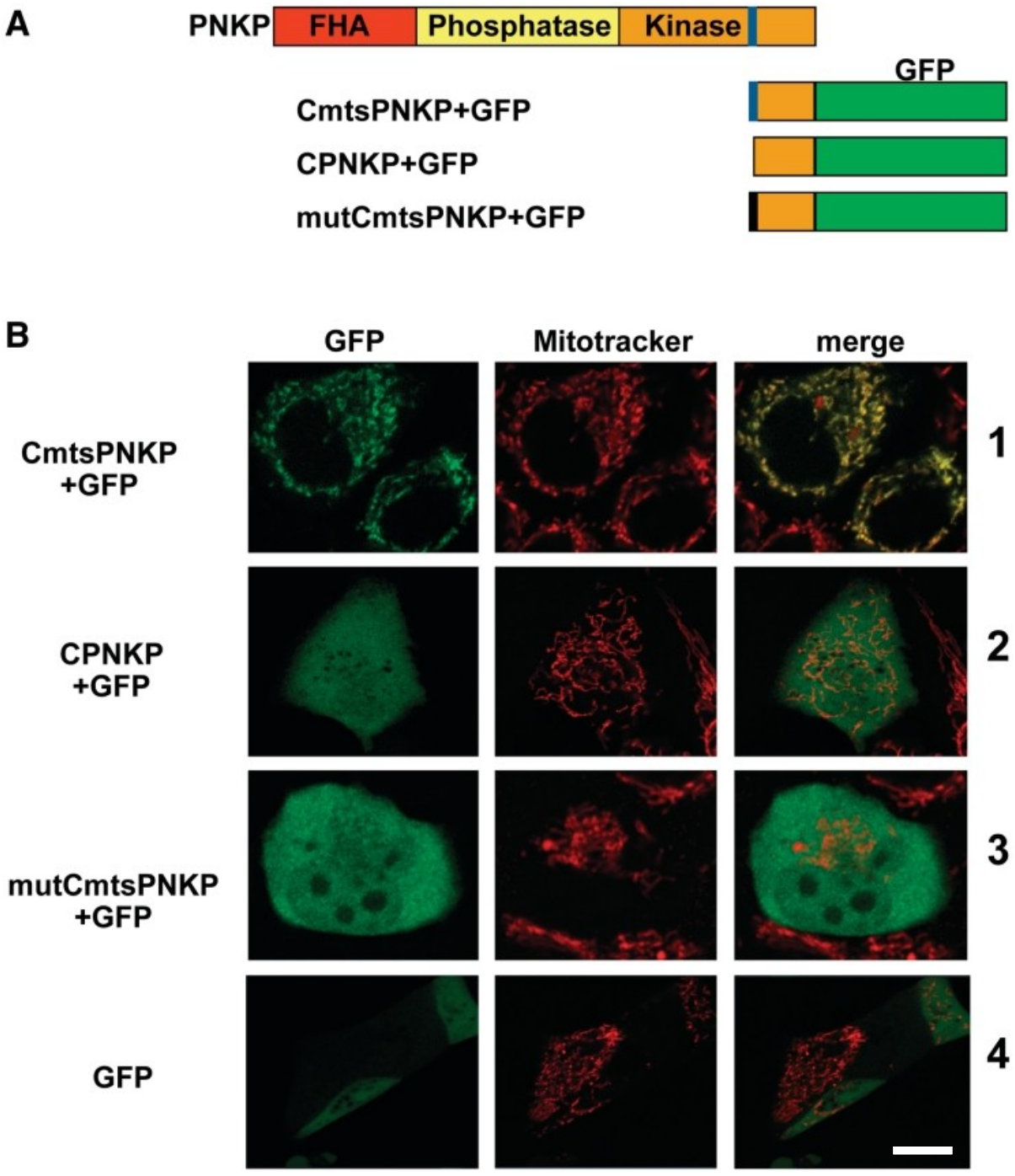
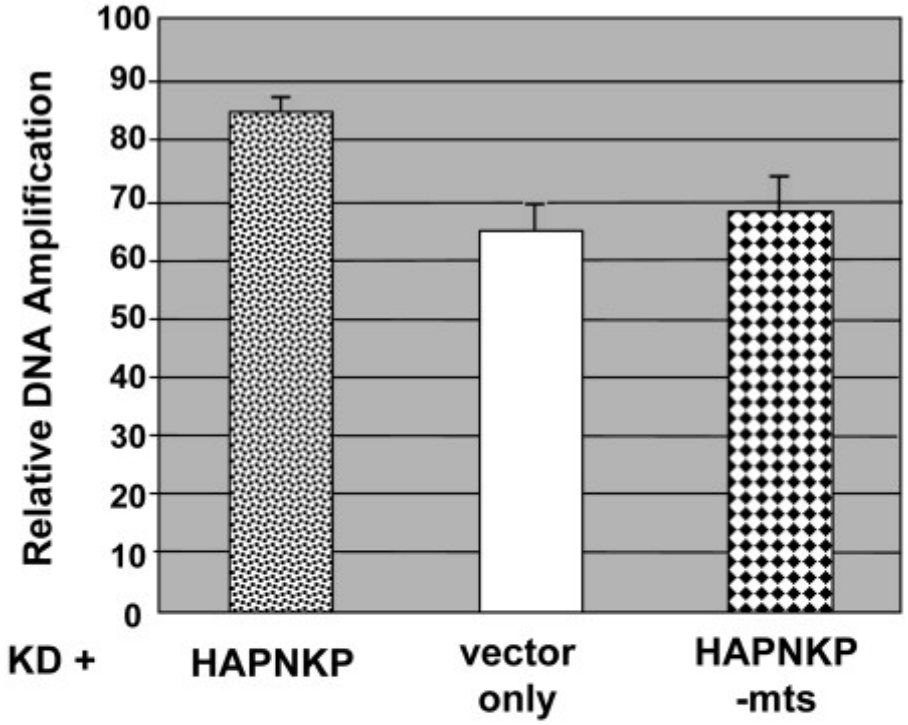

A
PNKP
FHA
Phosphatase
Kinase
GFP
CmtsPNKP+GFP
CPNKP+GFP
mutCmtsPNKP+GFP
B
GFP
Mitotracker
merge
CmtsPNKP +GFP
CPNKP +GFP
mutCmtsPNKP +GFP
GFP
1
2
3
4

Figure 12. Mitochondrial localization of PNKP is dependent on the presence of a mitochondrial-targeting signal (MTS) in proximity to its carboxy terminus[8]. (A) Computer programs (Mitoprot, Psort II and Predotar) predicted the presence of a mitochondrial-targeting signal (MTS) close to the C-terminus of PNKP (shown in blue). To further examine this potential MTS we generated three constructs (i) CmtsPNKP + GFP containing the GFP fused to the PNKP C-terminus including the putative MTS, (ii) CPNKP + GFP, which is essentially the same as CmtsPNKP + GFP but lacking the MTS sequence and (iii) mutCmtsPNKP + GFP, a mutated form of CmtsPNKP + GFP with the first three amino acids of the putative PNKP MTS mutated as follows: A432D, R433G and Y434D. In all cases a methionine was included at the amino-terminus. (B) The constructs were transfected into A549 cells and the cellular localization of GFP was monitored. Only the GFP fusion protein containing the wild-type MTS localized to mitochondria (panel 1) as shown by colocalization with Mitotracker Orange (Molecular Probes). CPNKP + GFP, mutCmtsPNKP + GFP and GFP alone showed a diffuse signal throughout the cell (panel 2–4). Scale bar: 20 μm. Reproduced by permission from Oxford University Press: Nucleic Acids Research. Tahbaz N., Subedi S. and Weinfeld M., Copyright © **2**011.

Fig. 13


Relative DNA Amplification
100
90
80
70
60
50
40
30
20
10
0
KD +
HAPNKP
vector only
HAPNKP -mts

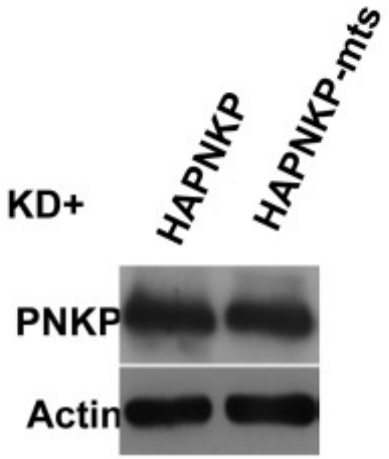

KD+
HAPNKP
HAPNKP-mts
PNKP
Actin

Figure 13. The MTS of PNKP is required for its function in mtDNA repair[8]. XL-qPCR was used to monitor mtDNA repair in PNKP-depleted A549 cells treated with hydrogen peroxide as described in Section 2.1.2. Transient complementation of the cells with PNKP mutated in the first three amino acids of the MTS (HAPNKP-mts), as opposed to the wild-type protein (HAPNKP), reduces DNA repair in mitochondria to a level similar to the vector only control. The western blot of whole-cell extracts shown at the bottom of the figure indicates that similar levels of HAPNKP-mts and HAPNKP proteins were expressed in the A549 cells.

3.3 PNKP Nuclear Localization Signal (NLS) experiments

While the MTS of PNKP was firmly established, two putative NLSs were identified *in silico* using an NLS predicting online tool NucPred. One of these sequences, NLS1, resided at residues 301-304, (RKKK) and the second, NLS2, at 137-139 (KKR). NLS1-residues were also predicted as a putative NLS in a previously conducted study[111]. To further investigate these putative NLS sites, PNKP tagged with green fluorescent protein (GFP) was mutated at the sites using site directed mutagenesis and then transfected into HeLa cells. GFP was then tracked for localization using live confocal fluorescent microscopy. In the first phase of the experiment, we either created mutants by introducing random or alanine-substituted mutations at the NLSs. Figure 15 is a map showing all the regions within full-length PNKP that were considered during our investigation of NLS.

Live imaging of the cells clearly indicated that the residues at NLS1 did not contribute to nuclear localization of PNKP (Fig. 16, Panel 2). As observed, NLS1 showed a robust localization in the nucleus identical to a wild type PNKP that was used as a positive control (Fig. 16, Panel 1). NLS1 was therefore eliminated as a putative NLS of PNKP.

Interestingly, residues at NLS2 that were not predicted as a putative NLS previously, when mutated showed a very diverse localization pattern of PNKP-GFP in the cells (Fig. 16, Panel 3). A majority of cells showed a uniform distribution of GFP throughout the cells (nucleocytoplasmic distribution). On the other hand, the cells that displayed an overexpression of GFP showed a very bright signal in the nucleus as well as

in the cytoplasm. A third category of GFP expressing cells showed more cytoplasmic localization of GFP and a substantially diminished GFP signal in the nucleus. The data therefore indicated that NLS2-residues do contribute to localization of PNKP in the nucleus. Although the localization pattern was disrupted due to NLS2 mutation, it did not result in complete sequestration of PNKP in the cytoplasm. Therefore we reasoned NLS2 residues might be part of a bipartite NLS.

Considering NLS2 residues could be a part of a potential bipartite NLS, we created additional mutations at the residues R141, K142 present downstream ($NLS2_{down}$) and K130, R131 upstream ($NLS2_{up/down}$) of the NLS2 residues. In general, a bipartite NLS is composed of two basic amino acids, a spacer region of 10–12 amino acids and a basic cluster in which three of five amino acids are basic[135]. The $NLS2_{up/down}$ mutant generated was then transfected into HeLa cells and tracked for GFP localization. The localization patterns observed with these mutants were identical to the NLS2 mutants (Fig. 16, Panel 4).

In addition, we employed cNLS Mapper, an *in silico* approach that can predict bipartite as well as monopartite NLSs. While cNLS Mapper did not indicate any presence of a bipartite NLS, it pointed out an entire PNKP sequence of ^{135} LPKKRMRKSNPGW ^{147} to be the potential NLS of PNKP. In order to experimentally verify the prediction, we created a deletion mutant of PNKP by removing all the residues in question. Thus constructed PNKP-GFP-NLSΔ135-147 was tracked for localization. Interestingly, ~90% (based on observation) of the transfected cells showed nuclei completely devoid of PNKP indicating the deleted sequence to be the potential NLS of PNKP (Fig. 17).

One major concern with such a large deletion of resides (13-residues) was the potential impact that could cause on the protein structure, i.e. the deletion could affect protein folding as well as result in mis-localization of protein within the cell. Amongst the residues that were deleted, residues 140-147 pack in the catalytic domain and W147 packs in the hydrophobic core of PNKP. In order to address this concern, we decided to create a second deletion mutant (PNKP-GFP-NLSΔ135-142) with deletions from residue 135 to 142 (^{135}LPKKRMRK^{142}) sparing the aforementioned critical residues. . In addition, we also introduced identical deletion on PNKP cDNA that was cloned into a pShooter vector (more about pShooter vectors to follow in Section 3.4). To our surprise, the construct failed to express any protein (Fig. 18, lane 2-5) whereas the bigger deletion PNKP-GFP-NLSΔ135-147 was expressed (Lane 6) at similar levels to wtPNKP (Lane 1). The results were verified twice with live immunofluorescence and immunoblots. Table 4 summarizes the nature of mutations introduced in the NLS regions and their outcome on PNKP localization.

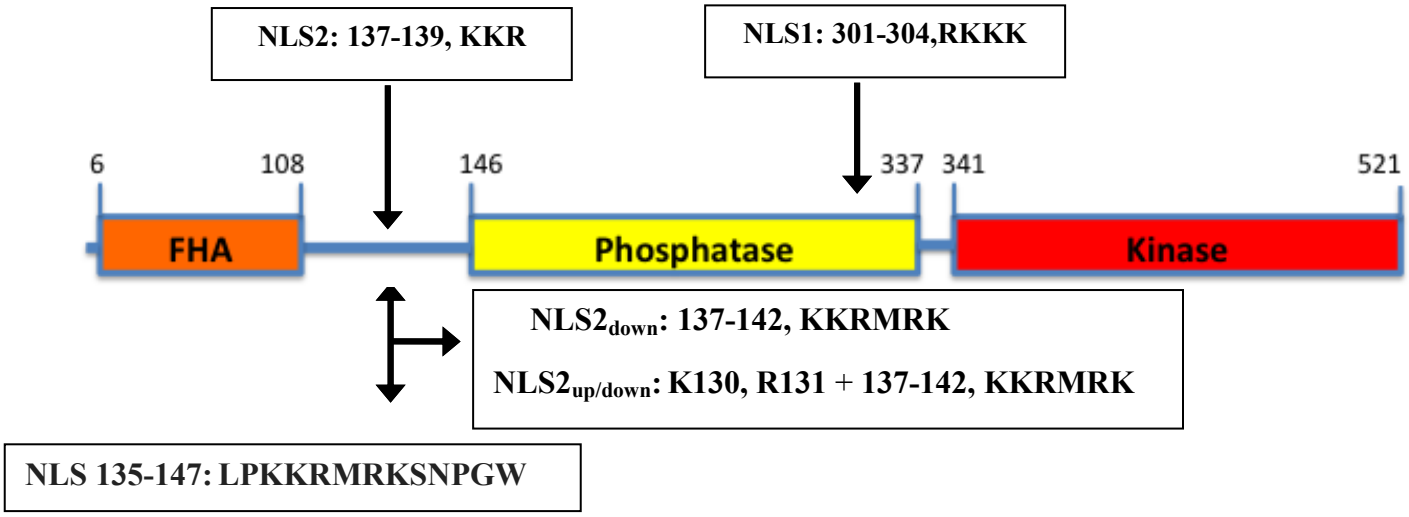


Figure 14. Mapping the putative nuclear localization signal (NLS) of PNKP. Computer program, NucPred predicted two Lys-Arg rich regions as the potential (NLSs), one within the phosphatase domain (NLS1) and the second within the linker region (NLS2) of PNKP. Additionally, Lys-Arg residues proximal to NLS2 were also considered ($NLS2_{down}$ and $NLS2_{up/down}$) presuming the presence of a bi-partite NLS. cNLS Mapper, another computer based NLS-predicting program, did not indicate the presence of a bi-partite NLS but indicated the presence of a larger stretch of residues (NLS 135-147) as a putative NLS of PNKP. Positions of all the residues within the predicted regions are indicated on a full-length PNKP.

Fig. 15

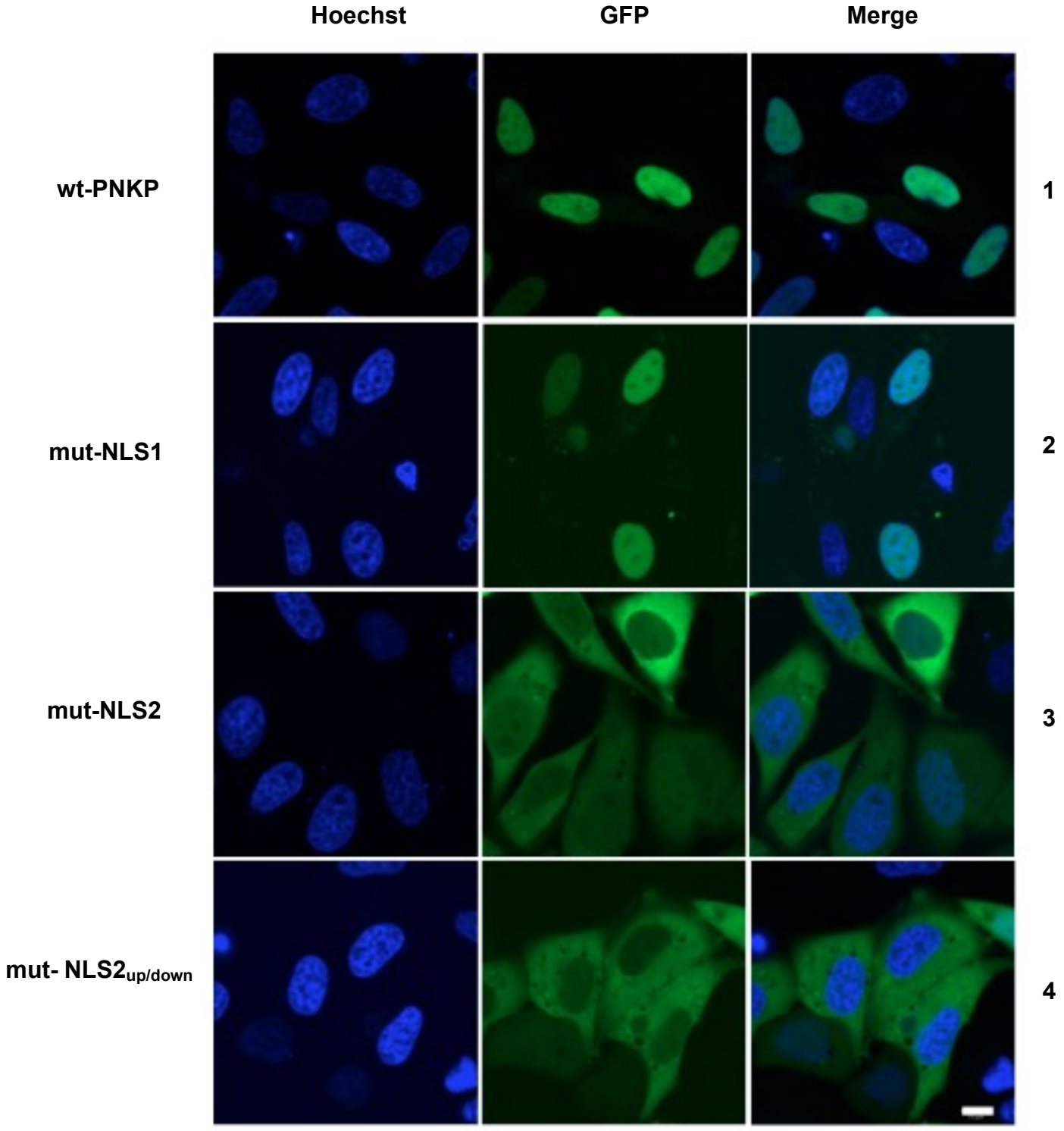

Hoechst
GFP
Merge
wt-PNKP
1
mut-NLS1
2
mut-NLS2
3
mut- NLS2up/down
4

Figure 15. The NLS may be present close to the N-terminus of PNKP. Based on the predictions of the computer program NucPred, GFP-tagged constructs of PNKP were generated by introducing either random or alanine-substituted mutations at the putative NLS sites. The sites of putative NLSs where the mutations were introduced are described in Section 3.3. The constructs were transfected into HeLa cells and cellular localization of GFP was monitored. GFP fusion protein containing the wild-type NLS (PNKP+GFP) localized predominantly in the nucleus as expected (Panel 1) as shown by co-localization pattern with Hoechst-dye. NLS1-mutant protein showed a localization pattern identical to the wild-type PNKP, as shown in Panel 2. However, NLS2-mutant showed a diverse localization pattern (Panel 3): the majority of cells showed a uniform or diffuse pattern of GFP throughout the cells (nucleocytoplasmic distribution), a second category of cells that were overexpressing GFP displayed a very bright signal in the nucleus as well as in the cytoplasm, and a third category of cells showed more cytoplasmic localization of GFP and a substantially diminished GFP signal in the nucleus. $NLS2_{up/down}$-mutant, bearing additional mutations at the Lys-Arg residues upstream and downstream of the NLS2, showed a localization pattern identical to the NLS2-mutant (Panel 4). Scale bar: 10 μm.

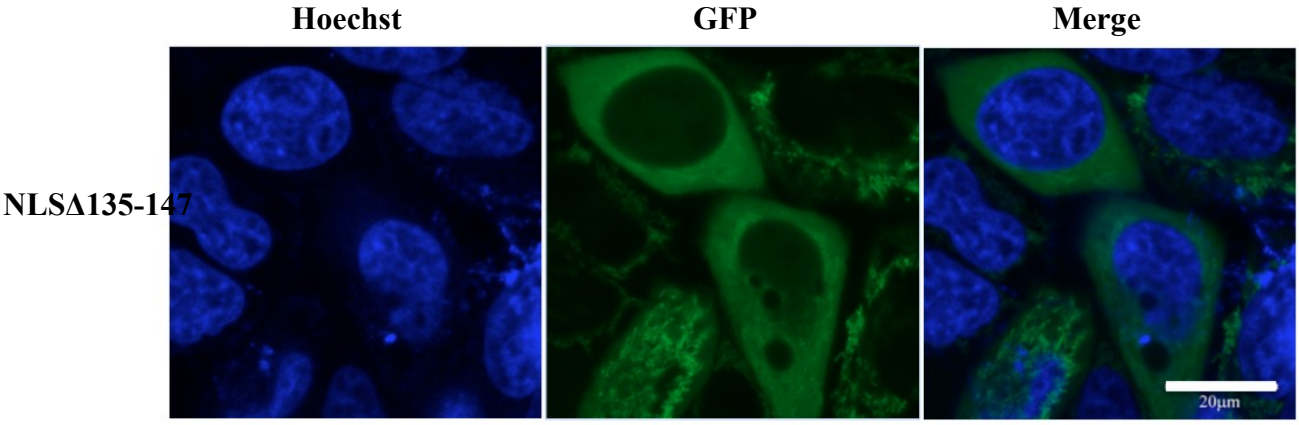


Figure 16. NLS-deletion mutant resulted in a complete sequestration of PNKP in the cytoplasm. An independent experiment was conducted where the ^{135}LPKKRMRKSNP^{147} residues of PNKP were deleted. The resulting construct NLSΔ135-147 was transfected into HeLa cells and GFP was tracked for PNKP localization. As shown, the deletion caused a complete sequestration of PNKP in the cytoplasm. Almost 90% of cells showed nuclei completely devoid of GFP with a few cells still showing a diffuse pattern of GFP. Scale bar: 20 μm.

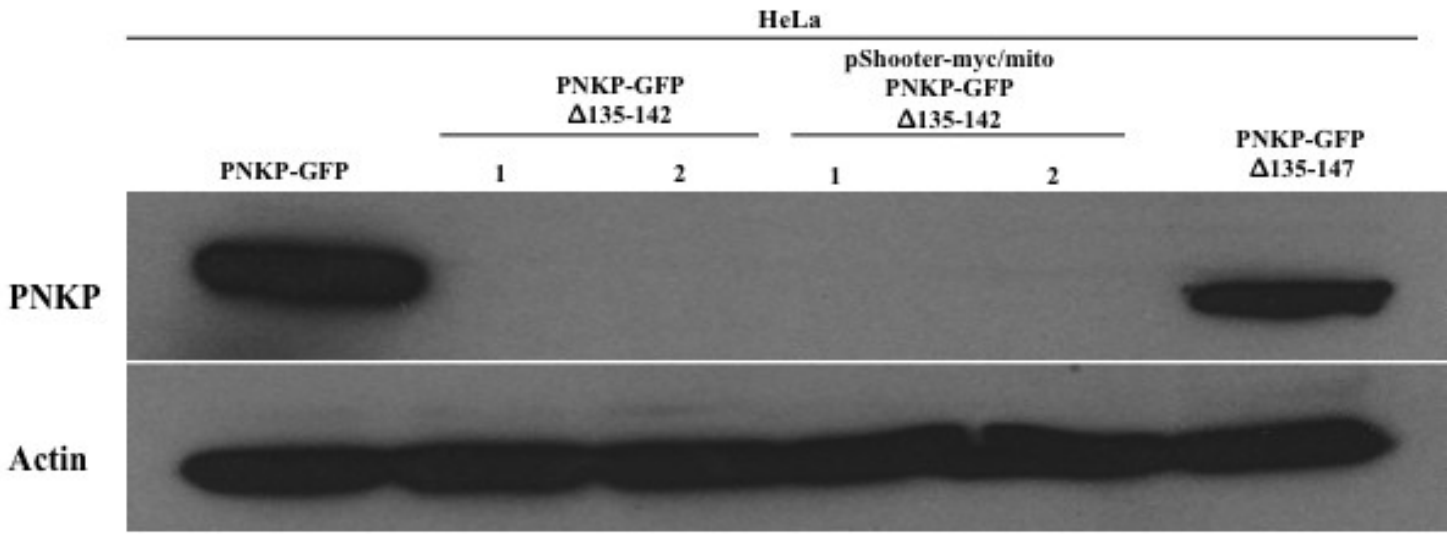


Figure 17. Expression levels of NLS-deletion mutants: Pre-plated HeLA cells were transfected with vectors carrying wild-type PNKP+GFP, deletion mutants PNKP-GFP NLSΔ135-147 and NLSΔ135-142. An additional vector pShooter-myc/mito (details on the vector available in Sections 2.3.1 and 3.4) bearing 135-142-deletion on PNKP (pShooter-myc/mito-NLSΔ135-142) was also included. The transfected cells were incubated overnight and whole cell lysates were prepared on the following day. 50 µg of cell lysates were loaded and electrophoresed in 10% SDS-PAGE gel. The proteins were transferred to a nitrocellulose membrane and immunostained with anti-PNKP antibody (H101). Actin was used as a loading control. As expected, the wild type PNKP showed a robust signal as seen in Lane 1. Lane 6 shows the expression level of Mutant PNKP-GFP NLSΔ135-147. Surprisingly, cells transfected with 135-142-deletion failed to show expression (Lanes 2-5).

Random Mutations			
Mutant	**Position on PNKP**	**Mutation**	**Localization**
NLS1 (R)	^{301}RKKK^{304}	^{301}WMMK^{304}	Nuclear
NLS2 (R)	^{137}KKR^{139}	^{137}EER^{139}	Nucleocytoplasmic
NLS1/2 (R)	^{301}RKKK^{304}/^{137}KKR^{139}	^{301}WMMK^{304}/^{137}EER^{139}	Nucleocytoplasmic
NLS2$_{down}$ (R)	^{137}KKRMRK^{142}	^{137}EERMAA^{142}	Nucleocytoplasmic
Alanine-substituted Mutations			
Mutant	**Position on PNKP**	**Mutation**	**Localization**
NLS1	^{301}RKKK^{304}	^{301}AAAK^{304}	Nuclear
NLS2	^{137}KKR^{139}	^{137}AAA^{139}	Nucleocytoplasmic
NLS2$_{down}$	^{137}KKRMRK^{142}	^{137}AAAMAA^{142}	Nucleocytoplasmic
NLS2$_{up/down}$	^{130}KR$^{131.137}$KKRMRK^{142}	^{130}AA$^{131.137}$AAAMAA^{142}	Nucleocytoplasmic
NLSΔ135-147	^{135}LPKKRMRKSNP^{147}	Δ^{135}LPKKRMRKSNPG^{147}	Cytoplasmic
NLSΔ135-142	^{135} LPKKRMRK^{142}	Δ^{135} LPKKRMRK^{142}	did not express

Table 4: A summary of localization of NLS mutants. The table displays the positions of the putative NLSs within PNKP, the mutations that were introduced in them and the corresponding pattern of cellular localization of the mutants.

3.3.1 Fluorescence recovery after photobleaching (FRAP)

In order to further examine the effect of NLS2 mutation on localization of PNKP, we devised an experiment where we employed the FRAP technique with HeLa cells transfected with GFP-tagged NLS2-mutants. In FRAP, a small region within a larger area of a cell is defined and briefly illuminated with high laser intensity. For instance, a region maybe defined within a cell nucleus or even entire nucleus. Photobleaching renders the majority of the GFP-tagged proteins within the defined region to irreversibly lose their fluorescent properties. In a setup where all GFP-tagged proteins are mobile, proteins from outside will diffuse into the bleached region increasing the degree of fluorescent signal until the signal inside the is equal to the signal outside the bleached region. In contrast, immobile proteins are will not diffuse into the strip resulting in an incomplete recovery of the fluorescent signal inside the bleached region relative to the remainder of the area[136].

We photo-bleached the entire nucleus of the cells and recorded the recovery of GFP back into the nucleus at different time intervals (Fig. 19). The results were intriguing since the nuclear recovery of GFP was indeed very slow but the experiment posed its own limitation as the experiment lacked a proper control. Using wt-PNKP-GFP transfected cells as a positive control had its own problem since the cells showed a predominant nuclear and no cytoplasmic localization, and bleaching their nuclei showed no discernible GFP recovery. As a second control we used a vector carrying GFP-only. GFP showed a uniformly distributed localization pattern as expected when transfected into HeLa cells. Although it showed quick recovery post bleaching, it could not be considered as an ideal control because GFP is a protein with a molecular weight of ~27 kDa and proteins which

have a molecular weight of less than 40 kDa reportedly enter the nucleus through passive diffusion through nuclear pores[99, 100]. With these concerns, the experiments did not provide definitive answers. However, the experiment suggested that the NLS2 mutations might have disrupted the PNKP import in the nucleus as indicated by an extremely slow recovery of GFP in the nucleus.

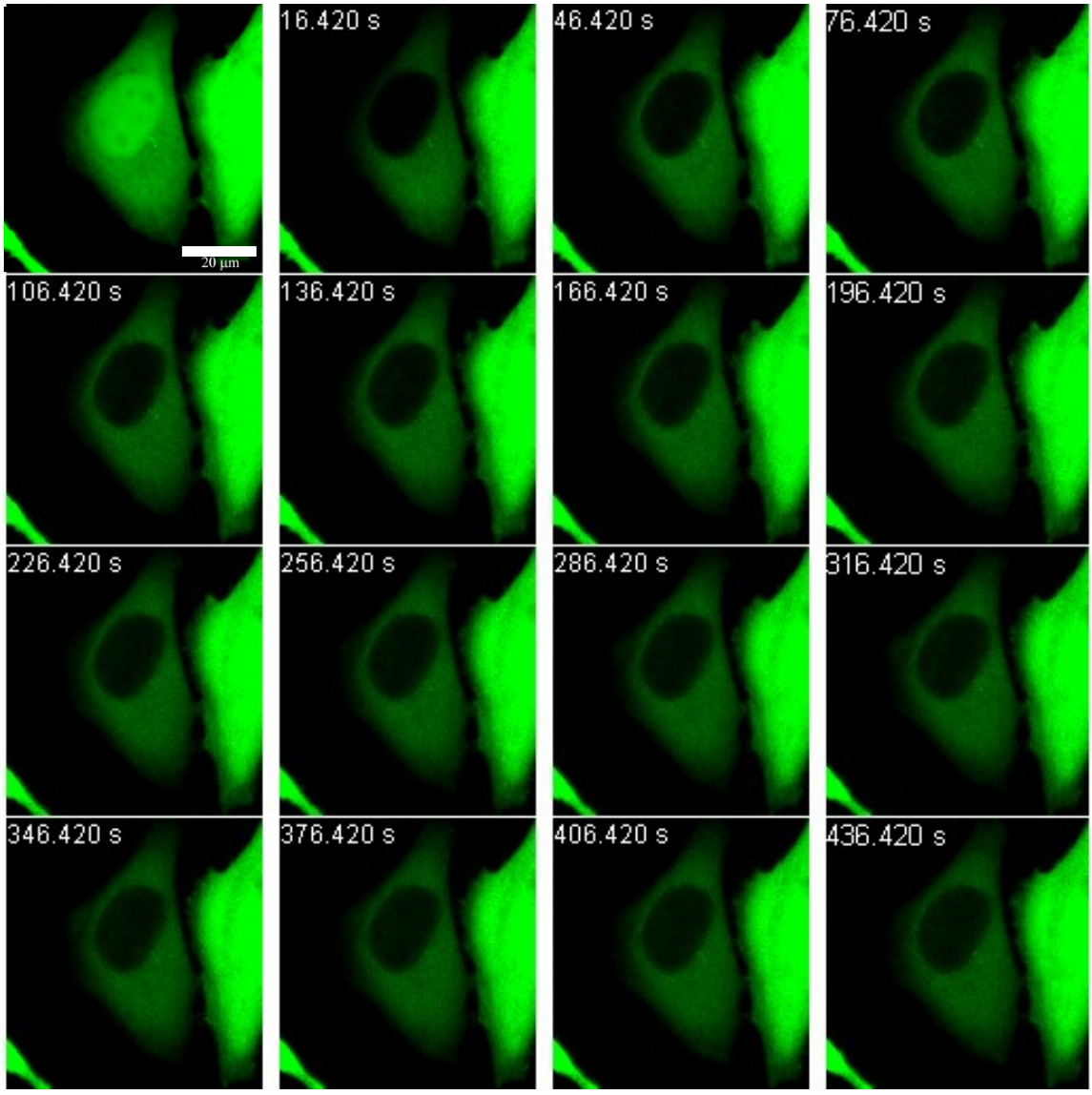


Figure 18. FRAP experiment results conducted with the NLS2-mutant. Cells pre-plated on 35-mm glass-bottom plates were transfected with mutant-NLS2+GFP. 24 h post transfection, nuclei of the cells were bleached using multi-photon laser and the nuclear GFP recovery was recorded every 10, 15 and 20 seconds (The figure shows recovery interval of 20 seconds). As shown, even after >7 min, there was no visibly discernible recovery of GFP in the nucleus. Scale bar: 20 µm.

3.4 pShooter PNKP Experiments

In order to effectively redirect PNKP into particular organelles, we resorted to pShooter vectors that bear sequences known to bind to the receptors responsible for targeting proteins into those organelles. PNKP, both wild type and NLS2-mutant, with a GFP tag were cloned into pShooter myc/mito and pShooter myc/nuc vectors. Live confocal imaging was conducted on cells transfected with these constructs. As expected, pShooter-myc/nuc-PNKP-GFP showed a very prominent localization in the nucleus (Fig. 20, Panel 2). pShooter-myc/mito-NLS2-GFP showed a diverse population of cells in terms of GFP localization. Despite being able to localize proteins to the mitochondria, as seen with GFP alone, when PNKP was expressed in the pShooter-myc/mito vector, GFP was detected in the mitochondria as well as the cytoplasm and nucleus. Hence while localization was enhanced, it was not complete (Fig. 20, Panel 1).

In order to segregate cells expressing PNKP only in the mitochondria, we employed fluorescence activated cell sorting (FACS) to isolate GFP expressing cells (cells transfected with pShooter-myc/mito-NLS2-GFP) as single cells into 96-well plates. Two sets of 96-well plates were prepared, one receiving GFP overexpressing cells and the other receiving GFP non-overexpressing cells. The cells were incubated until colonies developed in each well. Six random colonies from each set were trypsinized and grown on glass-bottom 35-mm plates. These plates (a total of twelve) were used for live confocal imaging (Fig. 21).

Fig. 19

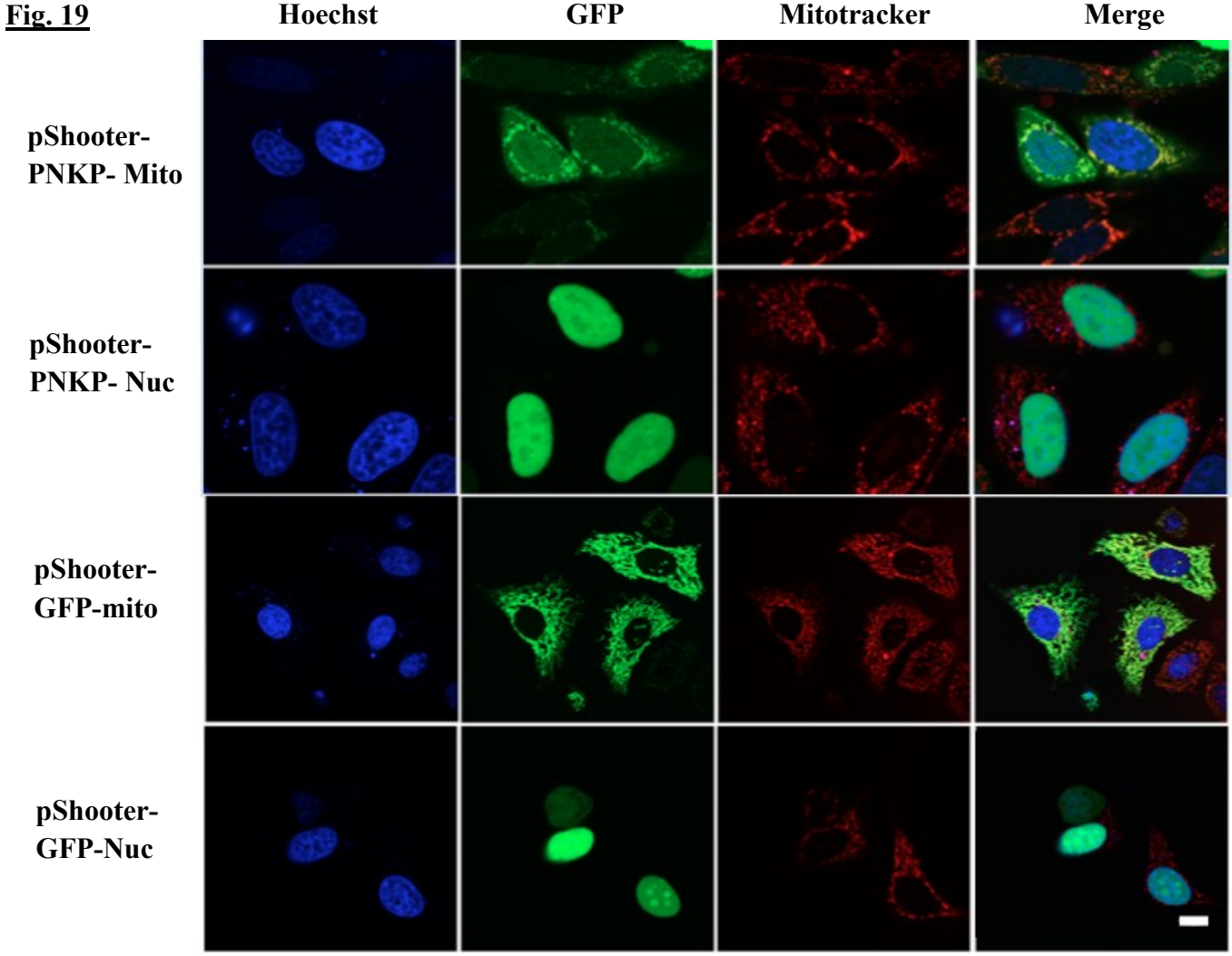

Hoechst
GFP
Mitotracker
Merge
pShooter-PNKP- Mito
pShooter-PNKP- Nuc
pShooter-GFP-mito
pShooter-GFP-Nuc

Figure 19. Subcellular localization of GFP-tagged PNKP on pShooter vectors: HeLa cells were transfected with pShooter - PNKP – Nuc (pShooter-myc/nuc-PNKP-GFP) and pShooter – PNKP – Mito (pShooter-myc/mito-NLS2-GFP). Cells were visualized live by confocal microscopy. In parallel, cells were transfected with the appropriate pShooter vectors encoding GFP, which served as positive controls (Panels 3 and 4). Co-localization of pShooter – PNKP – Nuc and pShooter – PNKP – Mito with the specific organelle markers, Hoechst and MitoTracker Orange, was determined. As seen in Panel 1, pShooter-Mito vector was able to direct PNKP (NLS2-version) into the mitochondria, evident by GFP co-localization with Mitotracker Orange. Similarly, Panel 2 show the pShooter-Nuc vector successfully directing PNKP into the nucleus, evident by GFP co-localization with Hoechst dye. Scale bar: 10 μm.

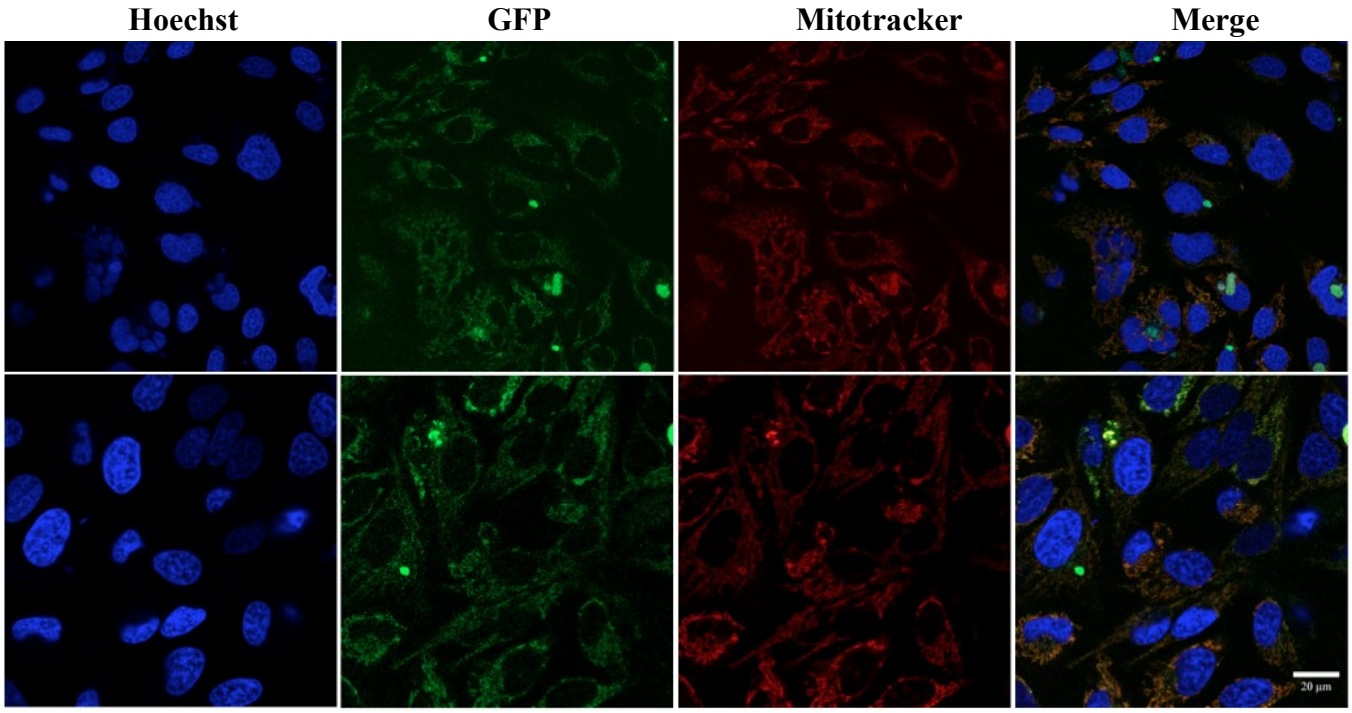


Figure 20. Enriching population of cells with mitochondrial PNKP (mtPNKP): HeLa cells were transfected with pShooter – PNKP – Mito (pShooter-myc/mito-NLS2-GFP) and incubated overnight. On the following day, cells were trypsinized and resuspended in Basic sorting buffer. In a 96-well plate, single cells were isolated in each well and incubated until colonies developed. The colonies were subsequently trypsinized and plated on 35-mm glass-bottom plates. Cells were visualized live by confocal microscopy. As shown, a population of cells expressing PNKP exclusively in the mitochondria were obtained (GFP lane). Mitotracker Orange and Hoechst dye were used to stain the mitochondria and nucleus, respectively. Scale bar: 20 µm.

Chapter III

Discussion

3.1 Determining the localization pattern of PNKP under different fixatives

Our first objective in this study was to determine the localization pattern of PNKP under two different fixation conditions using two different types of antibodies. The purpose of fixation is to immobilize antigens within cells while retaining the cellular and subcellular structure, which allows the antibodies to gain access into cellular organelles and bind to the epitopes on the protein under study. However, fixation can also lead to epitope masking/damage as in the case of formaldehyde-based fixatives, and extraction of antigens as in the case of methanol/acetone fixation approaches[137, 138]. Therefore in either of the approach, there is a probability of immunostaining being compromised. Therefore, to examine the sub-cellular localization of PNKP, we decided to use two different fixatives most commonly used in research laboratories.

One of most common fixatives used in laboratories is PFA, a crosslinking agent that chemically crosslinks proteins and molecules to the cytoskeleton preserving overall architecture of the cell. Whereas PFA can promptly fix the cells, a secondary detergent-based permeabilization step (Triton-X100 or NP40) is required for enabling antibodies to gain access into sub-cellular structures. We adopted 4% PFA and 0.2% Triton-X100 for fixing and permeabilizing our A549-monolayer cells. We used mouse-monoclonal (H101) antibodies prepared in the lab and commercially available rabbit-polyclonal antibodies (Sigma) against PNKP to track the sub-cellular localization of PNKP in fixed cells. We observed a predominant PNKP signal in the nucleus with minimal signal in the mitochondria, and almost identical patterns with both the antibodies that were used. Interestingly, a pattern of nucleolar localization was also evident in the images with the anti-PNKP polyclonal but further experiments will be necessary for verification. While

the predominant nuclear localization of PNKP was not surprising, only a relatively weak mitochondrial signal was present. We reasoned that to be because of the lower accessibility of antibodies through the mitochondrial membrane that can be better addressed by a brief exposure of the cell to an additional organic agent such as methanol or acetone.

We therefore also adopted an organic solvent-based approach using methanol:acetone (1:1) with the same pair of antibodies. This technique eliminates an extra step of permeabilization since both acetone and methanol are able to fix as well as permeabilize the cells. In contrast to the PFA results, we observed a predominant mitochondrial signal of PNKP with a relatively weaker signal in the nucleus.

A comparative study conducted by Hoetelmans et al in MCF7 cells using different fixatives suggested that methanol and acetone could compromise the nuclear and cytoplasmic structure of the cells[139]. Moreover, a combined use of PFA/methanol was observed to retain the nuclear integrity better than when methanol was used alone. Especially in the acetone fixed cells, the nuclear envelope integrity was observed to be compromised leading to loss of nuclear contents[139].

We concluded by re-emphasizing that PNKP localizes in both the nucleus and the mitochondria although we need to perfect our staining technique by adopting a mixed approach utilizing both formaldehyde and organic solvent-based fixatives.

3.2 Determining the MTS of PNKP

PNKP has long been established as a bi-functional end-processing enzyme that operates in multiple DNA repair pathways such as BER, SSBR and NHEJ. PNKP catalyzes the dephosphorylation of 3′-P termini and the phosphorylation of 5′-OH termini to yield DNA polymerase compatible 5′-P and 3′-OH ends, playing a key step in those repair processes. Mitochondria are the source of reactive oxygen species (ROS), which are known to inflict SSBs that bear unligatable termini such as 3′-P and 5′-OH. In this light, it is logical to assume that mtDNA sustains even more ROS-induced unligatable termini compared to nuclear DNA. Moreover, earlier studies report various SSBR proteins such as UNG, NEIL1, NEIL2, TDP1, PARP, Aprataxin, APE1 and DNA ligase III to be present in the mitochondria[84, 140-144] indicating the potential of PNKP being a part of specialized mtDNA repair machinery. We have conducted a study providing evidence that indicate that a full-length PNKP localizes in the mitochondria (Appendix B).

The mitochondrial proteins that are encoded by nuclear DNA are known to often bear a canonical N-terminus MTS that is recognized by mitochondrial receptor proteins. However, 50% of the mitochondrially-targeted proteins do not use this classical import mechanism. Many nuclear proteins such as TDP1, NFkB, p53, BRCA1, PARP1 and AP1 that have been discovered in mitochondria lack the canonical MTS[142, 143, 145-148]. APE1 was discovered to possess an MTS on its C-terminus but also required proteolytic cleavage of an N-terminus segment containing its NLS for mitochondrial homing[149, 150]. Also, yeast DNA helicase Hmip and human DNA2 proteins are known to possess a C-terminal MTS[151, 152]. On the other hand, Aprataxin was discovered to have an N-

terminal MTS[144]. These reports suggest that there is a lot of diversity in MTSs of dually targeted proteins and there must be mitochondrial transport mechanisms other than the classical pathway yet to be discovered.

Here we show that a non-canonical MTS resides close to the C-terminus of PNKP that enables trafficking of PNKP into the mitochondria. We analyzed PNKP protein sequence with Mitoprot II and Predotar MTS predicting programs, which indicated a sequence, ^{432}ARYVQCARAA^{441}, close to the C-terminus as the putative MTS of PNKP. To verify the prediction, we generated truncations of PNKP (GFP-tagged) so that the putative MTS was placed at the N-terminus of the resulting protein. The sequence was able to target the protein into the mitochondria. When the sequence was either deleted or mutated, the protein failed to localize in the mitochondria. We also introduced MTS mutations on a full-length PNKP and analyzed it effects on mtDNA repair employing XL-qPCR technique. We determined this sequence to be the functional MTS of PNKP and required for PNKP repair activity in the mitochondria.

3.3 Determining the NLS of PNKP

Our next objective was to determine the NLS of PNKP. A previous study suggested a Lys-Arg motif, ^{301}RKKK^{304} (NLS1) as the putative NLS. However, analysis of PNKP protein sequence with NucPred and cNLS mapper programs indicated another sequence ^{137}KKR^{139} (NLS2) located within the linker region of the PNKP, as a better candidate to be an NLS. Here we show that ^{301}RKKK^{304} region predicted earlier as a putative NLS does not contribute to nuclear import of PNKP. However, the predicted

^{137}KKR^{139} region, although not able to completely sequester PNKP in the nucleus, is critical for the nuclear import of PNKP. Introducing mutations (random mutations or alanine-substitutions) at the ^{137}KKR^{139} region substantially disrupts the nuclear targeting of the protein as indicated by the live immunofluorescence results (Fig. 16). Surmising that PNKP might bear two independent NLSs, as was observed with SRY and SOX9 proteins previously[153], we also produced a PNKP construct combining both NLS1 and NLS2 mutants. The results of live immunofluorescence were identical to the results obtained with NLS2-mutants (data not shown). This result emphasized that NLS1-residues play absolutely no role in nuclear import of PNKP. Furthermore, FRAP experiments conducted on NLS2 mutants supported our finding that NLS2-residues could be the potential NLS of PNKP.

A bi-partite NLS was first described for Nucleoplasmin[154], and has been discovered in several proteins since then. In general, a bi-partite NLS comprises of two basic amino acids, a 10-15 amino acid spacer followed by three basic amino acids. Considering this, additional mutations were introduced on lysine and arginine residues upstream and downstream of the NLS2. The additional mutations did not further affect the localization pattern of PNKP eliminating the possibility of a bi-partite NLS. Furthermore, cNLS mapper did not indicate the presence of a bi-partite NLS. However, it did indicate a larger segment spanning the NLS2 region as a putative NLS, which is a considerably longer stretch of amino acids compared to conventional NLSs. But the length is not unusual considering that a much longer stretch of amino acids has qualified as an NLS, such as the 38-amino acid long NLS of Plk1 protein[155]. Deletion of ^{135}LPKKRMRKSNP^{147} residues did result in a complete sequestration of PNKP in the

cytoplasm in almost 90% of the cells. However, it raised concerns about the potential structural alterations that the protein can sustain as a result, as mentioned in Section 3.3. A mutant coding for a smaller amino acid deletion was also constructed (^{135}LPKKRMRK^{142}), but this failed to express the protein when transfected into cells (Fig. 18).

We therefore concluded that the amino acids located within the 11 amino acid residues (^{35}LPKKRMRKSNP^{147}) that we investigated as critical for import of PNKP into the nucleus.

3.4 Overall Impact on the objective

Previous studies suggest that depletion of PNKP increases the sensitivity of cells against ionizing radiation and H_2O_2 amongst other chemotherapeutic agents[156]. We now have successfully shown full-length PNKP localizes in the mitochondria contributing to the maintenance of mtDNA integrity. Our objective was to examine the relative importance of nuclear versus mitochondrial PNKP in response to H_2O_2 and other ROS inducing agents. For this purpose, we set out to generate two cell lines that expressed PNKP either only in the nucleus or mitochondria. In order to obtain those cell lines, we decided to determine the MTS and NLS of PNKP first. Once the signals were determined, the next step was to reconstitute NLS-negative and MTS-negative versions of PNKP in PNKP-knockdown cells. The following step was to expose the cells to various genotoxic and chemotherapeutic agents and record the cell survival.

Although we firmly established the MTS and located the NLS, we were not successful to produce a cell line with nuclei completely devoid of PNKP. Therefore as an

alternative approach, we adopted pShooter-vectors as a cellular system to selectively target PNKP to nucleus or mitochondria. The pShooter-vector system has been previously used in studies such as determining the role of mitochondrial localization of mutant SOD1 (superoxide dismutase 1) in cell death, analyzing the effect of subcellular localization of PKCδ (protein kinase δ) on its apoptotic functions, examining the impact of intracellular localization of tissue transglutaminase on cell death and studying the interaction of Chromogranin with $InsP_3R$ (Inositol 1, 4, 5-triphosphate Receptor[157-160]. We therefore cloned PNKP into the nuclear targeting (pShooter-myc-nuc) and mitochondrial targeting (pShooter-myc-mito) pShooter vectors. PNKP targeting into the nucleus was easily achieved. In order to facilitate the mitochondrial targeting of PNKP to mitochondria, we introduced NLS2-mutations (^{37}KKR^{139}) on the pShooter-myc-mito-PNKP construct to achieve a population of cells that contained PNKP localized only in the mitochondria. The revised workflow of the project after the introduction of pShooter vectors has been explained in Chart1.

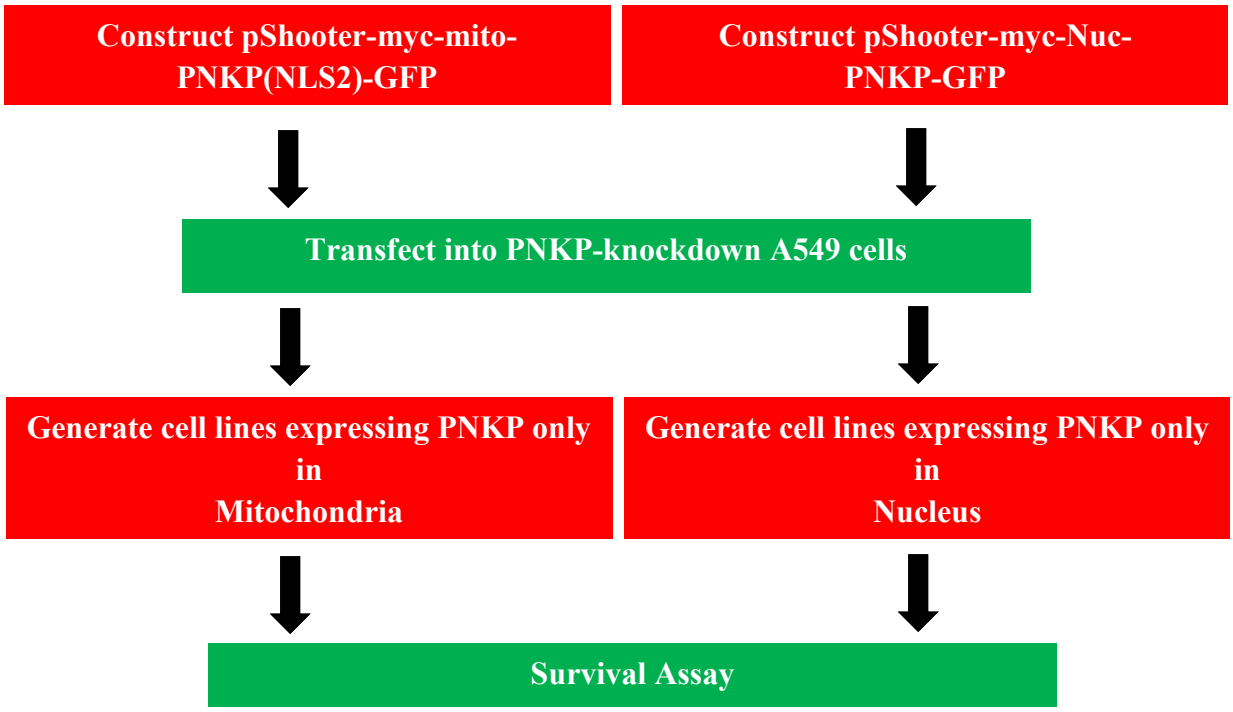


Figure 21. Revised thesis focus.

Part II

Chapter I

Introduction

1.1 CRISPR interference in prokaryotes

Bacteria are the most abundant organisms found in many natural habitats where they are under intermittent threats posed by predatory viruses and bacteriophages. As a result, they have evolved various defense mechanisms to resist such invaders. One such defense mechanism is a form of adaptive immune defense termed CRISPR (Clusters of regularly inter-spaced short palindromic repeat) interference. CRISPR interference is a reprogrammable defense system that involves a sequence specific, short RNA-mediated targeting and degradation of invading nucleic acids that mostly originate from phages or foreign plasmids[15].

1.1.1 CRISPR Locus

A CRISPR locus within a host is a structure primarily composed of partially palindromic identical DNA repeats that occur at regular intervals in the bacterial genome. Each of these recurring 'repeats' varies in length between 23 to 55-bp and are separated from each other by non-repetitive 'spacers' each ranging from 21 to 72-bp in length[15, 161, 162]. (The significance of repeats and spacers is explained below in Section 1.1.2). Therefore, a typical CRISPR array is an assembly of a few to hundred repeat-spacer units thus exceeding several thousand base pairs. Within the locus, the array is flanked on one side by an AT-rich 'leader' sequence comprising a promoter sequence responsible for CRISPR RNA transcription and the other side by a set of CRISPR-associated (*cas)* genes (Fig. 1)[163]. *Cas* genes encode protein complexes that work in consort with the small RNAs as a part of the CRISPR machinery. Together the products of the CRISPR locus manufacture a molecular apparatus that provides the host with a form of adaptive immune

system that involves recognizing and degrading the genetic material deposited by intruding viruses and plasmids in the host cell[164].

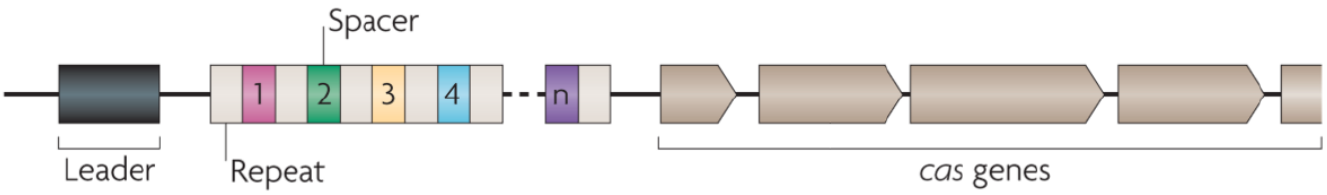


Figure 1. CRISPR Locus: CRISPR locus consists of an AT-rich leader sequence followed by multiple repeat and spacer sequence. Repeats are the recurring palindromic sequences that are separated by spacer sequences. Spacers are non-repetitive sequences (colored), which share homology with the nucleic acid sequences of various foreign plasmids and bacteriophages (explained in Section 1.1.2). *Cas*-genes code for necessary proteins that complete CRISPR-machinery[15]. Reproduced by permission from Nature Publishing Group: Nature Reviews Genetics. Marraffini L.A. and Sontheimer E.J, Copyright © **2**010.

1.1.2 Type II CRISPR system

Depending upon the genetic content, structural and functional differences, CRISPR systems are broadly categorized as Type I, Type II and Type III[15]. For the purpose of this thesis, we will only focus on Type II CRISPR-Cas9 system of *Streptococcus pyogenes*.

In the Type II CRISPR-Cas9 system, Cas9 endonuclease and an RNA complex comprising a hybrid between crRNA (CRISPR RNA) and a tracrRNA (transacting RNA) in concert execute the process of accomplishing immunity that involves recognizing and degrading of foreign genetic material[165-168]. In a bacterial life cycle, the Type II CRISPR-Cas9 interference is achieved in following steps (Fig. 2):

1. *Spacer Acquisition or Immunization phase*: In this step, new nucleic acid fragments are acquired from the genome deposited by the invading bacteriophages as a consequence of infection and stored as molecular signatures by integrating them into the CRISPR locus. These captured fragments are called 'spacers' and each one is flanked on either side by identical repeat regions. Thus the CRISPR spacers with variable sequences are the outcomes of uptake and amassing of various genetic materials encountered over time by the host. Spacers therefore can be homologous to a number of different foreign genetic elements, mostly originating from bacteriophages[10, 169].

2. *Immunity phase*: Upon infection by a previously encountered virus, transcription of the CRISPR locus is triggered in the host. Next, tracrRNAs hybridize with the repeat regions on pre-crRNA (combination of repeats and spacers). The hybrid RNA thus formed is subsequently processed by endogenous RNAse III to generate mature CRISPR RNAs or crRNA. Next, crRNA associates with Cas9, the effector nuclease, to form an active DNA endonuclease, often called a 'dualRNA-Cas9 complex', which locates the target foreign nucleic acid. The targeting is achieved when the spacer-transcript region within the dualRNA-Cas9 complex base pairs with the complementary region, known as a 'protospacer', present on the target. In the next step, Cas9 cleaves both strands thus preventing virus assembly thereby accomplishing defense[10, 169]. The cleavage of foreign nucleic acid is mediated by two catalytic endonuclease domains of Cas9, the RuvC-like domain that cleaves the non-complementary strand and the HNH domain that cleaves the complementary strand Fig. 3[167].

The specificity of target recognition is determined by two associations: first association established due to complementarity between the RNA chimera and target DNA and second resulting from the binding of Cas9 with a short motif of sequence present on the target DNA termed as PAM (Protospacer Adjacent Motif)[170, 171]. PAM sequence is known to vary between 2 to 5-bp depending upon the type of CRISPR system and the organism[170]. Interestingly, the importance of PAM is underscored by the fact that the two-endonuclease domains of Cas9 cleave on the segment of protospacer region

2 to 3-nt upstream of the PAM[11]. One of the examples of a PAM in a Type II system is a tri-nucleotide sequence 5'-NGG-3', where N can be any nucleotide[10].

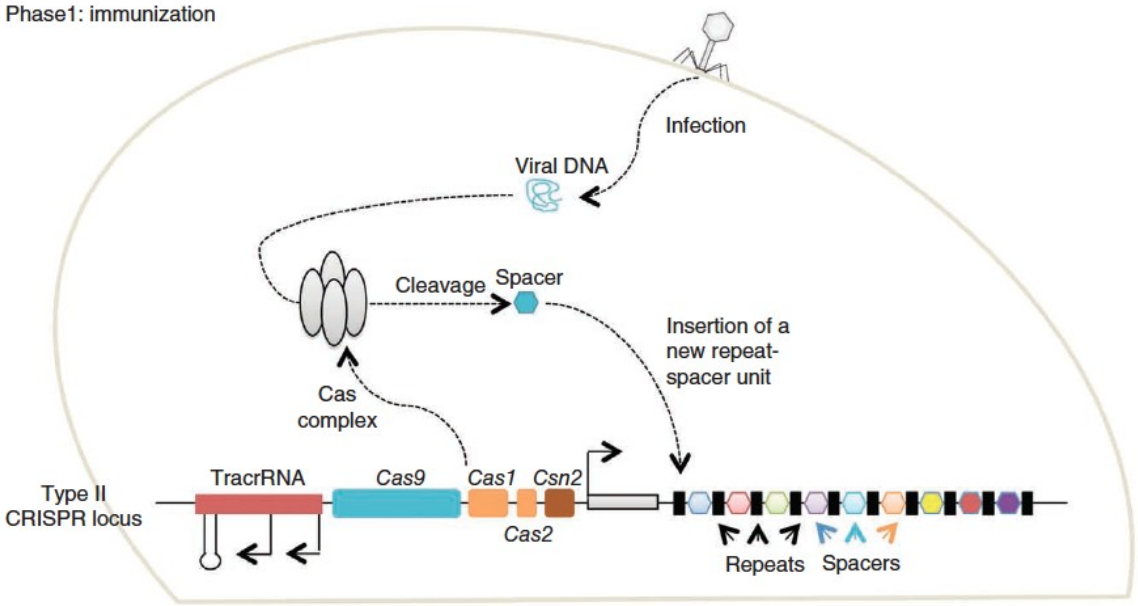


Figure 2a. Bacterial Type II CRISPR-Cas9 system – Spacer Acquisition/Immunization Phase: In this phase, nucleic acid of invading bacteriophage is processed and stored as spacers (colored) in the CRISPR locus. Spacers thus serve as molecular signatures of the previous infections[10]. Reproduced by permission from Nature Publishing Group: Nature Methods. Mali P., Esvelt K.M. and Church G.M, Copyright © **2**013.

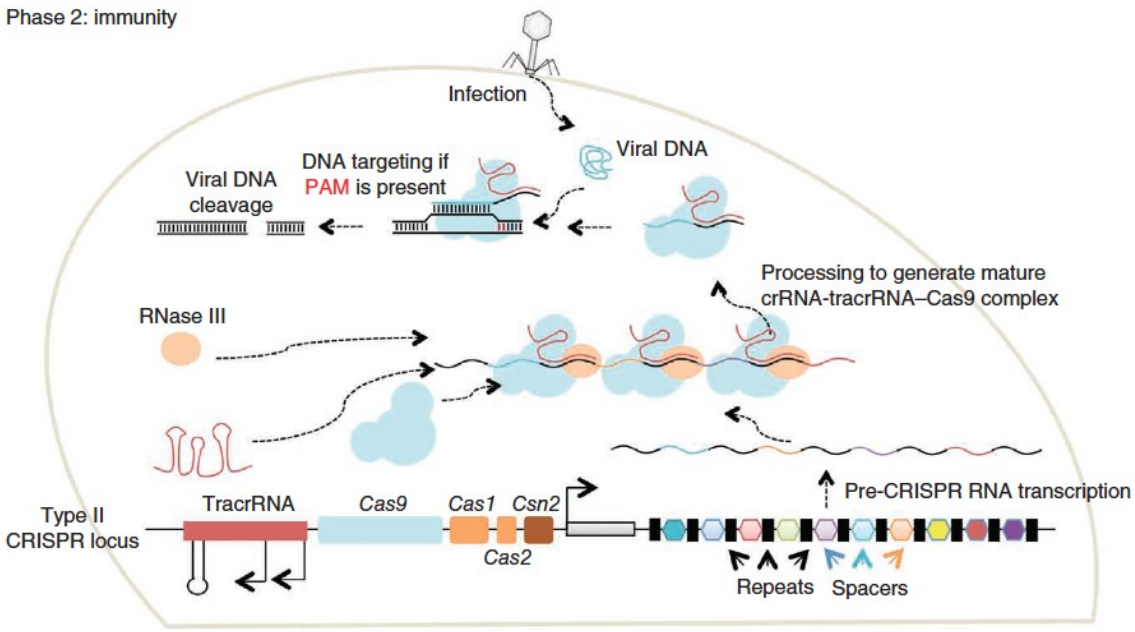


Figure 2b. Bacterial Type II CRISPR-Cas9 system – Immunity phase: In this phase, the host uses the stored molecular signature or spacer as a tool of defense against invading bacteriophages. Upon viral infection, the CRISPR locus is transcribed to produce CRISPR RNAs and Cas9-nuclease, which in concert (as dualRNA-Cas9 complex) locate the region within the foreign nucleic acid (protospacer) complementary to the spacer and introduces strand breaks to accomplish defense[10]. Reproduced by permission from Nature Publishing Group: Nature Methods. Mali P., Esvelt K.M. and Church G.M, Copyright © **2**013.

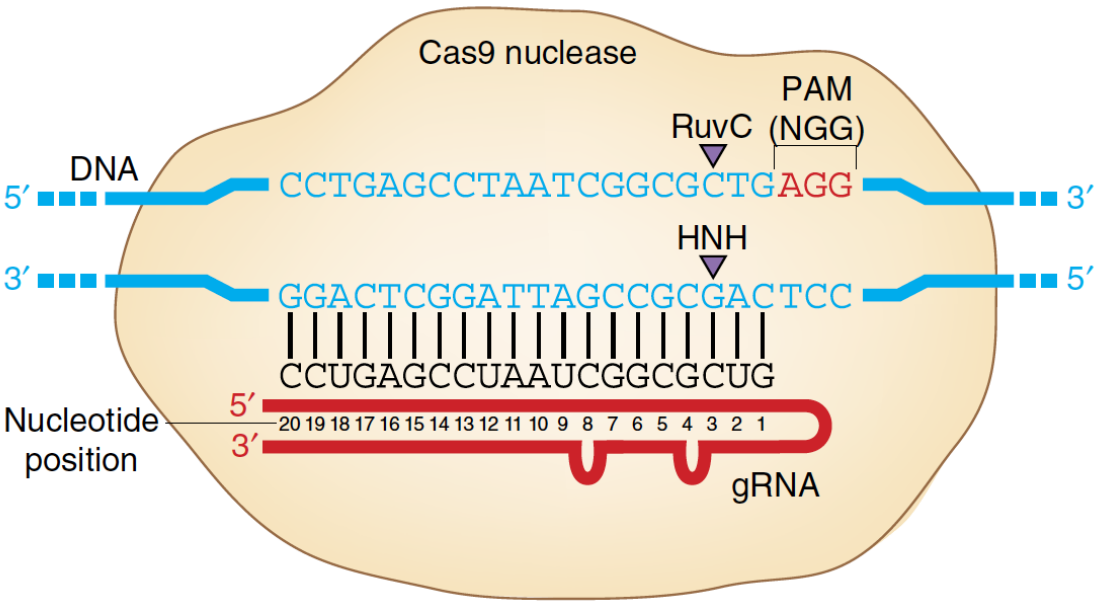


Figure 3. Cas9 nuclease domains and PAM: Once the dualRNA-Cas9 complex locates the protospacer region within the target nucleic acid, the RuvC-like and HNH domains of Cas9 introduce cleavage 2 to 3-bp upstream of the PAM sequence. PAM is a tri-nucleotide 5'-NGG-3' sequence on the target nucleic acid, where N can be any nucleotide[11]. Adapted by permission from Nature Publishing Group: Nature Biotechnology. Sander J.D. and Joung J.K., Copyright © **2**014.

1.2 Type II CRISPR-Cas9 system in genome engineering

Recently, there have been some remarkable breakthroughs in using Type II CRISPR-Cas9 system as a gene-editing tool. Various research groups have manipulated the system to achieve site-specific modification of eukaryotic and microbial genomes. We focus here only on the genome editing of mammalian cells using the CRISPR-Cas9 system emphasizing its utility in knocking out a gene.

Type II CRISPR-Cas9 machinery comprises of two primary elements. In a human setting, it includes a human codon optimized version of Cas9-nuclease and a single guide RNA (sgRNA), which is a minimal chimera structure of tracrRNA-crRNA[172, 173]. An sgRNA harbors a complementary region specific to the gene of interest, a Cas9-handle required for Cas9 binding and a transcription termination hairpin structure Fig. 6[14]. When these elements are introduced in human cells, sgRNA acts as a guide that directs Cas9 to the specific site on the appropriate chromosome where Cas9 introduces a site-directed DSB[11]. A DSB triggers the activation of the DSB repair machinery of the host cell. Both HR and NHEJ DSB repair pathways have been exploited for the purpose of genome editing:

- In the presence of a homology repair template that is supplied exogenously along with the CRISPR machinery, DSB repair introduces sequences carried by the template into the site of breaks through HR-mediated recombination events[174].
- In the absence of an exogenous repair template, an error-prone NHEJ pathway is triggered. NHEJ can introduce insertions and deletions while attempting to repair the Cas9-created DSBs thereby introducing frame-shifts or pre-mature stop codons on the gene. This approach is employed in knocking out a gene[174, 175].

1.3 Thesis Focus – Type II CRISPR-Cas9 knockout of PNKP

In this part, we will focus on describing the approaches we undertook in order to knockout PNKP in A549 lung adenocarcinoma cells.

In one of our approaches, we developed stable A549 cell lines expressing Cas9 protein (Fig. 5a). We also generated cell lines expressing the mutant versions of Cas9: Cas9-D10A-nickase or Cas9n with RuvC-like domain inactivated, which is only capable of forming single stranded DNA nicks[176] and Cas9-D10A/H840A-catalytically dead or dCas9 with both RuvC-like and HNH domains inactivated, which is incapable of introducing any form of breaks on DNA[177]. Use of a wild-type Cas9 has been often linked with off target cleavage and disruption of unintended sites on chromosomes[178]. Therefore we decided to create and use the mutant versions of Cas9 because they have reportedly been able to demonstrate more specificity when compared to wild-type Cas9. Subsequently, these cell lines were transfected with plasmids bearing the sgRNA machinery containing a PNKP specific complementary sequence (Fig. 5b).

In a slightly different approach, we adopted a Lentivirus-based method where we used a bicistronic vector bearing both Cas9 gene and the sgRNA machinery (Fig. 6) in A549 cells.

In either of these approaches, PNKP knockout should occur when the sgRNA-Cas9 chimera binds to the target site of the PNKP gene on chromosome 19q13.3 and creates a DSB proximal to the PAM. In turn, the erroneous NHEJ pathway is triggered that results in introduction of indels creating frameshifts or introduction of premature stop codons[176].

Chapter II

Materials and Methods

2.1 Cells

A549 (human lung adenocarcinoma cells) and HeLa (human cervical adenocarcinoma cells) were obtained from the American Type Culture Collection (Manassas, VA). The cells were cultured in a 1:1 mixture of Dulbecco's modified Eagle's medium/nutrient mixture and F12 (DMEM/F12) supplemented with 10% fetal bovine serum (FBS), penicillin (50 U/ml), streptomycin (50 μg/ml), L-glutamine (2 mM), non-essential amino acids (0.1 mM) and sodium pyruvate (1 mM), and maintained at 37^{o}C under 5% CO_2 in a humidified incubator. All culture supplements were purchased from Invitrogen, USA.

2.2 Transient transfections

Approximately 5 X 10^5 HeLa cells were plated in a 35-mm dish (Sarstedt, Germany) in DMEM/F12 media without antibiotics and allowed 24 h to adhere in a humidified incubator at 37°C and 5% CO_2. 500 ng of hCas9 plasmid DNA was added to 250 μL total reaction volume in Opti-MEM (Invitrogen, USA). Simultaneously, a 1:50 dilution of Lipofectamine 2000 Transfection Reagent (Invitrogen) in Opti-MEM was allowed to incubate at room temperature for 5 min, to provide a final volume of 5 μL of transfection reagent per well. The two-transfection solutions were then combined and held at room temperature for 20 min. The media was then removed from the pre-plated cells and 500 μL of the transfection mixture was added per dish and incubated at 37°C and 5% CO_2 for 24 h before subsequent steps.

2.3 G-Block approach

2.3.1 gRNA sequence identification

All 23-bp genomic sites of the form 5'-N20NGG-3' (where NGG is a PAM) on the exon of PNKP (both template and non-template strand) were identified using Serial Cloner Software. Using NCBI blast, sequences unique to PNKP gene of the form 5'-NNNNN NNNNN NNNNN NNNNN NGG…3' were identified and listed.

2.3.2 G-block synthesis

19 bp from the selected target sequence, excluding the PAM sequence, were incorporated into the following DNA sequence as demonstrated:

(For example: 5'-**CACTGGCCCCAGTGACTGGNGG**-3' sequence targeting exon 5 of PNKP was incorporated into the DNA fragment as indicated below.)

TGTACAAAAAAGCAGGCTTTAAAGGAACCAATTCAGTCGACTGGATCCG
GTACCAAGGTCGGGCAGGAAGAGGGCCTATTTCCCATGATTCCTTCATA
TTTGCATATACGATACAAGGCTGTTAGAGAGATAATTAGAATTAATTTGA
CTGTAAACACAAAGATATTAGTACAAAATACGTGACGTAGAAAGTAATAA
TTTCTTGGGTAGTTTGCAGTTTTAAAATTATGTTTTAAAATGGACTATCAT
ATGCTTACCGTAACTTGAAAGTATTTCGATTTCTTGGCTTTATATATCTTG
TGGAAAGGACGAAACACCCACTGGCCCCAGTGACTGGTTTTAGAGCTAG
AAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGC
ACCGAGTCGGTGCTTTTTTT

The 455-nt fragment bears components necessary for gRNA expression, U6 promoter, target sequence, guide RNA scaffold and termination signal. The 455-nt DNA fragment was purchased as a double stranded blunt ended G-Block from IDT, Coralville, IA. G-Blocks (455-nt fragment) that were used in the experiments are listed in Appendix I.

Figure 4. Vectors used for G-block approach: a) Plasmid hCas9 carries a human codon optimized Cas9 cassette bearing a nuclear localization signal (NLS) on its C-terminus. b) Plasmid used for sgRNA expression harbors all necessary components for expressing an sgRNA. Refer Fig 6 for the structure of sgRNA chimera.

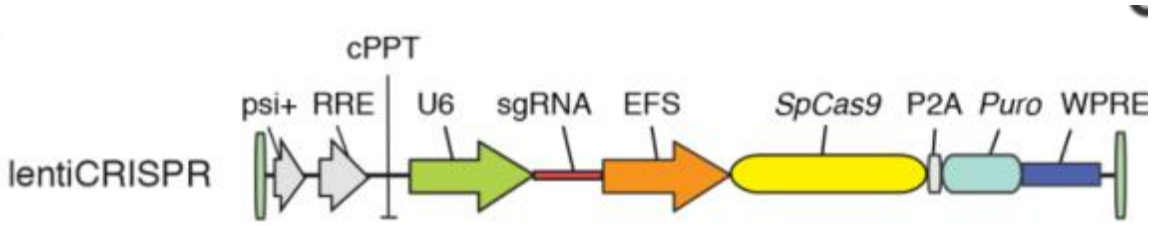


Figure 5. Vector used for Lentiviral approach: The pLentiCRISPR v2 is a one-vector system that contains both the Cas9 and the chimeric sgRNA expression cassettes. Refer to Figure 6 for the structure of the sgRNA chimera[2]. Reproduced by permission from American Association for the Advancement of Science: Science. Shalem O. et al., Copyright © 2014.

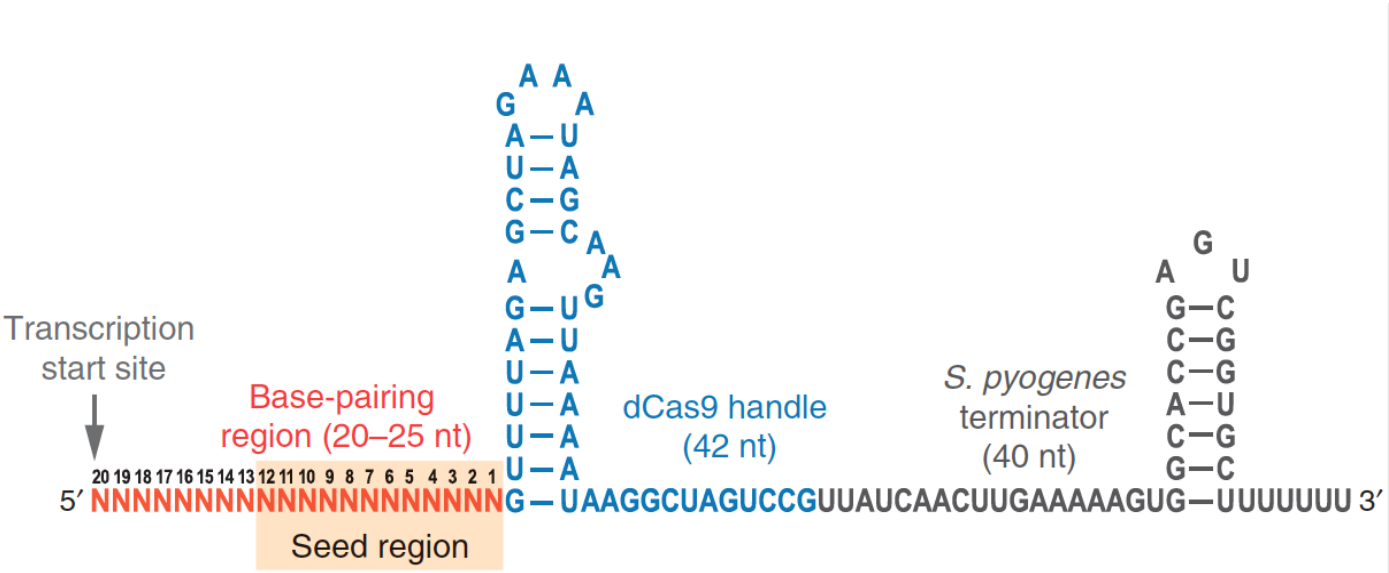


Figure 6. sgRNA chimera: The sgRNA sequence optimized for gene editing comprises a 20-25-nt complementary base pairing region for specific binding to the target DNA (*orange*), a 42-nt long Cas9 handle required for binding with Cas9 protein *(blue)* and a 40-nt long transcription terminator hairpin *(gray)* derived from *Streptococcus pyogenes*[14]. Reproduced by permission from Nature Publishing Group: Nature Protocols. Larson M.H. et al., Copyright © **2**013.

2.3.3 G-Block cloning

All the G-Blocks (G-blocks 3, 7, F, RC and Scramble) were cloned into an empty pCR-Blunt II-TOPO vector (Invitrogen) following the manufacturer's protocol.

2.4 Stable transfections

Approximately 200,000 cells were plated in DMEM/F12 media without antibiotics and allowed to adhere overnight in a 35-mm dish (Sarstedt, Germany) in a humidified incubator at 37°C and 5% CO_2. The transfection mixture was prepared from two separate solutions, the first containing 1 μg of plasmid DNA dissolved in 250 μL total of Opti-MEM (Invitrogen), and the second 6 μL of Lipofectamine 2000 (Invitrogen) in 250 μL total Opti-MEM. The two solutions were held at room temperature for 5 min, combined and then held at room temperature for another 20 min. The media from the pre-plated cells was replaced and the transfection mixture was added. The cells were allowed to incubate for 24 h at 37°C and 5% CO_2. On the following day, cells were trypsinized and replated into 100-mm plates in DMEM/F12 containing 500 μg/mL G418 (Calbiochem, USA) or 350 μg/mL Zeocin (Invivogen, USA). The selective media was replaced every 3 days to allow for growth of individual colonies. When single-clone colonies formed after 10-18 days, the colonies were picked and expanded.

2.5 RNA isolation and cDNA synthesis

Cells were seeded in a 60-mm tissue culture plate (Sarstedt, Germany) until 80% confluent. Thereafter, the total RNA was isolated using TRIzol reagent, according to the manufacturer's instructions (Invitrogen) and quantified by measuring the absorbance at 260 nm. RNA quality was determined by measuring the 260/280 ratio and the first strand

cDNA was synthesized using the High-Capacity cDNA reverse transcription kit (Applied Biosystems®, Inc. Foster City, CA) according to the manufacturer's instructions. Briefly, 1.5 μg of total RNA from each sample was added to a mix of 2 μL of 10X reverse transcriptase buffer, 0.8 μL of 25X dNTP mix (100 mM), 2 μL of 10X reverse transcriptase random primers, 1 μL of MultiScribe™ reverse transcriptase, and 4.2 μL of nuclease-free water. The final reaction mix was kept at 25 °C for 10 min, heated to 37°C for 120 min, heated for 85°C for 5 sec, and finally cooled to 4°C[179].

2.6 Quantification of mRNA expression by real-time polymerase chain reaction (real-time-PCR)

Real-time PCR reactions were performed for Cas9 using the Mastercycler® ep realplex real-time PCR system (Eppendorf, Germany). The reaction mixture (20 μL) contained 0.06 μL of 25 μM forward primer and 0.06 μL of 25 μM reverse primer (75 nM final concentration of each primer), 10 μL of Fast SYBR Green Universal Master Mix, 9 μL of nuclease-free water, and 1 μL of cDNA sample. The primers of Cas9 were selected manually, whereas the primers of the reference gene *β*-actin were chosen from a previously published study[179] and were purchased from Integrated DNA Technologies (IDT). Assay controls were incorporated on the same plate, namely, no-template controls to test for the contamination of any assay reagents. The thermocycling program was initiated at 95 °C for 20 min, followed by 50 PCR cycles of denaturation at 95 °C for 1 s and annealing/extension at 60°C for 20 s. The PCR product was monitored by fluorescence of SYBR Green dye and the level of expression of each mRNA was normalized to simultaneously assayed *β*-actin expression. A melting curve (dissociation stage) was performed at the end of each test to ascertain the specificity of the primers and

the purity of the final PCR product. The real-time PCR data were analyzed using the relative gene expression (ΔΔCt) method, as described in Applied Biosystems User Bulletin No. 2[180]. Briefly, the data are presented as the fold change in gene expression normalized to the reference gene (*β*-actin) and were determined using the equation fold change=$2^{-\Delta\ (\Delta Ct)}$, where ΔCt=Ct (target)—Ct (*β*-actin) and Δ (ΔCt)=ΔCt (treated)—ΔCt (untreated).

Table 1: RTPCR-primers

Gene	Forward primer	Reverse primer
Cas9	5'-AAAGCTCAAAGGGTCTCCCG-3'	5'-TTATCGAGGTTAGCGTCGGC-3'
β-Actin	5′-CTGGCACCCAGCACAATG-3′	5′-GCCGATCCACACGGAGTACT-3′

2.7 Lentiviral approach

2.7.1 Oligo synthesis

Once the target site on PNKP was selected, as explained in section 2.3.1, two oligonucleotides were synthesized in the following manner: from the PNKP gene the following sequence was located on the template strand of exon 2. The underlined sequence indicates the PAM on the chromosomal target site shown below (both target strand and its complementary strand are shown) and was not included as a part of the oligonucleotides denoted in **bold** letters.

5' . . . TTCTGG**AGCACTTCCGGTCCGTAACC**TGGGTCAGGGG. . . 3'
3' . . . AAGACC**TCGTGAAGGCCAGGCATTGG**ACCCAGTCCCC. . . 5'

Additionally, BsmBI restriction enzyme recognition sequences were added at the 5' end of each oligonucleotide. The oligos were annealed so that they can form BsmBI overhangs.

5'- caccgAGCACTTCCGGTCCGTAACC - 3' Oligo 1
3' - TCGTGAAGGCCAGGCATTGGcaaa – 5' Oligo 2

Oligonucleotides, Oligo1 and Oligo 2, were purchased from IDT.

2.7.2 Phosphorylation and annealing of oligonucleotides

500 μM of Oligo 1 and Oligo 2 were mixed with 10X T4 polynnucleotide kinase buffer (New England Biolabs, USA), 1μl T4 PNK (New England Biolabs, USA) and 2 mM ATP and incubated at 37ºC for 30 min to achieve phosphorylation. To anneal the phosphorylated oligonucleotides, the sample was heated at 95ºC for 5 min and gradually allowed to cool to room temperature.

To ensure phosphorylation and annealing, the reactions were simultaneously carried out in the presence of [γ-^{32}P] ATP (PerkinElmer, USA). The samples were then electrophoresed on a 12% polyacrylamide gel and exposed to X-ray film with an intensifying screen for autoradiography.

2.7.3 Vector digestion

LentiCRISPRv2 vector was digested with FastDigest BsmBI (Fermentas, USA) and then gel purified using Gel/PCR DNA Extraction Kit (Geneaid, Taiwan) following the manufacturer's protocol.

2.7.4 Ligation

A ligation reaction was set up with BsmBI digested LetiCRISPRv2 vector, annealed oligos, 2X Quick Ligase Buffer and Quick Ligase (New England Biolabs, USA)

and incubated at room temperature for 20 min. The ligation mixture was transformed into NEB Turbo Competent *E.coli* cells (New England Biolabs, USA) and grown on Ampicillin – supplemented LB agar plates.

2.7.5 Generation of Lentiviral particles

70-80% confluent HEK293T were seeded a day before in 100-mm tissue culture dishes (Sarstedt, Germany) and allowed to adhere in a humidified incubator at 37^{o}C under 5% CO_2. Cells were co-transfected either with 5 µg of LentiCRISPRv2 vector-only (52961, Addgene) or LentiCRISPRv2-RC along with packaging vectors, 4 µg of psPAX (12260, Addgene) encoding *rev, gag/pol* genes and 5 µg of pVSVG (8454, Addgene) encoding *env* gene using Lipofectamine 2000 (Invitrogen) transfection reagent following the guidelines described in the section 'Transient Transfection'. The medium on the monolayers was replaced 16-18 h post transfection. After an additional 24 h, the medium containing lentiviral particles was collected, filtered through a 0.45µm Millex® Syringe Filter (Millipore, USA) and stored at -80ºC for later use.

2.7.6 Lentiviral infection

Approximately 5 X 10^5 A549 cells were plated per well in a 6-well cell culture plate (Corning Inc., USA) in DMEM/F12 media and allowed 24 h to adhere in a humidified incubator at 37°C and 5% CO_2. The collected medium containing lentivirus particles (as explained in Section 2.7.5) was added to the cells the following day supplemented with 8 µg/ml Polybrene (Sigma Aldrich, USA) to increase infection efficiency. After another round of infection, the infected cells were trypsinized and re-

plated into 100-mm cell culture plates in 10% DMEM/F12 medium and incubated overnight in a humidified incubator at 37°C and 5% CO_2. The media was removed the following day and replaced with 10% DMEM/F12 media containing 0.8 μg/mL Puromycin (Sigma Aldrich). The selective media was replaced every 3 days to allow for colony growth. After 10-18 days, the single-clone colonies were picked and expanded for further analysis.

2.8 Western Blotting

Approximately 10^6 cells were washed twice with ice cold PBS and resuspended in RIPA buffer (50 mM Tris-HCl, pH 7.6, 150 mM NaCl, 1% deoxycholate, 1% TritonX-100, 1 mM Na_3VO_4, 50 mM NaF, 1 mM phenylmethylsulfonyl fluoride, 1X Biotool Protease Inhibitor cocktail). Cells were then sonicated briefly and cell debris was spun down at 20,000 *g* for 20 min at 4°C. Determination of whole cell lysate protein concentration was then performed using Bradford Assay.

50 μg of protein was added to 1X sample buffer (50 mM Tris-HCl pH 6.8, 2% SDS, 10% glycerol, 1% β-mercaptoethanol, 12.5 mM EDTA, 0.02 % bromophenol blue)and boiled for 5 min. Samples were then separated by 10% SDS-PAGE (200 V for 55 min at room temperature) and transferred to a nitrocellulose membrane by wet transfer (100 V for 90 min at room temperature or 30 V overnight at 4°C). Non-specific binding was blocked with 5% skim milk in PBST (phosphate buffered saline Tween 20 solution) for 1 h at room temperature. Membranes were then immunostained with the following primary antibodies: mouse monoclonal anti-PNKP (home grown, diluted 1:300, overnight at 4°C), mouse anti-Cas9 (Active Motif, Carlsbad, 2 mg/ml diluted 1:1000, 2 h) and goat anti-

Actin (Santa Cruz Biotech, Dallas, 200 µg/ml diluted 1:2500, 1 h). Membranes were then subjected to 3 x 10 min washes in 5% milk before being incubated with the appropriate HRP-conjugated secondary antibody (Goat Anti-Mouse IgG, 28.4 mg/ml and Goat Anti-Rabbit IgG, 50 mg/ml, Jackson Immunoresearch Laboratory, PA) at a 1:5000 dilution (in 5% milk for 45 min at room temperature. Membranes were then washed 3 x 10 min in 5% milk followed by a 30 min 1X PBS wash. Then the blots were incubated with 2 ml Lumi-Light Western Blotting substrate (Roche, Mississauga, ON) for 5 min before autoradiography.

2.9 DNA sequencing

DNA sequencing was carried out using BigDye terminator v3.1 Cycle Sequencing Kit (Applied Biosciences, UK). Approximately 500 ng of plasmid DNA was combined with 1.6 picomoles of sequencing primer (Appendix II) and added to the mix of 0.5 µl Big Dye Terminator and 1X Big Dye Buffer. On a thermocycler, the sequencing reaction mixture was kept at 96ºC for 1 min followed by 35 cycles of: 96ºC for 10 sec, 50ºC for 5 sec and 60ºC for 4 min, and finally held at 4ºC.

For precipitation of PCR products, 2.5 µl of 125 mM EDTA was added to each sample and mixed well. 25 µl of absolute ethanol was added to the sample and held at room temperature for 15 min followed by centrifugation at 18000 *g* for 10 min at 4ºC. Absolute ethanol was removed carefully from the sample and washed with 70% ethanol. Ethanol was carefully drained and the samples were dried at 95ºC for 1 min on a heat block. Precipitated DNA samples were resuspended in Hi-Di Formamide Injection Buffer

(Life Technologies, USA) by heating at 95ºC for 5 min. The samples were cooled on ice for 5 min and loaded on an ABI 310 genetic analyzer for sequencing.

Chapter III

Results

3.1 Generation of PNKP knockout cell lines

3.1.1 G-Block Approach

In order to achieve PNKP knockout using the CRISPR/Cas9 system, we first set out to identify the appropriate target sites on chromosomal PNKP for gene editing. The method we employed for target identification was pinpointing all possible 19-bp sequences positioned 5' to the PAM sequence 'NGG', where N can be any nucleotide. Once we listed all the potential candidates 'N_{20}NGG' sequences on both template and non-template strands of PNKP, we performed a BLAST analysis on each candidate to check for any off-target cross-reactivity.

From a list of several candidates, we focused on four target sites on PNKP to begin with (Table 2). We then incorporated the sequences with the sequences of the sgRNA machinery and purchased them as blunt ended 453-bp double stranded G-blocks. The process of G-block synthesis is described in 'G-block synthesis' section of Chapter II. Subsequently, we cloned the G-blocks into empty pCR-Blunt II- TOPO vector. The ligation mixture was then transformed into competent cells and clones were expanded the following day. To confirm the G-block incorporation, we isolated plasmid DNA from the clones and digested with NotI restriction endonuclease. The linearized plasmid were analyzed by gel electrophoresis on a 2% agarose gel. The bands that appeared larger on the gel (~4kb) relative to the digested empty vector control (3519-bp) were selected (Fig. 7) and sequenced to confirm the directionality of the inserts.

Table 2.Targeted sequences on PNKP

Identity	Sequence	Exon
3	CACTGGCCCCAGTGACTGG	5
7	GAGGCTGTGGTGGAGAAGC	7
F	CAAGCCCTGGTCCTGGGCA	2
RC	GCACTTCCGGTCCGTAACC	2
Scramble*	GCACTTCCGGTCCGTAACC	-

Scramble* - The scramble sequence was manually constructed in such a way that it is minimally identical to any region of the chromosome. The identity was confirmed using NCBI Blast tool.

Simultaneously, we acquired hCas9-vector bearing a human codon optimized Cas9-cassette and created two different mutant forms of the gene, Cas9-D10A (nickase) or Cas9n and Cas9-D10A/H840A (catalytically inactive) or dCas9, using site-directed mutagenesis. Next, we stably expressed these constructs in A549 cells to develop Cas9 and Cas9-mutant cell lines. We confirmed the expression levels of Cas9 and mutants using RT-PCR (Chart 1).

In next step, we stably transfected Cas9, Cas9n and dCas9 cells with pCR-Blunt II vectors carrying the G-blocks. At the time of writing of this thesis, we are still waiting for the results of the selection experiment. Once the colonies form, the next step would be to expand the clones and examine PNKP expression levels using immunoblotting or RTPCR.

Fig. 7

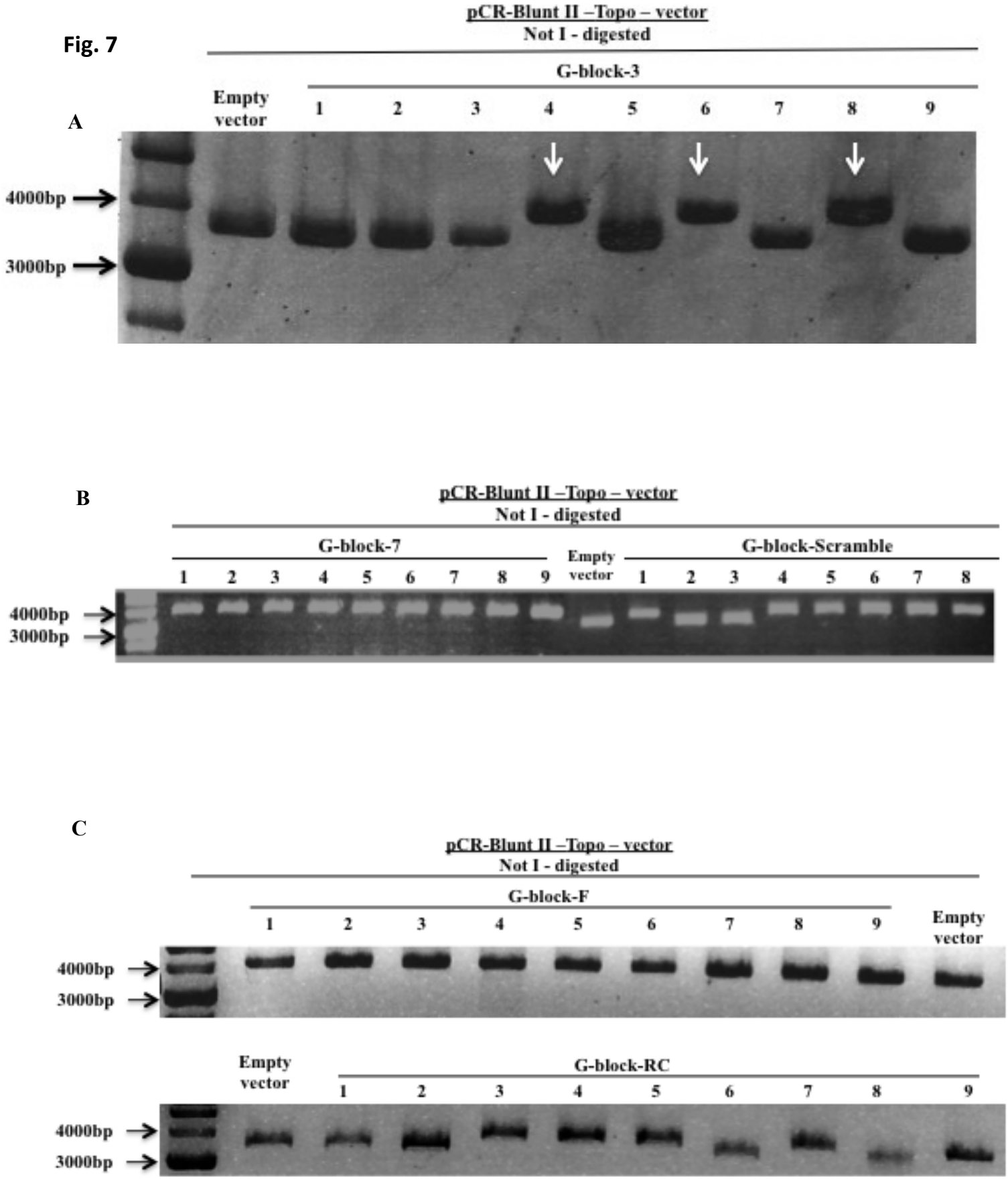

A
pCR-Blunt II –Topo – vector
Not I - digested
G-block-3
Empty vector
1 2 3 4 5 6 7 8 9
4000bp
3000bp
B
pCR-Blunt II –Topo – vector
Not I - digested
G-block-7
Empty vector
G-block-Scramble
1 2 3 4 5 6 7 8 9
1 2 3 4 5 6 7 8
4000bp
3000bp
C
pCR-Blunt II –Topo – vector
Not I - digested
G-block-F
1 2 3 4 5 6 7 8 9
Empty vector
4000bp
3000bp
Empty vector
G-block-RC
1 2 3 4 5 6 7 8 9
4000bp
3000bp

Figure 7. G-Block cloning: 435-bp long blunt double stranded G-blocks were cloned into pCR-Blunt II-Topo vector. At least 8 colonies were expanded and plasmid DNA was isolated and subjected to NotI restriction enzyme digestion and the linearized plasmids were electrophoresed at a low voltage in a 2% agarose gel. Linearized empty vector of length 3519-bp was run alongside as a negative control. (a) A successful integration of G-blocks was confirmed by visualizing a slightly larger fragment of 3954-bp (~4kb) on the gel as indicated by white arrows. The positive clones (lanes 4, 6 and 8) were then sequenced to confirm the directionality of the G-block inserts. An identical approach was undertaken with other G-blocks. (b) gel examing the clones obtained by cloning G-block 7 and G-block-scramble and (c) the clones obtained by cloning G-block-F (upper gel) and G-block-RC (lower gel).

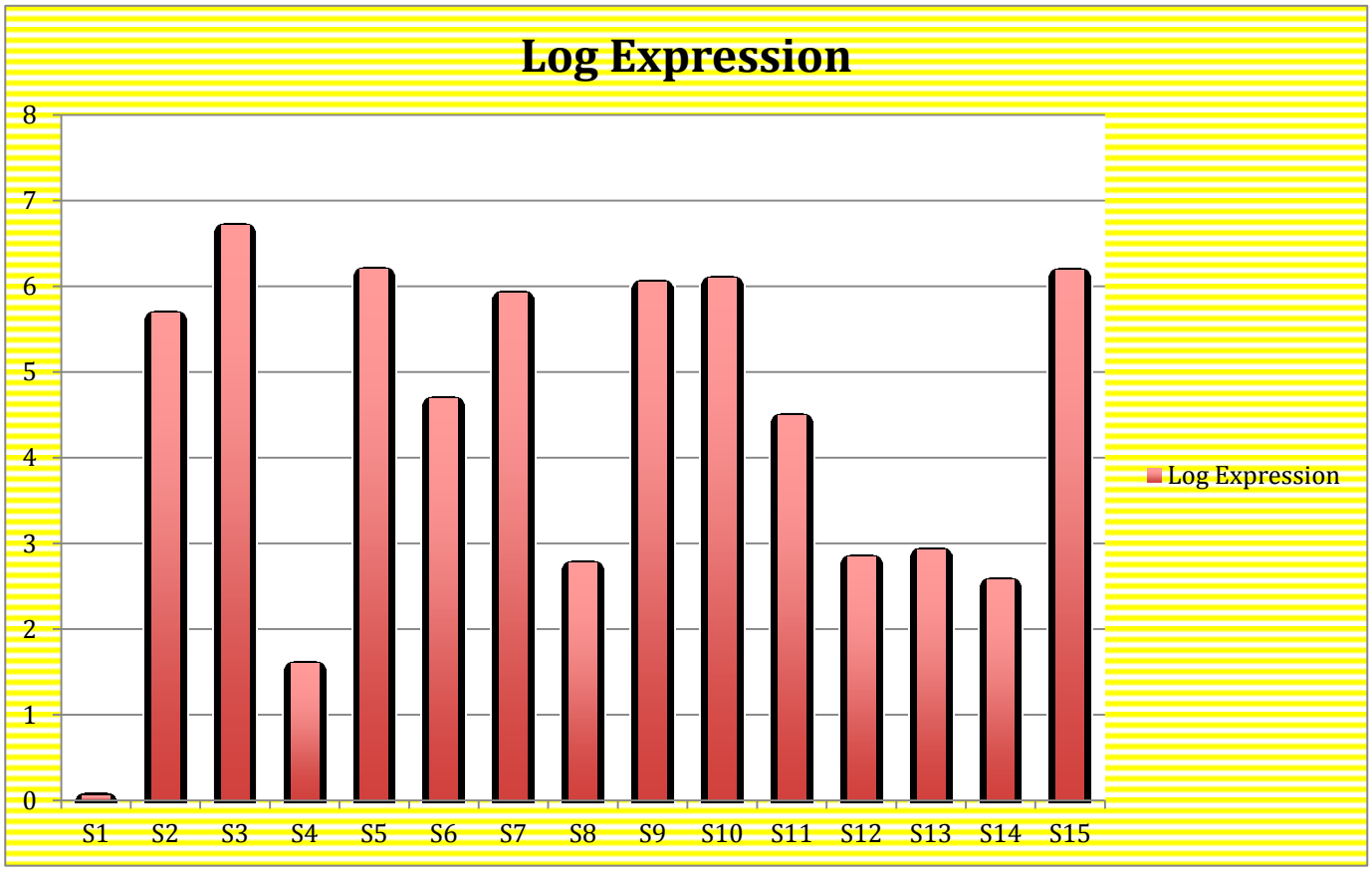


Figure 8. RT-PCR results showing expression levels of Cas9 and Cas9-mutants in A549 cells: Graph shows expression levels of Cas9 and Cas9-mutants in different clones in stably transfected A549 cells. S1: A549 cells, negative control; S2-S3: Cas9; S4-S11: Cas9n; S12-S15: dCas9. S3, S5 and S15 cell lines were selected for further experiments.

3.1.2 Lentiviral Approach

In this approach we made use of pLentiCRISPR v2 vector, a bicistronic lentivirus-based vector harboring cassettes for Cas9 and machinery for sgRNA. The target sites were identified in the same fashion as explained in section 3.1.1. Additionally, we consulted the data published by Shalem et al.[2] and concentrated on just one target site of PNKP (GCACTTCCGGTCCGTAACC) present on the template strand on exon 2. We first designed two oligonucleotides, Oligo 1 and Oligo 2, specific to the target site with BsmBI restriction sequence overhangs, hybridized them together to generate PNKP-RC fragments and cloned them into BsmBI digested pLentiCRISPRv2 vector to complete the sgRNA machinery. The successful integration of PNKP-RC fragment was confirmed by sequencing (Fig. 9). We named the vector carrying PNKP-RC fragment pLentiCRISPR-PNKP-RC.

To generate lentiviral particles encapsulating Cas9 and PNKP-specific sgRNA, we co-transfected HEK293T cells with pLentiCRISPR-PNKP-RC and packaging vectors psPAX2 and pVSVg. We used pLentiCRISPRv2 vector-only as a control for the experiment. The media from both control and experimental wells were collected from the culture and used to infect A549 cells. The infected cells were kept under Puromycin selection until colonies appeared. The clones were individually picked and expanded separately as independent cell lines.

To test whether PNKP knockout had been achieved, whole cell lysates were prepared from the expanded clones and analyzed using western blots. (In order to ascertain the specificity and effectiveness of newly purchased Cas9-antibody, a western blot analysis was conducted before hand on lysates isolated from HeLa cells that were

transiently transfected with Cas9 carrying vector (Fig. 10)). As shown in Figure 11, clones RC1 and RC2 showed a significant ablation of PNKP. In order to isolate a population of cells completely lacking PNKP, we plated RC1 cells at low dilution (~ 3 cells per well) in a 96-well plate. Then the plate was incubated until colonies were formed. Subsequently, the colonies were expanded and whole cell lysates prepared for western blot analysis. As shown in Figure 12, the RC1-subclones showed a complete ablation of PNKP when compared to wild type A549.

Fig. 9

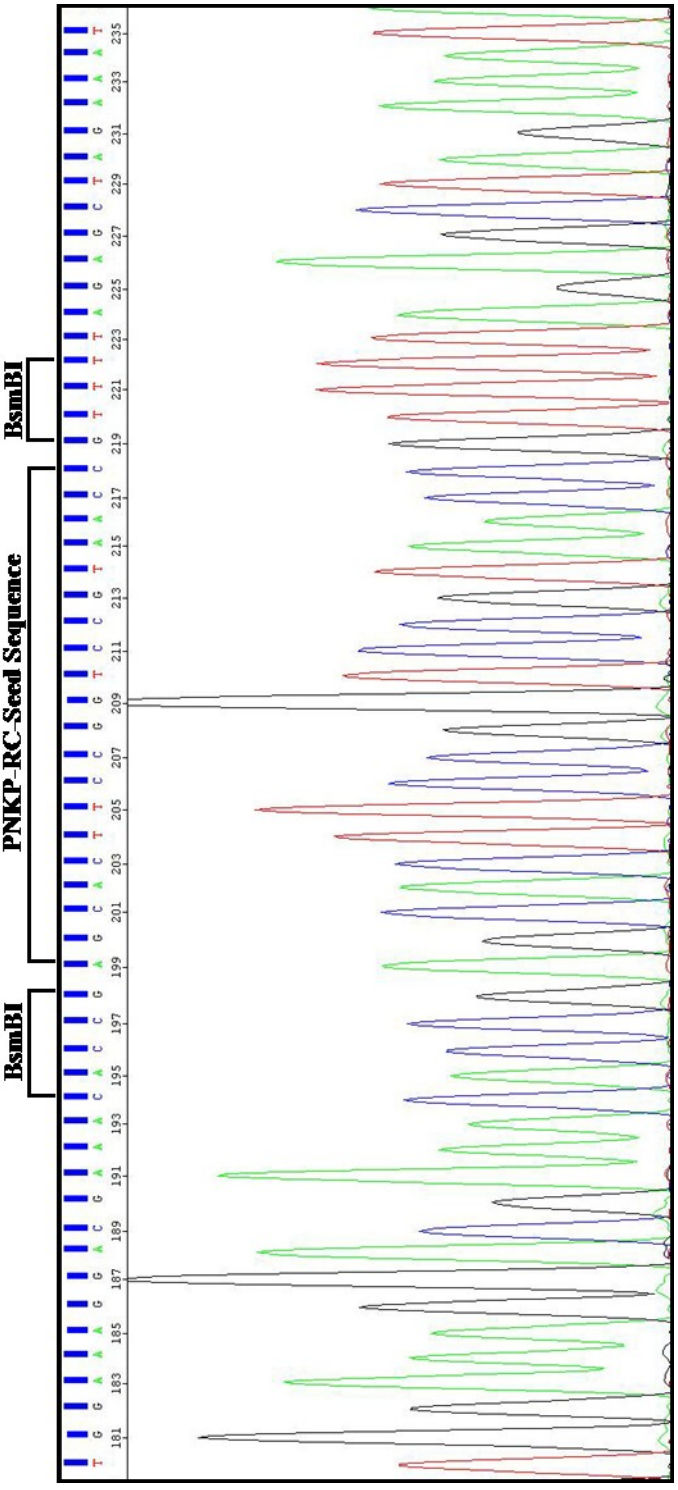


Figure 9. Sequencing pLentiCRISPR-PNKP-RC: PNKP-RC fragments were ligated into pLentiCRISPRv2 vectors and transformed into competent *E.coli* cells. 20 different clones were picked and plasmid DNAs were isolated. The sequencing PCR and precipitation were carried out as described in Section 2.9. The sequencing results were analyzed as depicted in the figure.

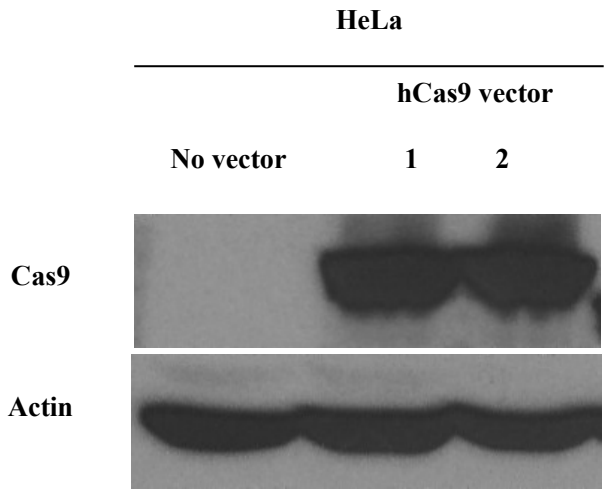


Figure 10. Expression of Cas9 in HeLa cells: HeLa cells were transfected with hCas9 vector and whole cell lysates were prepared. 50μg of cell lysates were loaded in duplicate and electrophoresed in 10% SDS-PAGE gels. The proteins were transferred to a nitrocellulose membrane and immunostained with anti-Cas9 antibodies. Actin was used as a loading control.

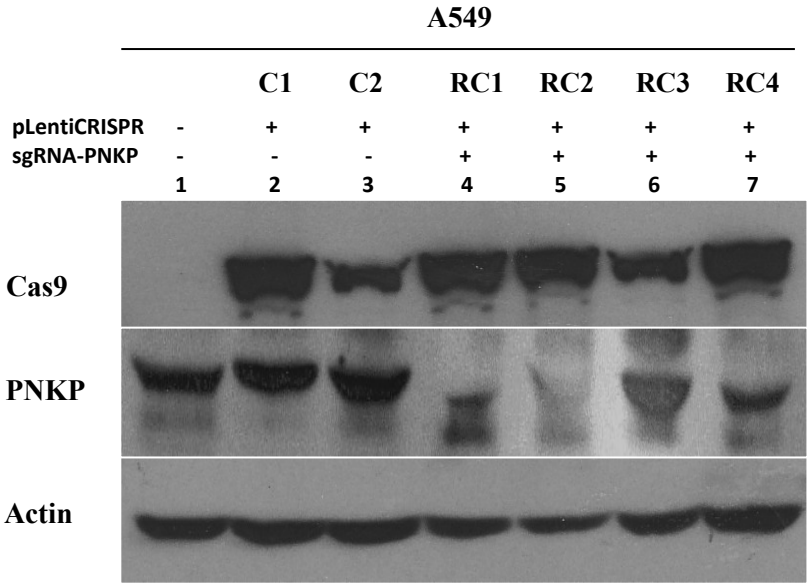


Figure 11. Expression levels of PNKP in pLentiCRISPR-PNKP-RC transduced A549 cells: Annealed Oligo 1/Oligo 2 (PNKP-RC-fragment) was integrated into the lentiviral vector pLentiCRISPR and lentiviral particles were generated using appropriate packaging vectors. pLentiCRISPRv2-vector-only was used to generate viral particles simultaneously, which acted as a negative control for the experiment. The viral particles encapsulating the CRISPR-machinery were used to infect A549 cells. After two-rounds of infection, the cells were cultured in a selective medium containing Puromycin until colonies developed. The colonies were subsequently picked and expanded as individual cell lines. Lane 1 of the western blot represents the cell lysates from wild-type A549. Lanes 2 and 3 represent the cells that were infected with viral particles containing vector only as indicated by robust expression of Cas9 protein and normal levels of PNKP. Lanes 4-7 (RC1 – RC4) represent the cells that were infected with viral particles carrying pLentiCRISPR-PNKP-RC plasmid. RC1 and RC2 show a marked reduction in the levels of PNKP protein compared to control cell lysates.

Fig. 12

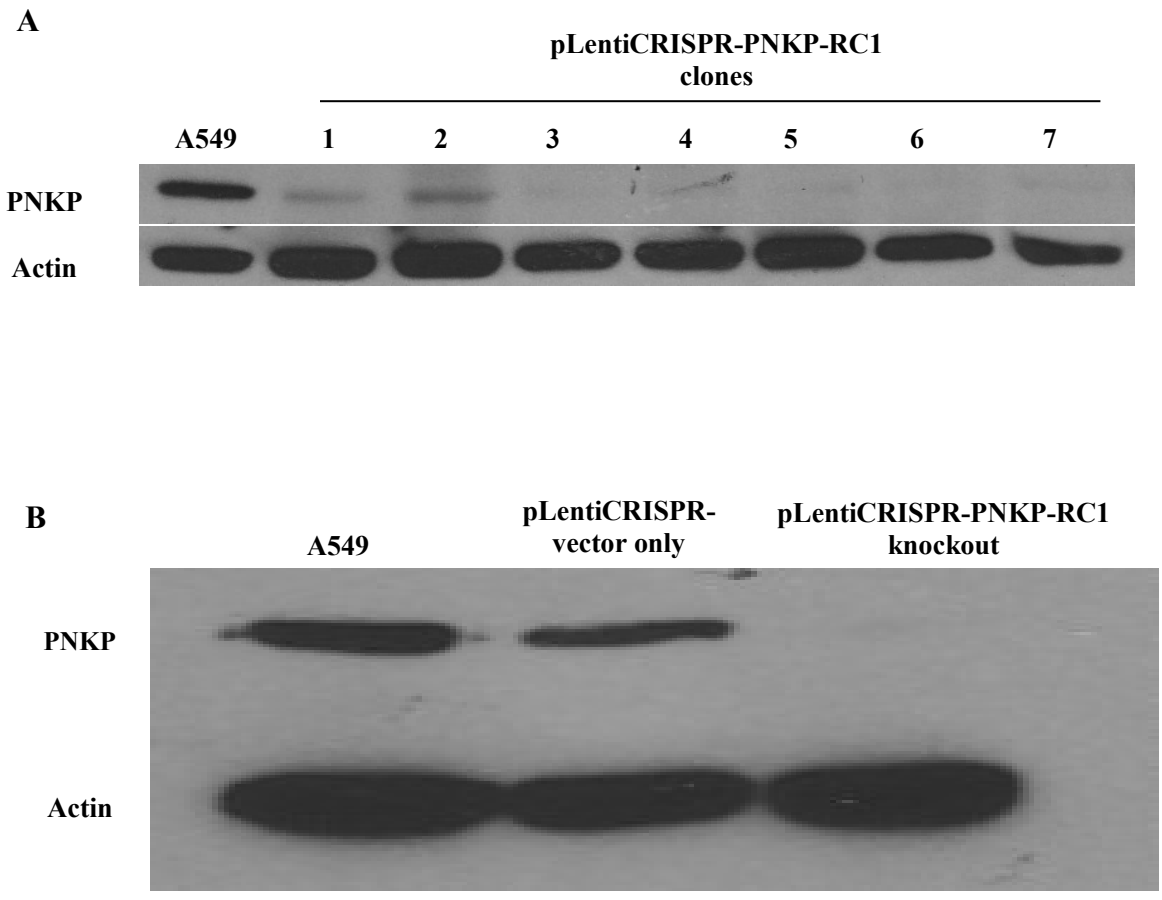

A
pLentiCRISPR-PNKP-RC1
clones
A549
1
2
3
4
5
6
7
PNKP
Actin
B
A549
pLentiCRISPR-
vector only
pLentiCRISPR-PNKP-RC1
knockout
PNKP
Actin

Figure 12. Expression levels of PNKP in serial diluted pLentiCRISPR-PNKP-RC1-transduced A549 cells: A. pLentiCRISPR-PNKP-RC1 cells were trypsinized, pelleted and resuspended in DMEM/F12 media. The cells were diluted to a low density in a 96-well cell culture plate and incubated until colonies became visible. Subsequently, the colonies were trypsinized and expanded as individual cell lines. The western blot of the whole cell lysates that were harvested from the colonies is shown. Lane 1 represents the cell lysate derived from wild-type A549 cells. Lanes 1-7 represent cell lysates derived from individual colonies isolated from pLentiCRISPR-PNKP-RC1-transduced A549 cells. Lanes 3-7 show an almost complete ablation of PNKP when compared to wild-type cells.

B. An immunoblot showing comparative expression levels of PNKP. Lane 3 represents the cells subjected to CRISPR-mediated knockout approach targeting PNKP where we see a robust ablation of PNKP. Lane 2 represents the vector-only cells and Lane 1 represents wild-type A549 cells.

Chapter IV

Discussion

4.1 Discussion

Over the past three years, there has been an explosion of interest in the field of gene editing using CRISPR/Cas9 method. Different research groups have engineered and perfected the method in mammalian cells amongst other organisms. This feat was accomplished by expressing codon optimizing *S. pyogenes*-derived Cas9 in the nucleus along with a minimal structure of noncoding CRISPR RNA, termed sgRNA. To date, the method has been used for various purposes including gene knockout, knock-in, site directed mutagenesis, gene interference, GFP-based site-directed chromosomal visualization, site-directed DNA purification, gene activation, epigenetic modulation and many more[10, 174]. And the field is continuously evolving to accommodate even more applications.

In this study, we used the CRISPR/ Cas9 method to knock out PNKP in A549 cells. Of the two approaches we used, the two-plasmid system (G-block approach) is still in progress at the time of writing of this thesis. However, the lentiviral approach using a bicistronic plasmid bearing both Cas9 and PNKP-specific sgRNA machinery (Lentiviral approach) has successfully resulted in ablation of PNKP. Utilizing antibiotic based selective medium and serial dilution technique, we were able to expand populations of cells that showed a complete repression of PNKP, as indicated by immunoblot results. However, there are a few necessary experiments required to be performed to support our results:

- Although immunoblot results are very convincing, an immediate RT-PCR assay needs to be performed to quantitate PNKP expression for confirmatory purposes.

- Since a knockout with the CRISPR/Cas9 approach is achieved by introducing frameshift mutations and premature stop codons at the target site, it would be very interesting to see the form of modification (indels) that the chromosomal PNKP has sustained post-NHEJ repair in different clones that we have established. Sequencing of chromosomal PNKP at location 19q13.3-q13.4 (base pairs 49,861,202 to 49,867,564 on chromosome 19) would be an immediate step to follow.
- A clonogenic cell survival assay needs to be performed in order to analyze the effect of PNKP-knock out on cell survival compared to wild type cells.

Incorporating different target seed sequences within sgRNA to obtain DSBs or editing at an intended site can be a time consuming and resource intensive work. Fortunately, there are assays to ascertain whether the DSBs or editing has occurred at the target site. One of the ways to test the effectiveness of various sgRNAs is by transiently transfecting the cells with plasmid(s) carrying Cas9 and sgRNA thereby generating a mixed population of edited and unedited cells (process explained in detail in [181]. In the following step, genomic DNA is isolated from the cells and only the target region within the gene is amplified using PCR. Next, the amplicons are denatured and re-hybridized using a thermocycler to generate heteroduplex pairs sustaining mismatches (e.g., indels). The sample is then digested with T7E1, a restriction enzyme known to selectively introduce cleavage at DNA mismatches[182]. A successful generation of DSBs or editing can be determined simply by electrophoresing the T7E1 digested DNA sample.

Alternatively, there are commercial kits available, such as GeneArt® Genomic Cleavage Detection Kit manufactured by Invitrogen, for the DSB/editing determination.

Intriguingly, CRISPR/Cas9 technique has also been developed for multiplex genome engineering[173]. That means more than one gene can be simultaneously knocked out from cells by using multiple sgRNAs bearing complementary sequences specific to multiple genes. For example, the multiplexing approach can be applied to knockout out PNKP (or any other gene under investigation) simultaneously with its synthetic lethal partners eliminating the need of co-transfection, which would have been the only option with the shRNA-based technique. Our lab has previously identified tumor suppressors PTEN and SHP1 as synthetic lethal partners of PNKP using an siRNA-based approach to screen ~7000 genes[128, 129]. Additionally, 423 other potential synthetic lethal candidates to PNKP were discovered including several tumor suppressors, which could be critical from the viewpoint of cancer therapy[128, 129]. Therefore, a CRISPR-mediated multiplexing knocking out strategy of genes would appear to be a viable approach for synthetic lethal studies.

Knocking-in of mutations or new DNA sequences at a desired location within a chromosome is another interesting application of CRISPR/Cas9-approach[183]. For knock-ins, an additional donor plasmid has to be introduced along with Cas9 and sgRNA. The donor plasmid contains the desired sequence modification (point mutations, insertions or deletions) flanked on either side by sequences homologous to the site on the chromosome where the knock-in is intended. Introduction of all three requisites triggers gene editing through HR repair thereby accomplishing the knock-in of the new sequence[183]. With regard to PNKP it should be possible to introduce sequences to

achieve selective inactivation of the key FHA, kinase and phosphatase domains of the protein. That way, we will be able to develop cell lines expressing versions of PNKP with particular domains inactivated eliminating the need of reconstituting plasmids with mutant versions of PNKP-domains in PNKP-null cells. Similarly, mutations within the localization signals of PNKP could be introduced in order to control its trafficking within the cell.

Part III

Future Directions

3.1 Determine the role of PNKP in the repair of nuclear and mtDNA

Now we plan to stably express PNKP-carrying pShooter vectors with inactivating mutations in either the NLS or the MTS in cells stably depleted of PNKP by the CRISPR approach. We will examine the relative effects of NLS and MTS mutations on DNA repair in cells challenged with IR, H_2O_2, and camptothecin, which is a topoisomerase I inhibitor. (Both nuclear and mitochondrial topoisomerase I can be inactivated by camptothecin[184]). We will employ single-cell gel electrophoresis under alkali and neutral conditions to monitor repair of SSBs and DSBs in nuclear DNA and "extra-large"-qPCR for mtDNA repair. Both of these techniques have been extensively used in our laboratory[8, 156]. It is likely that there is a greater backup repair capacity in the nucleus than in mitochondria[185], which would lead us to predict that the loss of PNKP will impair mtDNA repair more significantly than nuclear DNA repair.

3.2 Determine the relative roles of nuclear and mitochondrial PNKP in the survival response to DNA damaging agents

Although mitochondria play key roles in apoptosis and senescence, it remains an intriguing question as to whether mtDNA damage can trigger either process. We will use the cells described above in clonogenic assays to examine cell survival in response to IR, H_20_2 and camptothecin. Since H_20_2 and camptothecin will be ubiquitously located throughout the cell mtDNA damage may contribute to both H_20_2 and camptothecin induced cytotoxicity. Recent reports suggest that mtDNA is more susceptible to IR compared to nuclear DNA. mtDNA was found to sustain relatively greater amounts of

lesions and breaks when treated with IR[186]. In light of these facts, it would be interesting to observe how cells behave in terms of survival when exposed to various DNA damaging agents.

Reference

1. Caldecott, K.W., *DNA single-strand break repair and spinocerebellar ataxia.* Cell, 2003. **112**(1): p. 7-10.
2. Shalem, O., et al., *Genome-scale CRISPR-Cas9 knockout screening in human cells.* Science, 2014. **343**(6166): p. 84-7.
3. Bernstein, N.K., et al., *The molecular architecture of the mammalian DNA repair enzyme, polynucleotide kinase.* Mol Cell, 2005. **17**(5): p. 657-70.
4. Gorlich, D., *Transport into and out of the cell nucleus.* EMBO J, 1998. **17**(10): p. 2721-7.
5. Valerie, K. and L.F. Povirk, *Regulation and mechanisms of mammalian double-strand break repair.* Oncogene, 2003. **22**(37): p. 5792-812.
6. Bellance, N., P. Lestienne, and R. Rossignol, *Mitochondria: from bioenergetics to the metabolic regulation of carcinogenesis.* Front Biosci (Landmark Ed), 2009. **14**: p. 4015-34.
7. Helleday, T., et al., *DNA double-strand break repair: from mechanistic understanding to cancer treatment.* DNA Repair (Amst), 2007. **6**(7): p. 923-35.
8. Tahbaz, N., S. Subedi, and M. Weinfeld, *Role of polynucleotide kinase/phosphatase in mitochondrial DNA repair.* Nucleic Acids Res, 2012. **40**(8): p. 3484-95.
9. Weinfeld, M., et al., *Tidying up loose ends: the role of polynucleotide kinase/phosphatase in DNA strand break repair.* Trends Biochem Sci, 2011. **36**(5): p. 262-71.
10. Mali, P., K.M. Esvelt, and G.M. Church, *Cas9 as a versatile tool for engineering biology.* Nat Methods, 2013. **10**(10): p. 957-63.
11. Sander, J.D. and J.K. Joung, *CRISPR-Cas systems for editing, regulating and targeting genomes.* Nat Biotechnol, 2014. **32**(4): p. 347-55.
12. Hoeijmakers, J.H., *Genome maintenance mechanisms for preventing cancer.* Nature, 2001. **411**(6835): p. 366-74.
13. Becker, T., L. Bottinger, and N. Pfanner, *Mitochondrial protein import: from transport pathways to an integrated network.* Trends Biochem Sci, 2012. **37**(3): p. 85-91.
14. Larson, M.H., et al., *CRISPR interference (CRISPRi) for sequence-specific control of gene expression.* Nat Protoc, 2013. **8**(11): p. 2180-96.
15. Marraffini, L.A. and E.J. Sontheimer, *CRISPR interference: RNA-directed adaptive immunity in bacteria and archaea.* Nat Rev Genet, 2010. **11**(3): p. 181-90.
16. Pfeiffer, P., W. Goedecke, and G. Obe, *Mechanisms of DNA double-strand break repair and their potential to induce chromosomal aberrations.* Mutagenesis, 2000. **15**(4): p. 289-302.
17. Gostissa, M., F.W. Alt, and R. Chiarle, *Mechanisms that promote and suppress chromosomal translocations in lymphocytes.* Annu Rev Immunol, 2011. **29**: p. 319-50.
18. Lindahl, T., *Instability and decay of the primary structure of DNA.* Nature, 1993. **362**(6422): p. 709-15.

19. Shuck, S.C., E.A. Short, and J.J. Turchi, *Eukaryotic nucleotide excision repair: from understanding mechanisms to influencing biology.* Cell Res, 2008. **18**(1): p. 64-72.
20. Zharkov, D.O., *Base excision DNA repair.* Cell Mol Life Sci, 2008. **65**(10): p. 1544-65.
21. Friedberg, E.C., *DNA damage and repair.* Nature, 2003. **421**(6921): p. 436-40.
22. Lindahl, T., *DNA glycosylases, endonucleases for apurinic/apyrimidinic sites, and base excision-repair.* Prog Nucleic Acid Res Mol Biol, 1979. **22**: p. 135-92.
23. Dianov, G.L. and U. Hubscher, *Mammalian base excision repair: the forgotten archangel.* Nucleic Acids Res, 2013. **41**(6): p. 3483-90.
24. Ide, H. and M. Kotera, *Human DNA glycosylases involved in the repair of oxidatively damaged DNA.* Biol Pharm Bull, 2004. **27**(4): p. 480-5.
25. Matsumoto, Y. and K. Kim, *Excision of deoxyribose phosphate residues by DNA polymerase beta during DNA repair.* Science, 1995. **269**(5224): p. 699-702.
26. Boiteux, S. and J.P. Radicella, *The human OGG1 gene: structure, functions, and its implication in the process of carcinogenesis.* Arch Biochem Biophys, 2000. **377**(1): p. 1-8.
27. Aspinwall, R., et al., *Cloning and characterization of a functional human homolog of Escherichia coli endonuclease III.* Proc Natl Acad Sci U S A, 1997. **94**(1): p. 109-14.
28. Satoh, M.S. and T. Lindahl, *Role of poly(ADP-ribose) formation in DNA repair.* Nature, 1992. **356**(6367): p. 356-8.
29. Lilyestrom, W., et al., *Structural and biophysical studies of human PARP-1 in complex with damaged DNA.* J Mol Biol, 2010. **395**(5): p. 983-94.
30. Masaoka, A., et al., *DNA polymerase beta and PARP activities in base excision repair in living cells.* DNA Repair (Amst), 2009. **8**(11): p. 1290-9.
31. Sukhanova, M., S. Khodyreva, and O. Lavrik, *Poly(ADP-ribose) polymerase 1 regulates activity of DNA polymerase beta in long patch base excision repair.* Mutat Res, 2010. **685**(1-2): p. 80-9.
32. Ahel, I., et al., *Poly(ADP-ribose)-binding zinc finger motifs in DNA repair/checkpoint proteins.* Nature, 2008. **451**(7174): p. 81-5.
33. Gagne, J.P., et al., *Proteome-wide identification of poly(ADP-ribose) binding proteins and poly(ADP-ribose)-associated protein complexes.* Nucleic Acids Res, 2008. **36**(22): p. 6959-76.
34. El-Khamisy, S.F., et al., *A requirement for PARP-1 for the assembly or stability of XRCC1 nuclear foci at sites of oxidative DNA damage.* Nucleic Acids Res, 2003. **31**(19): p. 5526-33.
35. Audebert, M., B. Salles, and P. Calsou, *Involvement of poly(ADP-ribose) polymerase-1 and XRCC1/DNA ligase III in an alternative route for DNA double-strand breaks rejoining.* J Biol Chem, 2004. **279**(53): p. 55117-26.
36. Lindahl, T., et al., *Post-translational modification of poly(ADP-ribose) polymerase induced by DNA strand breaks.* Trends Biochem Sci, 1995. **20**(10): p. 405-11.
37. Mani, R.S., et al., *XRCC1 stimulates polynucleotide kinase by enhancing its damage discrimination and displacement from DNA repair intermediates.* J Biol Chem, 2007. **282**(38): p. 28004-13.

38. Jilani, A., et al., *Molecular cloning of the human gene, PNKP, encoding a polynucleotide kinase 3'-phosphatase and evidence for its role in repair of DNA strand breaks caused by oxidative damage.* J Biol Chem, 1999. **274**(34): p. 24176-86.
39. Demple, B. and L. Harrison, *Repair of oxidative damage to DNA: enzymology and biology.* Annu Rev Biochem, 1994. **63**: p. 915-48.
40. Ahel, I., et al., *The neurodegenerative disease protein aprataxin resolves abortive DNA ligation intermediates.* Nature, 2006. **443**(7112): p. 713-6.
41. Cortes Ledesma, F., et al., *A human 5'-tyrosyl DNA phosphodiesterase that repairs topoisomerase-mediated DNA damage.* Nature, 2009. **461**(7264): p. 674-8.
42. Caldecott, K.W., *Single-strand break repair and genetic disease.* Nat Rev Genet, 2008. **9**(8): p. 619-31.
43. Prasad, R., et al., *A review of recent experiments on step-to-step "hand-off" of the DNA intermediates in mammalian base excision repair pathways.* Mol Biol (Mosk), 2011. **45**(4): p. 586-600.
44. Balakrishnan, L., et al., *Long patch base excision repair proceeds via coordinated stimulation of the multienzyme DNA repair complex.* J Biol Chem, 2009. **284**(22): p. 15158-72.
45. Saleh-Gohari, N., et al., *Spontaneous homologous recombination is induced by collapsed replication forks that are caused by endogenous DNA single-strand breaks.* Mol Cell Biol, 2005. **25**(16): p. 7158-69.
46. Davis, A.J. and D.J. Chen, *DNA double strand break repair via non-homologous end-joining.* Transl Cancer Res, 2013. **2**(3): p. 130-143.
47. Hartlerode, A.J. and R. Scully, *Mechanisms of double-strand break repair in somatic mammalian cells.* Biochem J, 2009. **423**(2): p. 157-68.
48. Mari, P.O., et al., *Dynamic assembly of end-joining complexes requires interaction between Ku70/80 and XRCC4.* Proc Natl Acad Sci U S A, 2006. **103**(49): p. 18597-602.
49. Uematsu, N., et al., *Autophosphorylation of DNA-PKCS regulates its dynamics at DNA double-strand breaks.* J Cell Biol, 2007. **177**(2): p. 219-29.
50. Meek, K., V. Dang, and S.P. Lees-Miller, *DNA-PK: the means to justify the ends?* Adv Immunol, 2008. **99**: p. 33-58.
51. Nick McElhinny, S.A. and D.A. Ramsden, *Sibling rivalry: competition between Pol X family members in V(D)J recombination and general double strand break repair.* Immunol Rev, 2004. **200**: p. 156-64.
52. Mao, Z., et al., *DNA repair by nonhomologous end joining and homologous recombination during cell cycle in human cells.* Cell Cycle, 2008. **7**(18): p. 2902-6.
53. Moreno-Herrero, F., et al., *Mesoscale conformational changes in the DNA-repair complex Rad50/Mre11/Nbs1 upon binding DNA.* Nature, 2005. **437**(7057): p. 440-3.
54. Haince, J.F., et al., *PARP1-dependent kinetics of recruitment of MRE11 and NBS1 proteins to multiple DNA damage sites.* J Biol Chem, 2008. **283**(2): p. 1197-208.
55. Williams, R.S., et al., *Mre11 dimers coordinate DNA end bridging and nuclease processing in double-strand-break repair.* Cell, 2008. **135**(1): p. 97-109.

56. Lee, J.H. and T.T. Paull, *ATM activation by DNA double-strand breaks through the Mre11-Rad50-Nbs1 complex.* Science, 2005. **308**(5721): p. 551-4.
57. Carvalho, J.F. and R. Kanaar, *Targeting homologous recombination-mediated DNA repair in cancer.* Expert Opin Ther Targets, 2014. **18**(4): p. 427-58.
58. Sung, P. and H. Klein, *Mechanism of homologous recombination: mediators and helicases take on regulatory functions.* Nat Rev Mol Cell Biol, 2006. **7**(10): p. 739-50.
59. Bugreev, D.V., O.M. Mazina, and A.V. Mazin, *Bloom syndrome helicase stimulates RAD51 DNA strand exchange activity through a novel mechanism.* J Biol Chem, 2009. **284**(39): p. 26349-59.
60. Asin-Cayuela, J. and C.M. Gustafsson, *Mitochondrial transcription and its regulation in mammalian cells.* Trends Biochem Sci, 2007. **32**(3): p. 111-7.
61. Taylor, R.W. and D.M. Turnbull, *Mitochondrial DNA mutations in human disease.* Nat Rev Genet, 2005. **6**(5): p. 389-402.
62. Alexeyev, M., et al., *The maintenance of mitochondrial DNA integrity--critical analysis and update.* Cold Spring Harb Perspect Biol, 2013. **5**(5): p. a012641.
63. Ojala, D., J. Montoya, and G. Attardi, *tRNA punctuation model of RNA processing in human mitochondria.* Nature, 1981. **290**(5806): p. 470-4.
64. Anderson, L., *Identification of mitochondrial proteins and some of their precursors in two-dimensional electrophoretic maps of human cells.* Proc Natl Acad Sci U S A, 1981. **78**(4): p. 2407-11.
65. O'Rourke, T.W., et al., *Mitochondrial dysfunction due to oxidative mitochondrial DNA damage is reduced through cooperative actions of diverse proteins.* Mol Cell Biol, 2002. **22**(12): p. 4086-93.
66. Wallace, D.C., et al., *Mitochondrial DNA mutations in human degenerative diseases and aging.* Biochim Biophys Acta, 1995. **1271**(1): p. 141-51.
67. Barja, G. and A. Herrero, *Oxidative damage to mitochondrial DNA is inversely related to maximum life span in the heart and brain of mammals.* FASEB J, 2000. **14**(2): p. 312-8.
68. Mason, P.A., et al., *Mismatch repair activity in mammalian mitochondria.* Nucleic Acids Res, 2003. **31**(3): p. 1052-8.
69. Khan, S.K., et al., *Identification of a novel cryptochrome differentiating domain required for feedback repression in circadian clock function.* J Biol Chem, 2012. **287**(31): p. 25917-26.
70. Christmann, M., et al., *Mechanisms of human DNA repair: an update.* Toxicology, 2003. **193**(1-2): p. 3-34.
71. Myers, K.A., R. Saffhill, and P.J. O'Connor, *Repair of alkylated purines in the hepatic DNA of mitochondria and nuclei in the rat.* Carcinogenesis, 1988. **9**(2): p. 285-92.
72. Modrich, P., *Mechanisms in eukaryotic mismatch repair.* J Biol Chem, 2006. **281**(41): p. 30305-9.
73. Martin, S.A., et al., *DNA polymerases as potential therapeutic targets for cancers deficient in the DNA mismatch repair proteins MSH2 or MLH1.* Cancer Cell, 2010. **17**(3): p. 235-48.
74. de Souza-Pinto, N.C., et al., *Novel DNA mismatch-repair activity involving YB-1 in human mitochondria.* DNA Repair (Amst), 2009. **8**(6): p. 704-19.

75. Nakabeppu, Y., *Molecular genetics and structural biology of human MutT homolog, MTH1.* Mutat Res, 2001. **477**(1-2): p. 59-70.
76. Gilkerson, R.W., et al., *Mitochondrial nucleoids maintain genetic autonomy but allow for functional complementation.* J Cell Biol, 2008. **181**(7): p. 1117-28.
77. Kajander, O.A., et al., *Prominent mitochondrial DNA recombination intermediates in human heart muscle.* EMBO Rep, 2001. **2**(11): p. 1007-12.
78. Kazak, L., A. Reyes, and I.J. Holt, *Minimizing the damage: repair pathways keep mitochondrial DNA intact.* Nat Rev Mol Cell Biol, 2012. **13**(10): p. 659-71.
79. Sage, J.M., O.S. Gildemeister, and K.L. Knight, *Discovery of a novel function for human Rad51: maintenance of the mitochondrial genome.* J Biol Chem, 2010. **285**(25): p. 18984-90.
80. Duxin, J.P., et al., *Human Dna2 is a nuclear and mitochondrial DNA maintenance protein.* Mol Cell Biol, 2009. **29**(15): p. 4274-82.
81. Bacman, S.R., S.L. Williams, and C.T. Moraes, *Intra- and inter-molecular recombination of mitochondrial DNA after in vivo induction of multiple double-strand breaks.* Nucleic Acids Res, 2009. **37**(13): p. 4218-26.
82. Gredilla, R., *DNA damage and base excision repair in mitochondria and their role in aging.* J Aging Res, 2010. **2011**: p. 257093.
83. Hu, J., et al., *Repair of formamidopyrimidines in DNA involves different glycosylases: role of the OGG1, NTH1, and NEIL1 enzymes.* J Biol Chem, 2005. **280**(49): p. 40544-51.
84. Hazra, T.K., et al., *Identification and characterization of a human DNA glycosylase for repair of modified bases in oxidatively damaged DNA.* Proc Natl Acad Sci U S A, 2002. **99**(6): p. 3523-8.
85. Mandal, S.M., et al., *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): p. 2819-29.
86. Pagliarini, D.J., et al., *A mitochondrial protein compendium elucidates complex I disease biology.* Cell, 2008. **134**(1): p. 112-23.
87. Chacinska, A., et al., *Importing mitochondrial proteins: machineries and mechanisms.* Cell, 2009. **138**(4): p. 628-44.
88. Mokranjac, D. and W. Neupert, *Protein import into mitochondria.* Biochem Soc Trans, 2005. **33**(Pt 5): p. 1019-23.
89. Endo, T. and K. Yamano, *Multiple pathways for mitochondrial protein traffic.* Biol Chem, 2009. **390**(8): p. 723-30.
90. Boengler, K., G. Heusch, and R. Schulz, *Nuclear-encoded mitochondrial proteins and their role in cardioprotection.* Biochim Biophys Acta, 2011. **1813**(7): p. 1286-94.
91. Dudek, J., P. Rehling, and M. van der Laan, *Mitochondrial protein import: common principles and physiological networks.* Biochim Biophys Acta, 2013. **1833**(2): p. 274-85.
92. Endo, T. and D. Kohda, *Functions of outer membrane receptors in mitochondrial protein import.* Biochim Biophys Acta, 2002. **1592**(1): p. 3-14.
93. Truscott, K.N., et al., *Mitochondrial import of the ADP/ATP carrier: the essential TIM complex of the intermembrane space is required for precursor release from the TOM complex.* Mol Cell Biol, 2002. **22**(22): p. 7780-9.

94. Paschen, S.A., W. Neupert, and D. Rapaport, *Biogenesis of beta-barrel membrane proteins of mitochondria.* Trends Biochem Sci, 2005. **30**(10): p. 575-82.
95. Sideris, D.P., et al., *A novel intermembrane space-targeting signal docks cysteines onto Mia40 during mitochondrial oxidative folding.* J Cell Biol, 2009. **187**(7): p. 1007-22.
96. Bonnefoy, N., et al., *Roles of Oxa1-related inner-membrane translocases in assembly of respiratory chain complexes.* Biochim Biophys Acta, 2009. **1793**(1): p. 60-70.
97. Freitas, N. and C. Cunha, *Mechanisms and signals for the nuclear import of proteins.* Curr Genomics, 2009. **10**(8): p. 550-7.
98. Suntharalingam, M. and S.R. Wente, *Peering through the pore: nuclear pore complex structure, assembly, and function.* Dev Cell, 2003. **4**(6): p. 775-89.
99. Paine, P.L., L.C. Moore, and S.B. Horowitz, *Nuclear envelope permeability.* Nature, 1975. **254**(5496): p. 109-14.
100. Fahrenkrog, B. and U. Aebi, *The nuclear pore complex: nucleocytoplasmic transport and beyond.* Nat Rev Mol Cell Biol, 2003. **4**(10): p. 757-66.
101. Cronshaw, J.M., et al., *Proteomic analysis of the mammalian nuclear pore complex.* J Cell Biol, 2002. **158**(5): p. 915-27.
102. Mosammaparast, N. and L.F. Pemberton, *Karyopherins: from nuclear-transport mediators to nuclear-function regulators.* Trends Cell Biol, 2004. **14**(10): p. 547-56.
103. Lange, A., et al., *Classical nuclear localization signals: definition, function, and interaction with importin alpha.* J Biol Chem, 2007. **282**(8): p. 5101-5.
104. Poon, I.K. and D.A. Jans, *Regulation of nuclear transport: central role in development and transformation?* Traffic, 2005. **6**(3): p. 173-86.
105. Chook, Y.M. and G. Blobel, *Karyopherins and nuclear import.* Curr Opin Struct Biol, 2001. **11**(6): p. 703-15.
106. Gorlich, D., et al., *A 41 amino acid motif in importin-alpha confers binding to importin-beta and hence transit into the nucleus.* EMBO J, 1996. **15**(8): p. 1810-7.
107. Weis, K., U. Ryder, and A.I. Lamond, *The conserved amino-terminal domain of hSRP1 alpha is essential for nuclear protein import.* EMBO J, 1996. **15**(8): p. 1818-25.
108. Weis, K., *Regulating access to the genome: nucleocytoplasmic transport throughout the cell cycle.* Cell, 2003. **112**(4): p. 441-51.
109. Stewart, M., *Molecular mechanism of the nuclear protein import cycle.* Nat Rev Mol Cell Biol, 2007. **8**(3): p. 195-208.
110. Aggarwal, A. and D.K. Agrawal, *Importins and exportins regulating allergic immune responses.* Mediators Inflamm, 2014. **2014**: p. 476357.
111. Karimi-Busheri, F., et al., *Molecular characterization of a human DNA kinase.* J Biol Chem, 1999. **274**(34): p. 24187-94.
112. Karimi-Busheri, F., et al., *Repair of DNA strand gaps and nicks containing 3′-phosphate and 5′-hydroxyl termini by purified mammalian enzymes.* Nucleic Acids Res, 1998. **26**(19): p. 4395-400.

113. Aravind, L. and E.V. Koonin, *The HD domain defines a new superfamily of metal-dependent phosphohydrolases.* Trends Biochem Sci, 1998. **23**(12): p. 469-72.
114. Bernstein, N.K., et al., *Polynucleotide kinase as a potential target for enhancing cytotoxicity by ionizing radiation and topoisomerase I inhibitors.* Anticancer Agents Med Chem, 2008. **8**(4): p. 358-67.
115. Dobson, C.J. and S.L. Allinson, *The phosphatase activity of mammalian polynucleotide kinase takes precedence over its kinase activity in repair of single strand breaks.* Nucleic Acids Res, 2006. **34**(8): p. 2230-7.
116. Caldecott, K.W., *Mammalian DNA single-strand break repair: an X-ra(y)ted affair.* Bioessays, 2001. **23**(5): p. 447-55.
117. Reynolds, J.J., et al., *Impact of PNKP mutations associated with microcephaly, seizures and developmental delay on enzyme activity and DNA strand break repair.* Nucleic Acids Res, 2012. **40**(14): p. 6608-19.
118. McCullough, A.K., M.L. Dodson, and R.S. Lloyd, *Initiation of base excision repair: glycosylase mechanisms and structures.* Annu Rev Biochem, 1999. **68**: p. 255-85.
119. Interthal, H., J.J. Pouliot, and J.J. Champoux, *The tyrosyl-DNA phosphodiesterase Tdp1 is a member of the phospholipase D superfamily.* Proc Natl Acad Sci U S A, 2001. **98**(21): p. 12009-14.
120. Shen, J., et al., *Mutations in PNKP cause microcephaly, seizures and defects in DNA repair.* Nat Genet, 2010. **42**(3): p. 245-9.
121. Megnin-Chanet, F., M.A. Bollet, and J. Hall, *Targeting poly(ADP-ribose) polymerase activity for cancer therapy.* Cell Mol Life Sci, 2010. **67**(21): p. 3649-62.
122. Freschauf, G.K., et al., *Identification of a small molecule inhibitor of the human DNA repair enzyme polynucleotide kinase/phosphatase.* Cancer Res, 2009. **69**(19): p. 7739-46.
123. Dobzhansky, T., *Genetics of Natural Populations. Xiii. Recombination and Variability in Populations of Drosophila Pseudoobscura.* Genetics, 1946. **31**(3): p. 269-90.
124. Bryant, H.E., et al., *Specific killing of BRCA2-deficient tumours with inhibitors of poly(ADP-ribose) polymerase.* Nature, 2005. **434**(7035): p. 913-7.
125. Farmer, H., et al., *Targeting the DNA repair defect in BRCA mutant cells as a therapeutic strategy.* Nature, 2005. **434**(7035): p. 917-21.
126. McLornan, D.P., A. List, and G.J. Mufti, *Applying synthetic lethality for the selective targeting of cancer.* N Engl J Med, 2014. **371**(18): p. 1725-35.
127. Shaheen, M., et al., *Synthetic lethality: exploiting the addiction of cancer to DNA repair.* Blood, 2011. **117**(23): p. 6074-82.
128. Mereniuk, T.R., et al., *Synthetic lethal targeting of PTEN-deficient cancer cells using selective disruption of polynucleotide kinase/phosphatase.* Mol Cancer Ther, 2013. **12**(10): p. 2135-44.
129. Mereniuk, T.R., et al., *Genetic screening for synthetic lethal partners of polynucleotide kinase/phosphatase: potential for targeting SHP-1-depleted cancers.* Cancer Res, 2012. **72**(22): p. 5934-44.

130. Wu, C., et al., *SHP-1 suppresses cancer cell growth by promoting degradation of JAK kinases.* J Cell Biochem, 2003. **90**(5): p. 1026-37.
131. Yin, Y. and W.H. Shen, *PTEN: a new guardian of the genome.* Oncogene, 2008. **27**(41): p. 5443-53.
132. Lindahl, T. and B. Nyberg, *Rate of depurination of native deoxyribonucleic acid.* Biochemistry, 1972. **11**(19): p. 3610-8.
133. Santos, J.H., et al., *Quantitative PCR-based measurement of nuclear and mitochondrial DNA damage and repair in mammalian cells.* Methods Mol Biol, 2006. **314**: p. 183-99.
134. Fanta, M., et al., *Production, characterization, and epitope mapping of monoclonal antibodies against human polydeoxyribonucleotide kinase.* Hybridoma, 2001. **20**(4): p. 237-42.
135. Dingwall, C. and R.A. Laskey, *Nuclear targeting sequences--a consensus?* Trends Biochem Sci, 1991. **16**(12): p. 478-81.
136. van Royen, M.E., et al., *Fluorescence recovery after photobleaching (FRAP) to study nuclear protein dynamics in living cells.* Methods Mol Biol, 2009. **464**: p. 363-85.
137. Melan, M.A. and G. Sluder, *Redistribution and differential extraction of soluble proteins in permeabilized cultured cells. Implications for immunofluorescence microscopy.* J Cell Sci, 1992. **101 (Pt 4)**: p. 731-43.
138. Shi, S.R., R.J. Cote, and C.R. Taylor, *Antigen retrieval immunohistochemistry: past, present, and future.* J Histochem Cytochem, 1997. **45**(3): p. 327-43.
139. Hoetelmans, R.W., et al., *Effects of acetone, methanol, or paraformaldehyde on cellular structure, visualized by reflection contrast microscopy and transmission and scanning electron microscopy.* Appl Immunohistochem Mol Morphol, 2001. **9**(4): p. 346-51.
140. de Souza-Pinto, N.C., et al., *Mitochondrial DNA, base excision repair and neurodegeneration.* DNA Repair (Amst), 2008. **7**(7): p. 1098-109.
141. Liu, P. and B. Demple, *DNA repair in mammalian mitochondria: Much more than we thought?* Environ Mol Mutagen, 2010. **51**(5): p. 417-26.
142. Das, B.B., et al., *Role of tyrosyl-DNA phosphodiesterase (TDP1) in mitochondria.* Proc Natl Acad Sci U S A, 2010. **107**(46): p. 19790-5.
143. Rossi, M.N., et al., *Mitochondrial localization of PARP-1 requires interaction with mitofilin and is involved in the maintenance of mitochondrial DNA integrity.* J Biol Chem, 2009. **284**(46): p. 31616-24.
144. Sykora, P., et al., *Aprataxin localizes to mitochondria and preserves mitochondrial function.* Proc Natl Acad Sci U S A, 2011. **108**(18): p. 7437-42.
145. Bottero, V., et al., *Ikappa b-alpha, the NF-kappa B inhibitory subunit, interacts with ANT, the mitochondrial ATP/ADP translocator.* J Biol Chem, 2001. **276**(24): p. 21317-24.
146. De Bosscher, K., W. Vanden Berghe, and G. Haegeman, *The interplay between the glucocorticoid receptor and nuclear factor-kappaB or activator protein-1: molecular mechanisms for gene repression.* Endocr Rev, 2003. **24**(4): p. 488-522.
147. Marchenko, N.D., A. Zaika, and U.M. Moll, *Death signal-induced localization of p53 protein to mitochondria. A potential role in apoptotic signaling.* J Biol Chem, 2000. **275**(21): p. 16202-12.

148. Psarra, A.M. and C.E. Sekeris, *Nuclear receptors and other nuclear transcription factors in mitochondria: regulatory molecules in a new environment.* Biochim Biophys Acta, 2008. **1783**(1): p. 1-11.
149. Chattopadhyay, R., et al., *Identification and characterization of mitochondrial abasic (AP)-endonuclease in mammalian cells.* Nucleic Acids Res, 2006. **34**(7): p. 2067-76.
150. Li, M., et al., *Identification and characterization of mitochondrial targeting sequence of human apurinic/apyrimidinic endonuclease 1.* J Biol Chem, 2010. **285**(20): p. 14871-81.
151. Lee, C.M., et al., *The DNA helicase, Hmi1p, is transported into mitochondria by a C-terminal cleavable targeting signal.* J Biol Chem, 1999. **274**(30): p. 20937-42.
152. Zheng, L., et al., *Human DNA2 is a mitochondrial nuclease/helicase for efficient processing of DNA replication and repair intermediates.* Mol Cell, 2008. **32**(3): p. 325-36.
153. Sudbeck, P. and G. Scherer, *Two independent nuclear localization signals are present in the DNA-binding high-mobility group domains of SRY and SOX9.* J Biol Chem, 1997. **272**(44): p. 27848-52.
154. Dingwall, C., et al., *The nucleoplasmin nuclear location sequence is larger and more complex than that of SV-40 large T antigen.* J Cell Biol, 1988. **107**(3): p. 841-9.
155. Lee, M.S., et al., *Identification of a nuclear localization signal in the polo box domain of Plk1.* Biochim Biophys Acta, 2009. **1793**(10): p. 1571-8.
156. Rasouli-Nia, A., F. Karimi-Busheri, and M. Weinfeld, *Stable down-regulation of human polynucleotide kinase enhances spontaneous mutation frequency and sensitizes cells to genotoxic agents.* Proc Natl Acad Sci U S A, 2004. **101**(18): p. 6905-10.
157. Milakovic, T., et al., *Intracellular localization and activity state of tissue transglutaminase differentially impacts cell death.* J Biol Chem, 2004. **279**(10): p. 8715-22.
158. Gomel, R., et al., *The localization of protein kinase Cdelta in different subcellular sites affects its proapoptotic and antiapoptotic functions and the activation of distinct downstream signaling pathways.* Mol Cancer Res, 2007. **5**(6): p. 627-39.
159. Choe, C.U., et al., *Functional coupling of chromogranin with the inositol 1,4,5-trisphosphate receptor shapes calcium signaling.* J Biol Chem, 2004. **279**(34): p. 35551-6.
160. Takeuchi, H., et al., *Mitochondrial localization of mutant superoxide dismutase 1 triggers caspase-dependent cell death in a cellular model of familial amyotrophic lateral sclerosis.* J Biol Chem, 2002. **277**(52): p. 50966-72.
161. Mojica, F.J., et al., *Long stretches of short tandem repeats are present in the largest replicons of the Archaea Haloferax mediterranei and Haloferax volcanii and could be involved in replicon partitioning.* Mol Microbiol, 1995. **17**(1): p. 85-93.
162. Masepohl, B., K. Gorlitz, and H. Bohme, *Long tandemly repeated repetitive (LTRR) sequences in the filamentous cyanobacterium Anabaena sp. PCC 7120.* Biochim Biophys Acta, 1996. **1307**(1): p. 26-30.

163. Jansen, R., et al., *Identification of genes that are associated with DNA repeats in prokaryotes.* Mol Microbiol, 2002. **43**(6): p. 1565-75.
164. Barrangou, R., et al., *CRISPR provides acquired resistance against viruses in prokaryotes.* Science, 2007. **315**(5819): p. 1709-12.
165. Deltcheva, E., et al., *CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III.* Nature, 2011. **471**(7340): p. 602-7.
166. Gasiunas, G., et al., *Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria.* Proc Natl Acad Sci U S A, 2012. **109**(39): p. E2579-86.
167. Jinek, M., et al., *A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity.* Science, 2012. **337**(6096): p. 816-21.
168. Sapranauskas, R., et al., *The Streptococcus thermophilus CRISPR/Cas system provides immunity in Escherichia coli.* Nucleic Acids Res, 2011. **39**(21): p. 9275-82.
169. Horvath, P. and R. Barrangou, *CRISPR/Cas, the immune system of bacteria and archaea.* Science, 2010. **327**(5962): p. 167-70.
170. Deveau, H., et al., *Phage response to CRISPR-encoded resistance in Streptococcus thermophilus.* J Bacteriol, 2008. **190**(4): p. 1390-400.
171. Marraffini, L.A. and E.J. Sontheimer, *Self versus non-self discrimination during CRISPR RNA-directed immunity.* Nature, 2010. **463**(7280): p. 568-71.
172. Mali, P., et al., *RNA-guided human genome engineering via Cas9.* Science, 2013. **339**(6121): p. 823-6.
173. Cong, L., et al., *Multiplex genome engineering using CRISPR/Cas systems.* Science, 2013. **339**(6121): p. 819-23.
174. Hsu, P.D., E.S. Lander, and F. Zhang, *Development and applications of CRISPR-Cas9 for genome engineering.* Cell, 2014. **157**(6): p. 1262-78.
175. Jinek, M., et al., *RNA-programmed genome editing in human cells.* Elife, 2013. **2**: p. e00471.
176. Ran, F.A., et al., *Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity.* Cell, 2013. **154**(6): p. 1380-9.
177. Qi, L.S., et al., *Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression.* Cell, 2013. **152**(5): p. 1173-83.
178. Cho, S.W., et al., *Analysis of off-target effects of CRISPR/Cas-derived RNA-guided endonucleases and nickases.* Genome Res, 2014. **24**(1): p. 132-41.
179. El Gendy, M.A., et al., *Induction of quinone oxidoreductase 1 enzyme by Rhazya stricta through Nrf2-dependent mechanism.* J Ethnopharmacol, 2012. **144**(2): p. 416-24.
180. Livak, K.J. and T.D. Schmittgen, *Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method.* Methods, 2001. **25**(4): p. 402-8.
181. Roy, A., et al., *Generation of WNK1 knockout cell lines by CRISPR/Cas-mediated genome editing.* Am J Physiol Renal Physiol, 2015. **308**(4): p. F366-76.
182. Cho, S.W., et al., *Targeted genome engineering in human cells with the Cas9 RNA-guided endonuclease.* Nat Biotechnol, 2013. **31**(3): p. 230-2.
183. Auer, T.O., et al., *Highly efficient CRISPR/Cas9-mediated knock-in in zebrafish by homology-independent DNA repair.* Genome Res, 2014. **24**(1): p. 142-53.

184. Zhang, H., et al., *Human mitochondrial topoisomerase I.* Proc Natl Acad Sci U S A, 2001. **98**(19): p. 10608-13.
185. Sykora, P., D.M. Wilson, 3rd, and V.A. Bohr, *Repair of persistent strand breaks in the mitochondrial genome.* Mech Ageing Dev, 2012. **133**(4): p. 169-75.
186. Kam, W.W. and R.B. Banati, *Effects of ionizing radiation on mitochondria.* Free Radic Biol Med, 2013. **65**: p. 607-19.

Appendix-A

A1. Commonly Used Buffer Recipes

DMEM/F12 Medium

- 0.6g $NaHOC_3$ (or 8mL of 7.5% w/v solution)
- 5mL L-Glutamine (10x)
- 5mL MEM Non-Essential Amino Acids (10x)
- 5mL Sodium Pyruvate (10x)
- 5mL Penicillin and Streptomycin (10x)
- 50mL Fetal Bovine Serum (FBS)

Freezing Medium:

- 95% DMEM/F12
- 5% DMSO

SDS-PAGE (10% Resolving Gel)

- 2.5mL 40% Acrylamide/Bis
- 2.5mL 1.5 M Tris-HCl, pH 8.8
- 100µL 10% SDS
- 4.9mL ddH_20
- 5µL TEMED
- 50µL 10% Ammonium Persulfate (APS)

SDS-PAGE (4% Stacking Gel)

- 250µL 40% Acrylamide/Bis
- 630µL 0.5 M Tris-HCl, pH 6.8
- 25µL 10% SDS
- 1.6mL ddH_20
- 2.5µL TEMED

- 15μL 10% Ammonium Persulfate

10X Running Buffer with SDS

- 25 mM Tris
- 192 mM glycine
- 0.1% SDS
- pH to 8.3

Transfer Buffer

- 25mM Tris-HCl, pH 8.3
- 192mM glycine
- 10% methanol

1x Sample Buffer

- 50μL β-mercaptoethanol
- 950μL Lamelli Buffer

LB Broth:

- 10g Bacto-Tryptone
- 10g NaCl
- 5g Bacto-Yeast Extract
- 950mL ddH_20
- Raise pH to 7.00 with 1 tablet of NaOH

LB + Antibiotic Plates:

- 5g NaCl
- 5g Tryptone
- 5g Yeast Extract
- 10g Agar
- 500mL ddH_20
- 1:1000 Antibiotics
- pH to 7.4

RIPA Buffer:

- 50mM Tris-HCl, pH 7.6
- 150mM NaCl
- 1% deoxycholate
- 1% TritonX-100
- 1mM Na_3VO_4
- 50mM NaF
- 1mM phenylmethylsulfonyl fluoride

10x PBS Buffer

- 10.9g Na_2HPO_4
- 3.2g KH_2PO_4
- 90g NaCl
- 1L ddH_2O
- pH to 7.2 with NaOH

PNKP	
Primer	**Sequence**
Phos-R	CTGGTTGGTGAAGATCACCAG
MWPP4	GCCTTTGATCCGAGGACTGTC
MWPP5	TCAGCCCTCGGAGAACTGG
FOR20	TTAACCCCTCAACTACCGGG
292-R	CATTGACCAAATACAGTGTG
622+	TACAAGCTGGTGATCTTC
176+	ACAGTGGCAGTGAAACAGC
RNAi-R	TCAGCCCTCGGAGAACTGG
Cas9	
Primer	**Sequence**
Cas9- 250-R	TGCAGGTAGCAGATCCGATTC
Cas9-180-R	TCGGCCGTCTCCCCGGAGTCG
Cas9- 2318-F	GACAGAAGAACAGTAGGGAAA
Cas9-2718-F	CCTCGTTCAGCCTTAGTCAGA
MTS-mutant	
Primer	**Sequence**
Cmtsmut-R	AGAGAAGCCTTCAGCCAGC
Cmtsmut-F	AATGTCGTAACAACTCCGCC
LentiCRISPRv2-PNKP	
Primer	**Sequence**
lentiCRISPRv2-2630seq-F	CTTCATATTTGCATATACGA

Table 1. Primers used for sequencing.

G-block - F
TGTACAAAAAAGCAGGCTTTAAAGGAACCAATTCAGTCGACTGGATCCGGTACCAAG GTCGGGCAGGAAGAGGGCCTATTTCCCATGATTCCTTCATATTTGCATATACGATACA AGGCTGTTAGAGAGATAATTAGAATTAATTTGACTGTAAACACAAAGATATTAGTAC AAAATACGTGACGTAGAAAGTAATAATTTCTTGGGTAGTTTGCAGTTTTAAAATTAT GTTTTAAAATGGACTATCATATGCTTACCGTAACTTGAAAGTATTTCGATTTCTTGGC TTTATATATCTTGTGGAAAGGACGAAACACCGGCAAGCCCTGGTCCTGGGCAGTTTT AGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTG GCACCGAGTCGGTGCTTTTTTTCTAGACCCAGCTTTCTTGTACAAAGTTGGCATTA
G-block - RC
TGTACAAAAAAGCAGGCTTTAAAGGAACCAATTCAGTCGACTGGATCCGGTACCAAG GTCGGGCAGGAAGAGGGCCTATTTCCCATGATTCCTTCATATTTGCATATACGATACA AGGCTGTTAGAGAGATAATTAGAATTAATTTGACTGTAAACACAAAGATATTAGTAC AAAATACGTGACGTAGAAAGTAATAATTTCTTGGGTAGTTTGCAGTTTTAAAATTAT GTTTTAAAATGGACTATCATATGCTTACCGTAACTTGAAAGTATTTCGATTTCTTGGC TTTATATATCTTGTGGAAAGGACGAAACACCGAGCACTTCCGGTCCGTAACCGTTTTA GAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGC ACCGAGTCGGTGCTTTTTTTCTAGACCCAGCTTTCTTGTACAAAGTTGGCATTA
G-block - 3
TGTACAAAAAAGCAGGCTTTAAAGGAACCAATTCAGTCGACTGGATCCGGTACCAAG GTCGGGCAGGAAGAGGGCCTATTTCCCATGATTCCTTCATATTTGCATATACGATACA AGGCTGTTAGAGAGATAATTAGAATTAATTTGACTGTAAACACAAAGATATTAGTAC AAAATACGTGACGTAGAAAGTAATAATTTCTTGGGTAGTTTGCAGTTTTAAAATTAT GTTTTAAAATGGACTATCATATGCTTACCGTAACTTGAAAGTATTTCGATTTCTTGGC TTTATATATCTTGTGGAAAGGACGAAACACC**G**CACTGGCCCCAGTGACTGGGTTTTA GAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGC ACCGAGTCGGTGCTTTTTTTCTAGACCCAGCTTTCTTGTACAAAGTTGGCATTA
G-block - 7
TGTACAAAAAAGCAGGCTTTAAAGGAACCAATTCAGTCGACTGGATCCGGTACCAAG GTCGGGCAGGAAGAGGGCCTATTTCCCATGATTCCTTCATATTTGCATATACGATACA AGGCTGTTAGAGAGATAATTAGAATTAATTTGACTGTAAACACAAAGATATTAGTAC AAAATACGTGACGTAGAAAGTAATAATTTCTTGGGTAGTTTGCAGTTTTAAAATTAT GTTTTAAAATGGACTATCATATGCTTACCGTAACTTGAAAGTATTTCGATTTCTTGGC TTTATATATCTTGTGGAAAGGACGAAACACC**G**GAGGCTGTGGTGGAGAAGCGTTTTA GAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGC ACCGAGTCGGTGCTTTTTTTCTAGACCCAGCTTTCTTGTACAAAGTTGGCATTA
G-block - Scramble
TGTACAAAAAAGCAGGCTTTAAAGGAACCAATTCAGTCGACTGGATCCGGTACCAAG GTCGGGCAGGAAGAGGGCCTATTTCCCATGATTCCTTCATATTTGCATATACGATACA AGGCTGTTAGAGAGATAATTAGAATTAATTTGACTGTAAACACAAAGATATTAGTAC AAAATACGTGACGTAGAAAGTAATAATTTCTTGGGTAGTTTGCAGTTTTAAAATTAT GTTTTAAAATGGACTATCATATGCTTACCGTAACTTGAAAGTATTTCGATTTCTTGGC TTTATATATCTTGTGGAAAGGACGAAACACC**G**TAGATATCTATTTGCACGTGTTTTAG AGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCA CCGAGTCGGTGCTTTTTTTCTAGACCCAGCTTTCTTGTACAAAGTTGGCATTA

Table 2. The G-block sequences.

Antibody	Source	Poly/Monoclonal	Manufacturer	Catalog Number
PNKP 122	Rabbit	Polyclonal	Weinfeld Lab	
PNKP (H101)	Mouse	Monoclonal	Weinfeld Lab	
PNKP	Rabbit	Polyclonal	Sigma	HPA006782
Actin	Goat	Polyclonal	Santa Cruz	sc-1616
Cas9	Mouse	Monoclonal	Active Motif	61578
Mitofilin	Mouse	Monoclonal	MitoSciences	MSM02
VDAC	Rabbit	Polyclonal	Abcam	ab15895

Table 3. Antibodies.

Appendix-B

3484–3495 *Nucleic Acids Research, 2012, Vol. 40, No. 8* *Published online 30 December 2011*
doi:10.1093/nar/gkr1245

Role of polynucleotide kinase/phosphatase in mitochondrial DNA repair

Nasser Tahbaz[1,2], Sudip Subedi[1,2] and Michael Weinfeld[1,2,*]

[1]Department of Oncology, University of Alberta and [2]Experimental Oncology, Cross Cancer Institute, 11560 University Avenue, Edmonton, Alberta, Canada, T6G 1Z2

Received July 22, 2011; Revised November 30, 2011; Accepted December 1, 2011

ABSTRACT

Mutations in mitochondrial DNA (mtDNA) are implicated in a broad range of human diseases and in aging. Compared to nuclear DNA, mtDNA is more highly exposed to oxidative damage due to its proximity to the respiratory chain and the lack of protection afforded by chromatin-associated proteins. While repair of oxidative damage to the bases in mtDNA through the base excision repair pathway has been well studied, the repair of oxidatively induced strand breaks in mtDNA has been less thoroughly examined. Polynucleotide kinase/phosphatase (PNKP) processes strand-break termini to render them chemically compatible for the subsequent action of DNA polymerases and ligases. Here, we demonstrate that functionally active full-length PNKP is present in mitochondria as well as nuclei. Downregulation of PNKP results in an accumulation of strand breaks in mtDNA of hydrogen peroxide-treated cells. Full restoration of repair of the H_2O_2-induced strand breaks in mitochondria requires both the kinase and phosphatase activities of PNKP. We also demonstrate that PNKP contains a mitochondrial-targeting signal close to the C-terminus of the protein. We further show that PNKP associates with the mitochondrial protein mitofilin. Interaction with mitofilin may serve to translocate PNKP into mitochondria.

INTRODUCTION

Single- and double-strand DNA breaks are induced directly by external and internal genotoxic agents such as ionizing radiation (IR), UV light and reactive oxygen species (ROS), or indirectly as a result of aborted topoisomerase action or during base excision repair (BER). Radiation and ROS-induced strand breaks frequently bear strand-break termini that require processing before ligation can occur, including 3′-phosphate and phosphoglycolate and 5′-hydroxyl end groups (1–3). Similarly, trapping of topoisomerase 1 by agents such as camptothecin or the presence of abasic sites or nicks adjacent to the cleavage site, followed by tyrosyl-DNA phosphodiesterase 1 (TDP1)-mediated cleavage of the covalent bond linking the DNA to the topoisomerase, generates single-strand breaks with 3′-phosphate and 5′-hydroxyl termini (4,5). BER performed by bifunctional DNA glycosylases, such as the Nei family members, NEIL1, NEIL2 and NEIL3, remove damaged bases and then cleave the DNA at the abasic sites through a lyase activity that involves β,δ-elimination to generate 3′-phosphate termini (6–8). The damaged DNA termini have to be restored to 3′-hydroxyl and 5′-phosphate functionality prior to the completion of the repair process by DNA polymerases and DNA ligases.

Polynucleotide kinase/phosphatase (PNKP) plays a major role in the restoration of correct DNA termini following strand cleavage by IR, ROS or NEIL-dependent BER (3,7,9,10). PNKP contains a forkhead-associated domain, which is a protein–protein interaction domain required for the association of PNKP with CK2-phosphorylated XRCC1 and XRCC4 (11–14), and independent DNA 3′-phosphatase and 5′-kinase domains (15,16). It has been shown that the DNA 3′-phosphatase activity of PNKP takes precedence over its DNA 5′-kinase activity *in vitro* (17). Downregulation of PNKP sensitizes cells to IR and hydrogen peroxide (18,19).

In addition to damage to nuclear DNA, mitochondrial DNA (mtDNA) is also subject to DNA damage. MtDNA is a 16.5 kbp circular molecule, encoding 37 genes including 13 proteins, 22 tRNAs and 2 rRNAs. Eukaryotic cells can have more than 100 mitochondria and each mitochondrion may contain 10 mtDNAs. In general, mtDNA constitutes about 1% of the total cellular DNA. ROS produced in relatively large quantities in mitochondria during respiration are the major source of mtDNA lesions (20). Damage to mtDNA, if not repaired, can develop into mutations, and mutations of the mtDNA are associated with different diseases including diabetes (21,22), cancer (23), neurodegenerative disorders

*To whom correspondence should be addressed. Tel: +1 780 432 8438; Fax: +1 780 432 8428; Email: michael.weinfeld@albertahealthservices.ca

(24) and aging (25). The rate of mutations in some regions of mtDNA is 20- to 100-fold higher than the nuclear DNA (26). This could be explained by the lack of protection of mtDNA by chromosomal proteins and the proximity of mtDNA to the inner membrane that contains the electron transport chain, which is a constant source of ROS (27).

As in the nucleus, BER is the main DNA repair pathway in mitochondria that deals with ROS-induced DNA lesions (8,27,28). Several DNA glycosylases have been identified in mitochondria including Nth and Nei family members (27). Mitochondria also contain a truncated form of APE-1 that can process abasic sites and DNA ends produced by β-elimination by DNA glycosylases/lyases such as NTH1 (29). In mitochondrial BER, replacement of missing nucleotides at damaged sites is performed by DNA polymerase γ (Polγ) instead of DNA polymerase β found in the nucleus (30), and ligation of DNA at single strand breaks is mediated by DNA ligase III (31). Recent studies have highlighted the importance of mtDNA ligase III for cell survival (31). Topoisomerase 1 and TDP1 are also present and functional in mitochondria (32). The presence of ROS, an active BER pathway and topoisomerase 1/TDP1 pathway point strongly to a need for PNKP or similarly acting protein to correct strand break termini in mitochondria.

Here, we demonstrate that functionally active full-length PNKP is present in mitochondria. Downregulation of PNKP results in a decrease of both mitochondrial and nuclear PNKP and accumulation of DNA damage in mtDNA, in addition to the previously documented increase in the nuclear DNA damage (18). Furthermore, we demonstrate that PNKP contains a C-terminal mitochondrial-targeting signal (MTS) (33,34). This C-terminal MTS is functional and is required for the localization of PNKP to mitochondria. Our results also indicate that PNKP associates with the mitochondrial protein mitofilin.

MATERIALS AND METHODS

Cell culture and transfection

A549 cells (human lung adenocarcinoma) and MCF7 cells (breast carcinoma) were obtained from ATCC (Manassas, VA, USA). PNKP-depleted cells were prepared by stable transfection of A549 cells with a pSuper expression vector containing an shRNA directed against PNKP as previously described (18). Cells were grown in DMEM/F12 nutrient mixture (1 : 1) and supplemented with 10% FBS, 50 U ml^{-1} penicillin, 50 μg ml^{-1} streptomycin, 2 mM L-glutamine, 0.1 mM non-essential amino acids and 1 mM sodium pyruvate (Invitrogen, Burlington, ON, USA) in a 5% CO_2 humidified incubator at 37°C.

Cells were plated at ~80–90% confluency in normal media without antibiotics and Lipofectamine 2000 was used the next day for transient transfections, according to the manufacturer's instructions (Invitrogen).

Immunoprecipitations and western blots

Cells were harvested from >90% confluent dishes by trypsinization, washed in cold PBS, pH 7.4. Cell pellets were resuspended in IP lysis buffer (200 mM NaCl, 2.5 mM $MgCl_2$, 20 mM Tris–HCl pH 7.4 and 0.05% NP40) or lysis buffer (150 mM NaCl, 2 mM EDTA, 50 mM Tris–HCl, pH 7.4 and 1% NP40) plus 2 mM DTT, 1 mM 4-(2-aminoethyl)benzenesulfonyl fluoride hydrochloride (Sigma, St Louis, MO, USA), 1× complete protease inhibitor cocktail (Roche, Laval, QC, USA), followed by incubation on ice for 30–60 min. Next, cells were sonicated three times briefly in their corresponding buffer, centrifuged at 20 000*g* for 10 min at 4°C and the supernatants were snap-frozen and transferred to −80°C freezer for later use for either immunoprecipitation or western blots.

Cell lysates were precleared prior to immunoprecipitation by addition to Protein A-Sepharose beads (Sigma) and rotation at 4°C for 1 h followed by a spin to discard the beads. Then antibodies were added to the precleared lysates and rotated at 4°C for 1–3 h. Next, fresh beads were added to the lysate and rotation continued overnight. Finally, beads (now associated with antibodies) were harvested and washed three times with the lysis buffer, boiled in sample loading buffer and protein samples were run on a 10% PAGE gel. Samples were treated exactly the same for bead controls, except that no antibody was used. For immunoprecipitation of HA-tagged PNKP (HAPNKP), anti-HA affinity matrix (Roche) was used following the manufacturer's recommendations. Western blots were performed using an ECL kit (Roche). Western blot images were scanned and quantified using ImageQuant for Windows version 5.2. The reading from the PNKP knock-down (KD) cell line in each experiment was normalized to the reading from A549 control cells. The *P* value was calculated using a two-tailed student's *t* test.

Antibodies used included polyclonal and monoclonal antibodies against PNKP (35), monoclonal mitofilin antibody (cat. # MSM02, Mitosciences, Eugene, OR), PCNA (cat. # 9857) and actin (cat. # SC1616, Santa Cruz Biotechnology, Santa Cruz, CA), HA (cat. # 12013819001, Roche, Brampton, ON), COX IV (ab16056-100) and VDAC1 (ab15895-100, Abcam, San Francisco, CA).

Generation of constructs and site-directed mutagenesis

RNAi-resistant PNKP was prepared by mutating PNKP cDNA at the shRNA targeting sequence, using the following primers:

Forward 5′-CAACCGGTTTCGAGAAATGACCGA TTCCTCTC ATATCCCCG-3′
Reverse 5′-CGGGGATATGAGAGGAATCGGTCAT TTC TCGAAACCGGTTG-3′.

The kinase negative PNKP was generated by site-directed mutagenesis using RNAi-resistant PNKP cDNA and the following primers:

Forward 5′-GGGATTCCCTGGGGCCGGGGCCTC CACCTTTCTCAAGAAGC-3′

Reverse 5′-GGGATTCCCTGGGGCCGGGGCCTCC ACCTTTCTCAAGAAGC-3′

The phosphatase negative PNKP was generated by site-directed mutagenesis using RNAi-resistant PNKP cDNA and the following primers:

Forward 5′-AAGGTGGCTGGCTTTAATCTGAAC GGGACGCTCATCACC-3′
Reverse 5′-GGTGATGAGCGTCCCGTTCAGATTA AAGCCAGCCACCTT-3′

The RNAi-resistant PNKP bearing a C-terminal haemagglutinin epitope tag (PNKPHA) was generated using RNAi-resistant PNKP cDNA and the following primers:

Forward 5′-TGACTGAATTCGCACCCAGGATGG GCGAGGTGGAGCCC-3′
Reverse 5′-TGACTGCGGCCGCTCAAGCGTAATC TGGAACATCGTATGGGTACGGGAGCCTCTTGA CCGTC-3′.

The RNAi-resistant amino terminally-tagged HAPNKP was generated using RNAi-resistant PNKP cDNA and the following primers:

Forward 5′-TGACTGAATTCGCACCCAGGATGTA CCCATACGATGTTCCAGATTACG CTGGCGAGG TGGAGCCCCCGG-3′
Reverse 5′-TGACT GCGGCCGCTCAGCCCTCGGA GAA CTG GC-3′.

The RNAi-resistant HAPNKP mutated at the mitochondrial-targeting signal (HAPNKP-mts) was generated by site-directed mutagenesis using cDNA from RNAi-resistant HAPNKP and the following primers:

Forward 5′-AACCCAGACGCCGCGAGCCGCGAC GGGGACGTCCAGTGTGCCCGAGCC-3′
Reverse 5′- GGCTCGGGCACACTGGACGTCCCCG TCGCGGCTCGCGGCGTCTGGGTT-3′.

All of the above constructs were prepared in pIRESpuro3 from Clontech (Mountain View, CA, USA). Stable cell lines were prepared using the aforementioned cDNA constructs for transfection, followed by addition of 1 mM puromycin to the media and selection of surviving clones.

The constructs composed of the C-terminal region of PNKP, with and without the putative mitochondrial-targeting sequence, fused to GFP (CmtsPNKP + GFP and the CPNKP + GFP, respectively) were prepared using the pmEGFPN1 vector from Clontech. To prepare CmtsPNKP + GFP the following primers were used:

Forward 5′-TGACTAAGCTTGCACCCAGGATGG CCAGGTACGTCCAGTGTG-3′
Reverse 5′-GCTGATGGATCCCGGCCCTCGGAGA ACTGGCAG-3′

To prepare CPNKP + GFP we used the reverse primer from CmtsPNKP + GFP and the forward primer 5′- TGA CTAAGCTTGCACCCAGGATGGGCGTCCCCTGCC GCTG-3′

To mutate MTS in CmtsPNKP + GFP (mutCmtsPNKP + GFP) we used CmtsPNKP + GFP as a template. The site directed mutagenesis was accomplished using the following primers:

Forward 5′-GCTCAAGCTTGCACCCAGGATGGA CGGGGACGTCCAGTGTGCCCGAGCC-3′
Reverse 5′-GGCTCGGGCACACTGGACGTCCCCG TCCATCCTGGGTGCAAGCTTGAGC-3′

We used QickChange II XL kit (cat. # 200521-5, Stratagene) for all site-directed mutagenesis experiments.

Purification of mitochondria, trypsin and proteinase K treatments

Mitochondria were purified essentially as described (36), with minor modifications. Briefly, cells grown in thirty 150-mm tissue culture dishes were harvested by trypsinization at >80% confluency, resuspended in hypotonic buffer (20 mM HEPES-KOH pH 7.4, 5 mM $MgCl_2$, 5 mM KCl and 1 mM DTT) plus DTT and EDTA-free complete protease inhibitors (Roche) on ice (10 min), and homogenized in a Dounce homogenizer with 10–20 strokes. During homogenization cells were monitored by microscope to avoid break down of their nuclei. Immediately after homogenization, 2×MSH buffer (20 mM HEPES-KOH pH 7.4, 4 mM EDTA, 2 mM EGTA, 5 mM DTT, 420 mM mannitol, 140 mM sucrose) was added to the homogenate to stabilize the nuclei and samples were centrifuged two times at 1200*g* for 10 min each to prepare post-nuclear supernatant (PNS). The PNS then was centrifuged at 10 000 *g* to pull down crude mitochondria (CM). CM pellets were resuspended in 1×MSH/50% Percoll (Sigma) and loaded on top of a 1×MSH/50% Percoll column followed by ultracentrifugation at 50 000*g* for 70 min at 4°C. The white band of pure mitochondria that forms in the middle of the column was extracted by syringe, washed twice in 1×MSH and protease inhibitors, aliquoted, snap frozen and stored at −80°C. Prior to each experiment samples were thawed on ice and treated with trypsin (10 μg μl^{-1}) for 20 min at room temperature followed by trypsin inhibitor treatment (29), or proteinase K (36) as required. An Artek Sonication Dismembrator model 150 was used to sonicate mitochondrial samples briefly when needed.

For isolation of rat liver mitochondria, normal Wistar rats were sacrificed and livers were removed quickly, rinsed with MSH buffer and diced. Small pieces of rat liver were homogenized by 10 strokes in a Dounce homogenizer and spun three times at 1200*g* for 10 min to remove cell debris. PNS was then treated the same as described for the cell lines.

Kinase and phosphatase assays of mitochondrial preparations

Kinase and phosphatase assays were performed essentially as described (18). Briefly, pure mitochondrial preparations were thawed on ice, trypsin-treated, pelleted by centrifugation, sonicated or not and resuspended in buffer C (70 mM sucrose, 210 mM mannitol, 80 mM succinic acid, pH 5.5, 10 mM $MgCl_2$, 1 mM DTT and protease inhibitors) for kinase assays or in buffer D (70 mM sucrose, 210 mM mannitol, 50 mM Tris–HCl, pH 8.2, 10 mM

$MgCl_2$, 5 mM DTT, 1 mM spermidine) for phosphatase assays. A 21-mer oligonucleotide with a 5′-OH terminus and a 20-mer oligonucleotide bearing a 3′-phosphate group were used as the substrates for the kinase and phosphatase assays, respectively.

MtDNA repair assay using extra large-qPCR

H_2O_2 was added for 1 h to the media of PNKP knock down (KD) A549 cells transfected with different constructs or vector only (18). Whole-cell (genomic and mitochondrial) DNAs were extracted using a miniprep kit according to the manufacturer's instructions (Qiagen) from untreated (control) and H_2O_2-treated cells after 0, 0.5, 2 or 4 h of repair. Extra large qPCR (XL-qPCR) was performed using a Gene Amp kit (Applied Biosystems) following conditions and primers described (37). Fluorescence quantification of XL-qPCR products was achieved using Quant-iT Pico Green dsDNA assay kit (Invitrogen).

Immunofluorescence imaging

A549 and MCF7 cells were grown to 20% confluency on coverslips, fixed and stained as previously described (38). GFP constructs were transfected into A549 or MCF7 cells using Lipofectamine 2000 and cells were used for imaging within 24 h post-transfection. Fluorescently stained cells were viewed on a Zeiss LSM 710 laser scanning confocal microscope mounted on an AxioObserver Z1 inverted microscope (Carl Zeiss, Jena, Germany) with a plan Apochromat 40× (NA 1.3) oil immersion lens. Individual immunofluorescence channels were collected sequentially to avoid signal bleeding through. GFP and mitofilin (using an Alexa 488-conjugated secondary antibody) were detected using 488-nm laser excitation and collected with a bandpass filter of 493–552 nm. PNKP and COX IV were detected with a secondary antibody conjugated with rhodamine, which was imaged with 561 nm laser excitation and emission capture with a bandpass filter of 562–638 nm. DAPI-stained DNA was imaged with a 405 nm laser and a bandpass filter of 410–497 nm. Images were collected at Nyquist sample rate with pinhole of 1 Airy unit.

RESULTS

Full-length PNKP is present in mitochondria

Mitochondria were purified from rat liver and human A549 lung cancer cells by centrifugation through a Percoll gradient (36), and then subjected to limited trypsin-treatment to improve the purity of the mitochondrial preparation by digesting extramitochondrial proteins (29). Western blot analysis of the mitochondrial preparation revealed the presence of apparently full-length PNKP (Figure 1A). The purity of the preparation was confirmed by the absence of PCNA as a nuclear marker and the presence of COX IV (cytochrome C oxidase—subunit IV) or VDAC1 (Voltage Dependent Anion Channel) as mitochondrial markers. Notably no PNKP was detectable in mitochondria isolated from A549 cells

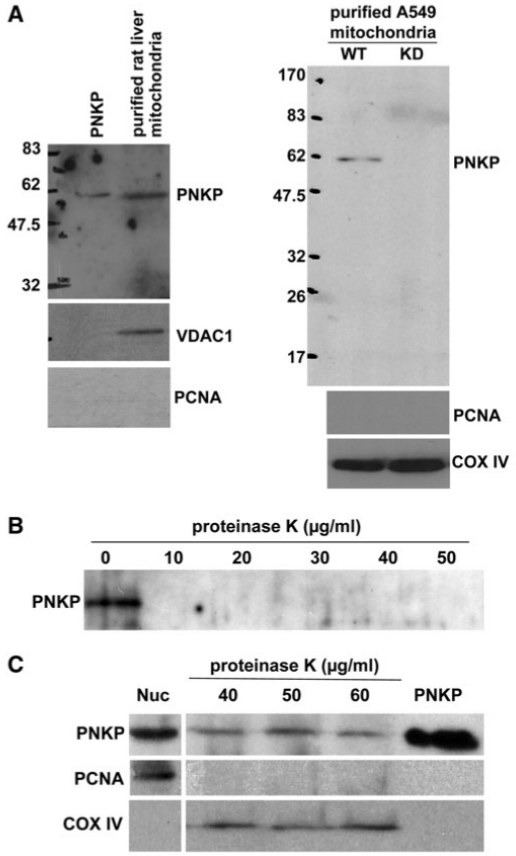


Figure 1. Full-length PNKP is present in mitochondria. (**A**) Mitochondria were isolated from rat liver and wild-type (WT) and PNKP knock-down (KD) A549 cells. Mitochondrial protein extracts were immunoblotted with a monoclonal antibody to PNKP. Antibodies against PCNA (nuclear marker) and COX IV or VDAC1 (mitochondrial markers) were used to ensure the purity of the mitochondrial preparation. (**B**) Purified human PNKP protein is sensitive to proteinase K, and is completely digested even at the lowest concentration used (10 μg ml^{-1}). (**C**) PNKP signal detected in purified mitochondria isolated from A549 cells is protected from proteinase K digestion. Proteins in nuclear extracts are shown on the left and purified PNKP protein on the right.

depleted of PNKP by shRNA. To further ensure that the PNKP signal came from inside the mitochondria and not due to contamination of the mitochondrial fraction with the nuclear PNKP, we also performed a proteinase K digestion assay. Purified human PNKP protein isolated from *Escherichia coli* was readily digested by proteinase K, being completely degraded by concentrations as low as 10 μg ml^{-1} (Figure 1B). However, mtPNKP recovered in the mitochondrial preparation was refractory to much higher concentrations of proteinase K (Figure 1C), indicating that the protein was protected by the mitochondrial membranes. Further titrations demonstrated that the

mtPNKP was resistant to proteinase K digestion up to 400 $\mu g\,ml^{-1}$ and significant digestion only started at 1 $mgml^{-1}$ (data not shown). That the mtPNKP is full-length was also confirmed by purification of mitochondria from A549 cell lines stably expressing either N- or C-terminally HA-tagged PNKP followed by trypsin treatment (Supplementary Figure S1).

Next we employed immunofluorescence confocal microscopy to localize the endogenous PNKP in A549 cells (Figure 2). PNKP in both nuclei and mitochondria was detected with a polyclonal antibody to PNKP, and the mitochondrial PNKP signal colocalized with the mitochondrial markers mitofilin and COX IV (Figure 2A). Similar data were obtained with a monoclonal antibody to PNKP (Supplementary Figure S2). To further demonstrate the specificity of the signal for PNKP, we repeated the experiment with a PNKP KD A549 cell line (Figure 2B). Comparison of upper and lower panels in Figure 2B demonstrates a marked reduction of both the nuclear and mitochondrial signals in the PNKP KD cells.

Based on western blotting mtPNKP appears to be the same size as the nuclear protein. However, we wanted to ensure that there is no major internal change in the functional domains of the enzyme that could interfere with their activities. We therefore directly

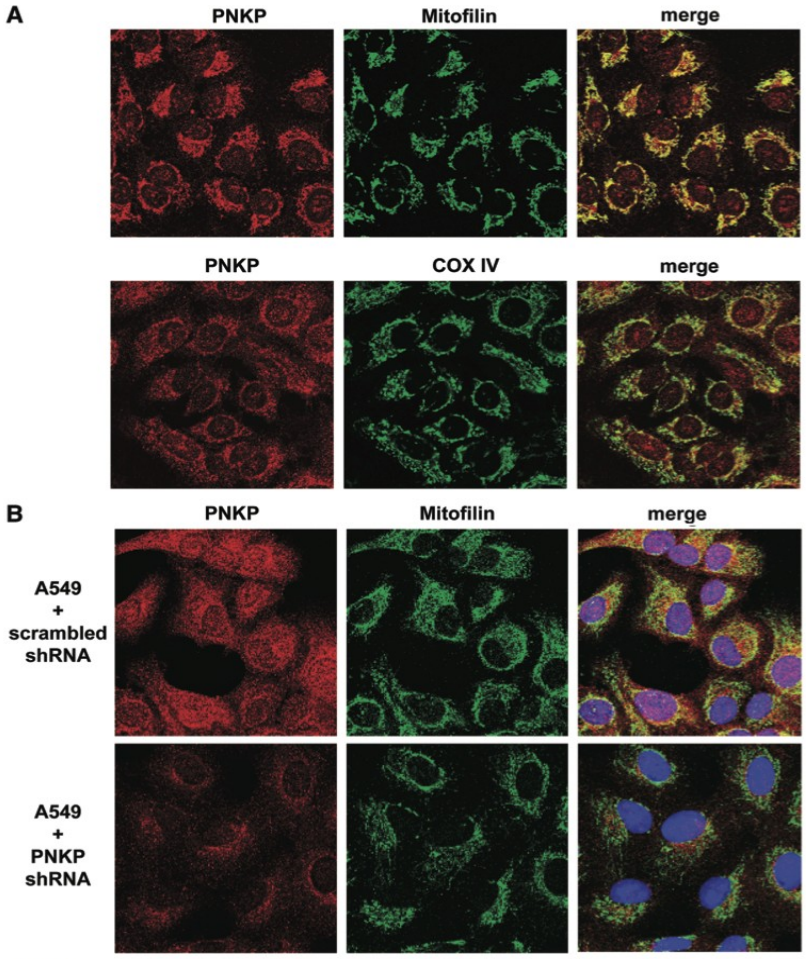


Figure 2. Endogenous PNKP colocalizes with mitochondrial markers. (**A**) Immunofluorescence of PNKP in A549 cells. PNKP localizes to mitochondria (in addition to the nuclei) as demonstrated by colocalization with the mitochondrial markers mitofilin and COX IV. (**B**) The shRNA KD of PNKP (lower panel) results in a decrease of the fluorescence signal for PNKP in both nuclei and mitochondria of A549 cells compared to the control cell line carrying scrambled shRNA (upper pannel).

examined the DNA kinase and phosphatase activities of mtPNKP (Figure 3A and B). Mitochondria were purified on a Percoll gradient, trypsin-treated and either sonicated or not before DNA kinase and phosphatase assays were performed (39). The kinase assay involves the transfer of radiolabeled phosphate from ATP to the 5′-terminus of an oligonucleotide, while the phosphatase assay monitors the loss of 3′-phosphate from a 5′-radiolabeled oligonucleotide (39). As shown in Figure 3, very minor kinase and phosphatase activities were observed in the absence of sonication of the mitochondrial membrane (probably due to release of mitochondrial proteins as a result of freeze-thawing of the samples), and both activities increased substantially following sonication. In contrast, no DNA kinase or phosphatase activity was detected with mitochondrial preparations isolated from PNKP KD cells (Figure 3). It is possible that other mitochondrial enzymes may have 5′-kinase or 3′-phosphatase activity (e.g. APE1), but presumably they do not act efficiently on the single-stranded substrates used to test PNKP activities. Thus this experiment indicates that mtPNKP possesses both DNA kinase and phosphatase activities and although we have not examined the protein at the amino acid sequence level, it would suggest that mtPNKP does not differ substantially, if at all, from the nuclear protein.

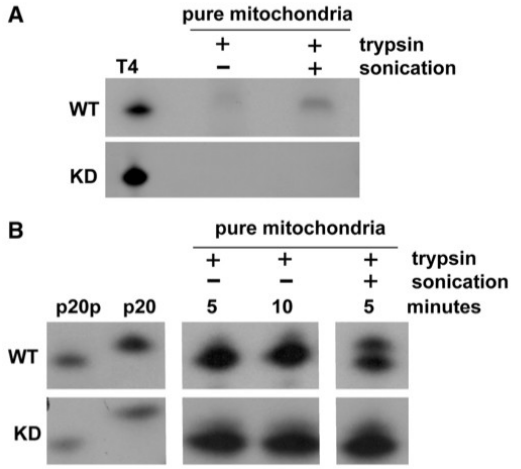


Figure 3. Mitochondrial PNKP displays both DNA kinase and phosphatase activities. (**A**) DNA kinase activity was measured by transfer of ^{32}P-phosphate from ATP to an oligonucleotide (21-mer) with a 5′-hydroxyl terminus. Mitochondria were purified from wild-type (WT) and PNKP KD A549 cells and treated with trypsin to digest any possible residual extramitochondrial protein contamination. Trypsin-treated mitochondria, either sonicated (+) or not sonicated (−), were used for DNA kinase assays. Sonication increased the DNA kinase activity of the purified mitochondrial preparation in WT cells. DNA kinase activity of T4 PNK is shown as a positive control. (**B**) DNA 3′-phosphatase activity in mitochondrial preparations from WT and KD A549 cells was determined by dephosphorylation of a 5′-^{32}P-labelled 3′-phosphorylated oligonucleotide (p20p), as substrate, resulting in conversion of p20p to p20 (markers shown on left). Sonication of the purified mitochondrial preparation substantially increased DNA phosphatase activity in WT cells.

mtPNKP is required for mtDNA repair

To show that mtPNKP is required for DNA repair in mitochondria, stable PNKP KD (A549 + PNKP shRNA) and control (A549 + scrambled shRNA) cell lines were exposed to 1.5 mM H_2O_2 for 1 h and repair was monitored 0, 2 and 4 h later. Total DNA was recovered from the cells (Supplementary Figure S3) and mtDNA repair was measured by (XL-qPCR) of an 8.9-kbp fragment of mtDNA (37). A 221-nt fragment of mtDNA was also amplified as an internal control and to show that the mtDNA was not degraded by H_2O_2, since such a small fragment would sustain only very limited damage under the conditions used. We employed picogreen to quantify the amount of large and small fragments amplified from mtDNA based on fluorescence reading. Finally, we compared the relative level of mtDNA amplification (large/small fragment) between the two cell lines. Our data (Figure 4A) clearly indicate that exposure of the PNKP-depleted (KD) cells to H_2O_2 ablated the signal from the 8.9-kbp fragment, strongly suggesting that PNKP is required for mtDNA repair. The decreased signal observed at 2 and 4 h in PNKP KD cell line (compared to the modest signal at 0 h) probably reflects an increase in the cleavage of mtDNA by DNA glycosylases/lyases during repair. Next, we wanted to determine the relative importance of the two enzymatic activities of PNKP for mtDNA repair in response to hydrogen peroxide. The PNKP KD cell line was transiently transfected with empty vector or HA-tagged wild-type PNKP (HAPNKP) as controls, and the kinase or phosphatase negative constructs (HAPNKPΔkin and HAPNKPΔphos), respectively. All the constructs carried mutations in the shRNA recognition site that prevented inhibition by shRNA but maintained the correct amino acid sequence (Supplementary Figure S4). The partially enzyme-inactivated PNKP constructs were prepared by site-directed mutagenesis of key residues in the kinase (K378A) or phosphatase (D171A and D173A) domains (17). The assay (Figure 4B) revealed intermediate activity for the kinase- and phosphatase-mutated PNKPs in dealing with H_2O_2-induced mtDNA damage in comparison to the wild-type PNKP and suggests that both the kinase and phosphatase activities of PNKP are required for the maintenance of mtDNA integrity.

PNKP contains a functional mitochondrial-targeting signal close to its carboxy-terminus

Mitochondrial-targeting signals (MTS) are most frequently found at the N-termini of proteins, but some of the mtDNA repair proteins reported to date do not contain an amino-terminal MTS. A computer-based analysis using multiple programs (including Mitoprot II and Predotar) showed that PNKP does not contain a canonical (N-terminal) MTS. However, a closer inspection of the sequence of PNKP revealed the presence of a 'cryptic' MTS close to the carboxy-terminus of the protein. This putative MTS consisted of amino acids 432–441 (ARYVQCARAA) and was identified by several computer programs, as mentioned. To determine if this MTS is functional, we fused the carboxy-terminus

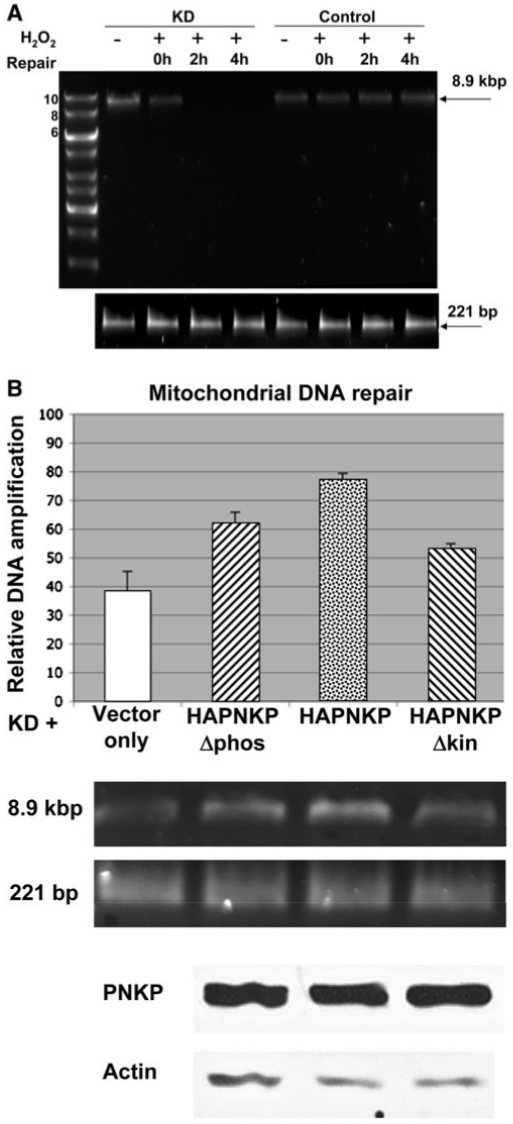


Figure 4. Functional PNKP is required for DNA repair in mitochondria following exposure to H_2O_2. (**A**) Total (i.e. nuclear + mitochondrial) DNA was purified from A549 cell lines stably transfected with PNKP shRNA (KD) or scrambled shRNA (control) (Supplementary Figure S3) following exposure to 1.5 mM H_2O_2 for 1 h and repair for 0, 2 and 4 h. Controls (–) were not exposed to H_2O_2. Upper panel: XL-qPCR performed on mtDNA, as described in Materials and Methods, amplified an 8.9-kbp PCR product following H_2O_2 exposure and repair in A549 cells expressing scrambled shRNA, but not in PNKP KD A549 cells. Lower panel: PCR of a 221-bp fragment from both cell lines indicating that both lines contain comparable amounts of mtDNA and that the H_2O_2 treatment did not degrade the DNA. (**B**) DNA repair in mitochondria in PNKP KD cells complemented with empty vector, wild-type (HAPNKP) or

of PNKP with (CmtsPNKP) or without the MTS (CPNKP) to GFP, generating CmtsPNKP + GFP and CPNKP + GFP (Figure 5A). The constructs were then transfected into A549 (Figure 5B) and MCF7 (data not shown) cells. In both cases an additional methionine was added to the amino terminus of the fusion protein. The carboxy-terminal region of PNKP incorporating the MTS was functional as a mitochondrial-localization signal and transferred most of the expressed CmtsPNKP + GFP into the mitochondria (Figure 5B, row 1). However, when the MTS was not included with the carboxy-terminal region of PNKP no CPNKP + GFP was transferred into mitochondria (Figure 5B, row 2), a result similar to the situation seen with GFP alone (Figure 5B, row 4). Further confirmation was provided by mutating the first three amino acids of the MTS as follows: A432D, R433G and Y434D. Computer analysis indicated that these mutations would dramatically decrease the capacity of the identified MTS to function as a true mitochondrial-localization signal. We observed that the protein expressed by this mutated construct (mutCmtsPNKP + GFP) failed to localize to the mitochondria (Figure 5B, row 3).

To test if the MTS is functional in the context of full-length PNKP, XL-qPCR was used to compare the functionality of HAPNKP-mts (MTS-mutated form of HAPNKP incorporating the same mutations to the MTS as described above) to HAPNKP in mtDNA repair. HAPNKP (positive control), vector only and HAPNKP-mts were transfected into PNKP KD cell lines prepared from A549 cells. As shown in Figure 6, the loss of the MTS resulted in a clear decrease in the activity of PNKP during mtDNA repair, similar to transfection with the vector only. The control western blot showed that the level of expression of HAPNKP-mts was the same as HAPNKP and thus the observed effect was not due to a lower level of protein expression of the mutated form of PNKP.

PNKP interacts with the mitochondrial proteins

It would be expected that the presence of PNKP in mitochondria would give rise to interaction of mtPNKP with mitochondrial proteins. We therefore expressed HAPNKP in an A549 PNKP KD background and immunoprecipitated the PNKP using anti-HA antibodies. We were able to identify mitofilin as one of the mitochondrial proteins that coimmunopreciptated with PNKP (Figure 7A). Interaction with mitofilin was further confirmed by reciprocal immunoprecipitation of HAPNKP by antibodies to mitofilin (Figure 7B). Mitofilin is a transmembrane protein of the inner membrane of mitochondria that has been shown to interact

Figure 4. Continued
kinase (HAPNKPΔkin) or phosphatase inactive (HAPNKPΔphos) PNKP. XL-qPCR was used to examine the level of DNA repair 30 min after exposure to 1 mM H_2O_2 for 1 h. The signal from the amplified 8.9 kbp mtDNA following XL-qPCR was normalized to the signal from the 221-bp fragment using Quant-it Pico Green Assay Kit. Error bars show the SD for three independent experiments. The western blot shows that similar levels of PNKP protein were expressed in the PNKP KD background.

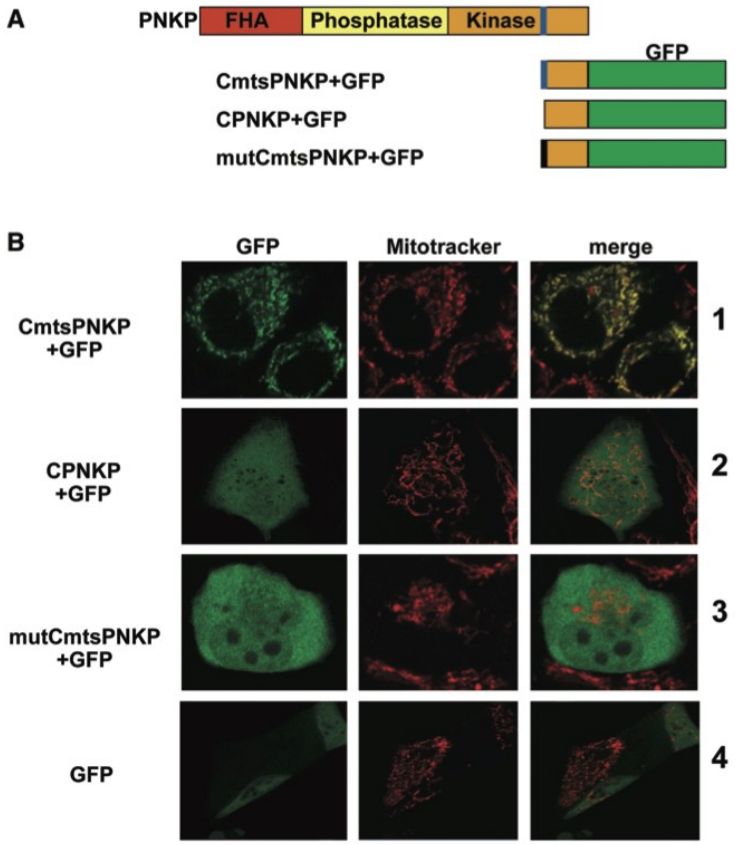


Figure 5. Mitochondrial localization of PNKP is dependent on the presence of a mitochondrial-targeting signal (MTS) in proximity to its carboxy terminus. (**A**) Computer programs (Mitoprot, Psort II and Predotar) predicted the presence of a mitochondrial-targeting signal (MTS) close to the C-terminus of PNKP (shown in blue). To further examine this potential MTS we generated three constructs (i) CmtsPNKP + GFP containing the GFP fused to the PNKP C-terminus including the putative MTS, (ii) CPNKP + GFP, which is essentially the same as CmtsPNKP + GFP but lacking the MTS sequence and (iii) mutCmtsPNKP + GFP, a mutated form of CmtsPNKP + GFP with the first three amino acids of the putative PNKP MTS mutated as follows: A432D, R433G and Y434D. In all cases a methionine was included at the amino-terminus. (**B**) The constructs were transfected into A549 cells and the cellular localization of GFP was monitored. Only the GFP fusion protein containing the wild-type MTS localized to mitochondria (row 1) as shown by colocalization with Mitotracker Orange (Molecular Probes). CPNKP + GFP, mutCmtsPNKP + GFP and GFP alone showed a diffuse signal throughout the cell (rows 2–4).

with PARP-1 (40), and DISC1 (disrupted-in-schizophrenia 1) (41). Interestingly, western blot analysis indicated that downregulation of PNKP by 70–90% in A549 cells reduced the level of mitofilin (Figure 7C). A similar reduction of the level of mitofilin in DISC1 KD cells has also been reported (41).

DISCUSSION

PNKP is a bifunctional end-processing enzyme involved in BER, and repair of single- and double-strand breaks. PNKP has been reported to be a nuclear protein by us and other laboratories (39,42). Here, we provide evidence that PNKP is also found in mitochondria. Given the abundance of ROS and the presence of other BER and single strand break repair proteins in mitochondria including uracil-DNA glycosylase, NEIL1, TDP1, APE1 and DNA ligase III (27,28,32), the appearance of PNKP in mitochondria is perhaps not surprising. Our data indicate that mtPNKP is the same size as the nuclear protein. Several of the BER proteins found in the mitochondria differ from their nuclear isoforms due to alternative splicing, as seen with mitochondrial uracil-DNA glycosylase (43), or post-translational processing, as observed with mitochondrial APE1 (29), or through the use of an alternative translation-initiation start site, as seen with mtDNA ligase III (44). Based on the western blots (Figure 1) coupled with observing both N- and C-terminally HA-tagged PNKP in the mitochondria (Supplementary Figure S1), we conclude that mtPNKP appears to be the same protein as nuclear PNKP.

Mitochondrial proteins often differ from their nuclear counterparts to incorporate a mitochondrial-targeting

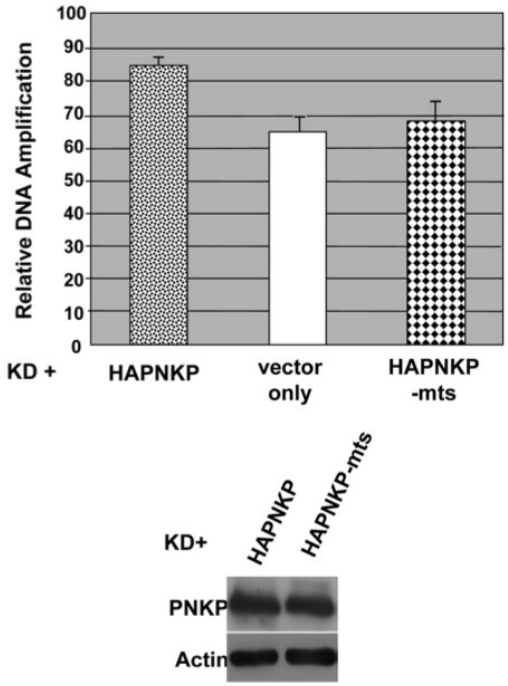


Figure 6. The MTS of PNKP is required for its function in mtDNA repair. XL-qPCR was used to monitor mtDNA repair in PNKP-depleted A549 cells treated with hydrogen peroxide as described in Figure 4B. Transient complementation of the cells with PNKP mutated in the first three amino acids of the MTS (HAPNKP-mts), as opposed to the wild-type protein (HAPNKP), reduces DNA repair in mitochondria to a level similar to the vector only control. The western blot of whole-cell extracts shown at the bottom of the figure indicates that similar levels of HAPNKP-mts and HAPNKP proteins were expressed in the A549 cells.

sequence (MTS) or to shed a nuclear localization signal (NLS). About 50% of mitochondrial proteins contain an MTS at their amino terminus. For example, aprataxin, which has recently been shown to play a role in mtDNA repair, possesses an amino-terminal MTS (45). However, localization to mitochondria does not necessarily require an amino-terminal MTS (46). Several mtDNA repair proteins, such as TDP1 and PARP-1, appear to lack an MTS (32,40). An MTS has been identified at the C-terminus of APE1 (47), and it appears that the redirection of APE1 into the mitochondria occurs after proteolytic removal by a mitochondria-associated N-terminal peptidase of the first 33 amino acids of APE1 that contains an NLS (29). A putative NLS has been previously identified in PNKP (amino acids 301–304), but it lies in the phosphatase domain of the protein (39,42,48,49) and is presumably retained in the mitochondrial PNKP. Alternatively, it has been shown that a C-terminal MTS can direct proteins to mitochondria in a carboxy to amino-terminal direction (34,50), and C-terminal MTSs have been found in several other mtDNA processing enzymes including the yeast DNA helicase Hmi1p (33) and the human nuclease/helicase DNA2 (51). An *in silico* analysis of the PNKP sequence uncovered a putative MTS located near the carboxy-terminus of the protein. We determined that this sequence is a functional MTS required for PNKP repair activity in mitochondria (Figure 6). It was also able to target GFP to mitochondria when placed at the N-terminus of GFP (Figure 5). Interestingly, using a similar *in silico* analysis with Mitoprot II and Predotar software, we detected a potential MTS in PARP-1 (amino acids 846–860) but this will require experimental verification.

We observed that mtPNKP retains both the DNA kinase and phosphatase activities (Figure 3), and that both activities are required for timely repair of mtDNA following damage induced by hydrogen peroxide (Figure 4). We have previously shown that shRNA-mediated depletion of PNKP sensitizes A549 cells to hydrogen peroxide (18). Further experimentation will be required to determine the relative importance of nuclear versus mitochondrial PNKP in response to hydrogen peroxide and other ROS, since this may have implications for the recently identified autosomal recessive neurological disorder, MCSZ, associated with mutations in PNKP (52). Although a direct comparison of the relative importance of PNKP kinase versus phosphatase activity for the repair of nuclear DNA has yet to be carried out, the phosphatase activity has been shown to play a critical role in single-strand break repair following oxidative damage by hydrogen peroxide and IR (19,53). The phosphatase activity of PNKP is also significantly more active than the kinase activity (17). Oxidative damage is more frequently associated with modifications to DNA 3′-strand-break termini than the 5′-termini (1,2). That the expression of a kinase-mutated PNKP had a similar outcome as the phosphatase-mutated PNKP on the repair of mtDNA suggests that 5′-hydroxyl termini are generated in similar quantities to modified 3′-termini, maybe through the intervention of trapped topoisomerase I complexes, or it may reflect differing levels of other mitochondrial repair proteins that may compensate for reduced PNKP activity. Alternatively, the relative importance of the kinase and phosphatase activities might be skewed by incomplete shRNA-mediated downregulation of endogenous PNKP that leaves cells with sufficient levels of the DNA phosphatase activity to partially compensate for the loss of phosphatase activity when expressing the phosphatase-mutated PNKP.

Our examination of an interaction between PNKP and mitofilin was predicated on a recent finding by Rossi *et al.* (40) that mitochondrial PARP-1 interacts with mitofilin. These authors further established that the presence of PARP-1 in mitochondria was dependent on the level of mitofilin and, based on the observation that mitofilin associates with proteins implicated in protein import into the mitochondria (54), they suggested that mitofilin may be involved in the translocation of PARP-1 across the mitochondrial membranes. Thus, interaction of mitofilin with PNKP could also assist with mitochondrial localization of PNKP. Knock down of PNKP resulted in a statistically significant decrease in the level of mitofilin. It is not readily apparent why the absence of PNKP should

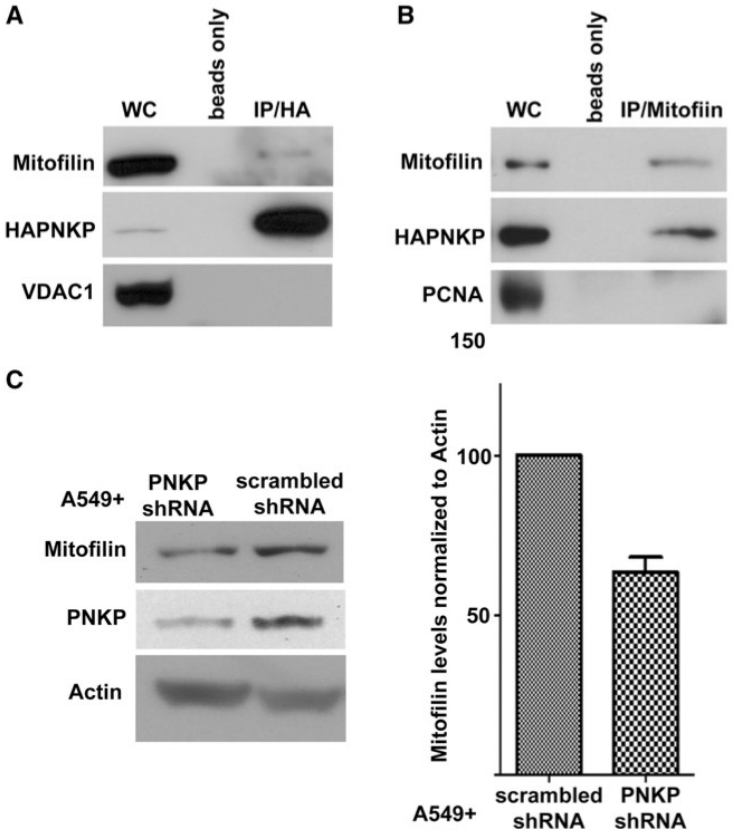


Figure 7. PNKP physically interacts with mitochondrial proteins. (**A**) Coimmunopreciptation of mitofilin by anti-HA antibody from whole-cell extract (WC) of A549 cells expressing HAPNKP. An antibody to the mitochondrial protein VDAC1 was used to ensure no non-specific pull-down of mitochondrial proteins. (**B**) Coimmunoprecipitation of PNKP with mitofilin from whole-cell extract of A549 cells expressing HAPNKP using a monoclonal antibody to mitofilin to immunoprecipitate mitofilin. In this case, the immunopreciptate was also tested for the presence of PCNA as a marker of potential contamination by nuclear proteins. The control lane 'beads only' indicates that no PNKP bound to the Sepharose beads in the absence of antibodies. (**C**) KD of PNKP by shRNA results in a decrease of about 40% in cellular mitofilin content. The graph shows the mean and SD values from three independent determinations ($P = 0.013$).

affect the level of mitofilin, but similar observations have been made with two other proteins, DISC1 (41) and the inner mitochondrial membrane protein ChChd3 (55). In the case of DISC1, its downregulation appears to increase proteasomal-mediated proteolysis of mitofilin (41). It remains to be determined whether downregulation of PNKP affects the level of mitofilin at the transcriptional or post-transcriptional level.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online: Supplementary Figures 1–4.

ACKNOWLEDGEMENTS

The authors thank Dr Xuejun Sun and Geraldine Barron for assistance with cell imaging, Dr Aghdass Rasouli-Nia, Mesfin Fanta, Ismail Abdou and Todd Mereniuk for other technical assistance and Dr Feridoun Karimi-Busheri for useful discussion and advice.

FUNDING

Canadian Institutes of Health Research (grant number 15385 to M.W.); a postdoctoral fellowship from the Alberta Cancer Foundation and an Incentive award from the Alberta Heritage Foundation for Medical Research (to N.T.). Funding for open access charge: Canadian Institutes of Health.

Conflict of interest statement. None declared.

REFERENCES

1. Henner,W.D., Grunberg,S.M. and Haseltine,W.A. (1982) Sites and structure of gamma radiation-induced DNA strand breaks. *J. Biol. Chem.*, **257**, 11750–11754.
2. Lennartz,M., Coquerelle,T., Bopp,A. and Hagen,U. (1975) Oxygen—effect on strand breaks and specific end-groups in DNA

of irradiated thymocytes. *Int. J. Radiat. Biol. Relat. Stud. Phys. Chem. Med.*, **27**, 577–587.
3. Weinfeld,M., Mani,R.S., Abdou,I., Aceytuno,R.D. and Glover,J.N. (2011) Tidying up loose ends: the role of polynucleotide kinase/phosphatase in DNA strand break repair. *Trends Biochem. Sci.*, **36**, 262–271.
4. Pourquier,P., Pilon,A.A., Kohlhagen,G., Mazumder,A., Sharma,A. and Pommier,Y. (1997) Trapping of mammalian topoisomerase I and recombinations induced by damaged DNA containing nicks or gaps. Importance of DNA end phosphorylation and camptothecin effects. *J. Biol. Chem.*, **272**, 26441–26447.
5. Plo,I., Liao,Z.Y., Barcelo,J.M., Kohlhagen,G., Caldecott,K.W., Weinfeld,M. and Pommier,Y. (2003) Association of XRCC1 and tyrosyl DNA phosphodiesterase (Tdp1) for the repair of topoisomerase I-mediated DNA lesions. *DNA Repair*, **2**, 1087–1100.
6. Bandaru,V., Sunkara,S., Wallace,S.S. and Bond,J.P. (2002) A novel human DNA glycosylase that removes oxidative DNA damage and is homologous to *Escherichia coli* endonuclease VIII. *DNA Repair*, **1**, 517–529.
7. Wiederhold,L., Leppard,J.B., Kedar,P., Karimi-Busheri,F., Rasouli-Nia,A., Weinfeld,M., Tomkinson,A.E., Izumi,T., Prasad,R., Wilson,S.H. *et al.* (2004) AP endonuclease-independent DNA base excision repair in human cells. *Mol. Cell*, **15**, 209–220.
8. Hegde,M.L., Hazra,T.K. and Mitra,S. (2008) Early steps in the DNA base excision/single-strand interruption repair pathway in mammalian cells. *Cell Res.*, **18**, 27–47.
9. Das,A., Wiederhold,L., Leppard,J.B., Kedar,P., Prasad,R., Wang,H., Boldogh,I., Karimi-Busheri,F., Weinfeld,M., Tomkinson,A.E. *et al.* (2006) NEIL2-initiated, APE-independent repair of oxidized bases in DNA: evidence for a repair complex in human cells. *DNA Repair*, **5**, 1439–1448.
10. Allinson,S.L. (2010) DNA end-processing enzyme polynucleotide kinase as a potential target in the treatment of cancer. *Future Oncol*, **6**, 1031–1042.
11. Loizou,J.I., El-Khamisy,S.F., Zlatanou,A., Moore,D.J., Chan,D.W., Qin,J., Sarno,S., Meggio,F., Pinna,L.A. and Caldecott,K.W. (2004) The protein kinase CK2 facilitates repair of chromosomal DNA single-strand breaks. *Cell*, **117**, 17–28.
12. Koch,C.A., Agyei,R., Galicia,S., Metalnikov,P., O'Donnell,P., Starostine,A., Weinfeld,M. and Durocher,D. (2004) Xrcc4 physically links DNA end processing by polynucleotide kinase to DNA ligation by DNA ligase IV. *EMBO J.*, **23**, 3874–3885.
13. Lu,M., Mani,R.S., Karimi-Busheri,F., Fanta,M., Wang,H., Litchfeld,D.W. and Weinfeld,M. (2010) Independent mechanisms of stimulation of polynucleotide kinase/phosphatase by phosphorylated and non-phosphorylated XRCC1. *Nucleic Acids Res.*, **38**, 510–521.
14. Mani,R.S., Yu,Y., Fang,S., Lu,M., Fanta,M., Zolner,A.E., Tahbaz,N., Ramsden,D.A., Litchfield,D.W., Lees-Miller,S.P. *et al.* (2010) Dual modes of interaction between XRCC4 and polynucleotide kinase/phosphatase: implications for nonhomologous end joining. *J. Biol. Chem.*, **285**, 37619–37629.
15. Pheiffer,B.H. and Zimmerman,S.B. (1982) 3′-phosphatase activity of the DNA kinase from rat liver. *Biochem. Biophys. Res. Commun.*, **109**, 1297–1302.
16. Habraken,Y. and Verly,W.G. (1988) Further purification and characterization of the DNA 3′-phosphatase from rat-liver chromatin which is also a polynucleotide 5′-hydroxyl kinase. *Eur. J. Biochem.*, **171**, 59–66.
17. Dobson,C.J. and Allinson,S.L. (2006) The phosphatase activity of mammalian polynucleotide kinase takes precedence over its kinase activity in repair of single strand breaks. *Nucleic Acids Res.*, **34**, 2230–2237.
18. Rasouli-Nia,A., Karimi-Busheri,F. and Weinfeld,M. (2004) Stable down-regulation of human polynucleotide kinase enhances spontaneous mutation frequency and sensitizes cells to genotoxic agents. *Proc. Natl Acad. Sci. USA*, **101**, 6905–6910.
19. Breslin,C. and Caldecott,K.W. (2009) DNA 3′-phosphatase activity is critical for rapid global rates of single-strand break repair following oxidative stress. *Mol. Cell Biol.*, **29**, 4653–4662.
20. Ralph,S.J., Rodriguez-Enriquez,S., Neuzil,J., Saavedra,E. and Moreno-Sanchez,R. (2010) The causes of cancer revisited: "mitochondrial malignancy" and ROS-induced oncogenic transformation—why mitochondria are targets for cancer therapy. *Mol. Aspects Med.*, **31**, 145–170.
21. Akbari,M., Skjelbred,C., Folling,I., Sagen,J. and Krokan,H.E. (2004) A gel electrophoresis method for detection of mitochondrial DNA mutation (3243 tRNA(Leu (UUR))) applied to a Norwegian family with diabetes mellitus and hearing loss. *Scand. J. Clin. Lab. Invest.*, **64**, 86–92.
22. van den Ouweland,J.M., Lemkes,H.H., Ruitenbeek,W., Sandkuijl,L.A., de Vijlder,M.F., Struyvenberg,P.A., van de Kamp,J.J. and Maassen,J.A. (1992) Mutation in mitochondrial tRNA(Leu)(UUR) gene in a large pedigree with maternally transmitted type II diabetes mellitus and deafness. *Nat. Genet.*, **1**, 368–371.
23. Brandon,M., Baldi,P. and Wallace,D.C. (2006) Mitochondrial mutations in cancer. *Oncogene*, **25**, 4647–4662.
24. Wallace,D.C. (2005) A mitochondrial paradigm of metabolic and degenerative diseases, aging, and cancer: a dawn for evolutionary medicine. *Annu. Rev. Genet.*, **39**, 359–407.
25. Trifunovic,A., Wredenberg,A., Falkenberg,M., Spelbrink,J.N., Rovio,A.T., Bruder,C.E., Bohlooly,Y.M., Gidlof,S., Oldfors,A., Wibom,R. *et al.* (2004) Premature ageing in mice expressing defective mitochondrial DNA polymerase. *Nature*, **429**, 417–423.
26. Pesole,G., Gissi,C., De Chirico,A. and Saccone,C. (1999) Nucleotide substitution rate of mammalian mitochondrial genomes. *J. Mol. Evol.*, **48**, 427–434.
27. de Souza-Pinto,N.C., Wilson,D.M. III, Stevnsner,T.V. and Bohr,V.A. (2008) Mitochondrial DNA, base excision repair and neurodegeneration. *DNA Repair*, **7**, 1098–1109.
28. Liu,P. and Demple,B. (2010) DNA repair in mammalian mitochondria: much more than we thought? *Environ. Mol. Mutagen*, **51**, 417–426.
29. Chattopadhyay,R., Wiederhold,L., Szczesny,B., Boldogh,I., Hazra,T.K., Izumi,T. and Mitra,S. (2006) Identification and characterization of mitochondrial abasic (AP)-endonuclease in mammalian cells. *Nucleic Acids Res.*, **34**, 2067–2076.
30. Hansen,A.B., Griner,N.B., Anderson,J.P., Kujoth,G.C., Prolla,T.A., Loeb,L.A. and Glick,E. (2006) Mitochondrial DNA integrity is not dependent on DNA polymerase-beta activity. *DNA Repair*, **5**, 71–79.
31. Simsek,D., Furda,A., Gao,Y., Artus,J., Brunet,E., Hadjantonakis,A.K., Van Houten,B., Shuman,S., McKinnon,P.J. and Jasin,M. (2011) Crucial role for DNA ligase III in mitochondria but not in Xrcc1-dependent repair. *Nature*, **471**, 245–248.
32. Das,B.B., Dexheimer,T.S., Maddali,K. and Pommier,Y. (2010) Role of tyrosyl-DNA phosphodiesterase (TDP1) in mitochondria. *Proc. Natl Acad. Sci. USA*, **107**, 19790–19795.
33. Lee,C.M., Sedman,J., Neupert,W. and Stuart,R.A. (1999) The DNA helicase, Hmi1p, is transported into mitochondria by a C-terminal cleavable targeting signal. *J. Biol. Chem.*, **274**, 20937–20942.
34. Folsch,H., Gaume,B., Brunner,M., Neupert,W. and Stuart,R.A. (1998) C- to N-terminal translocation of preproteins into mitochondria. *EMBO J.*, **17**, 6508–6515.
35. Fanta,M., Zhang,H., Bernstein,N., Glover,M., Karimi-Busheri,F. and Weinfeld,M. (2001) Production, characterization, and epitope mapping of monoclonal antibodies against human polydeoxyribonucleotide kinase. *Hybridoma*, **20**, 237–242.
36. Akbari,M., Visnes,T., Krokan,H.E. and Otterlei,M. (2008) Mitochondrial base excision repair of uracil and AP sites takes place by single-nucleotide insertion and long-patch DNA synthesis. *DNA Repair*, **7**, 605–616.
37. Santos,J.H., Meyer,J.N., Mandavilli,B.S. and Van Houten,B. (2006) Quantitative PCR-based measurement of nuclear and mitochondrial DNA damage and repair in mammalian cells. *Methods Mol. Biol.*, **314**, 183–199.
38. Crampton,N., Kodiha,M., Shrivastava,S., Umar,R. and Stochaj,U. (2009) Oxidative stress inhibits nuclear protein export by multiple mechanisms that target FG nucleoporins and Crm1. *Mol. Biol. Cell*, **20**, 5106–5116.
39. Karimi-Busheri,F., Daly,G., Robins,P., Canas,B., Pappin,D.J., Sgouros,J., Miller,G.G., Fakhrai,H., Davis,E.M., Le Beau,M.M.

et al. (1999) Molecular characterization of a human DNA kinase. *J. Biol. Chem.*, **274**, 24187–24194.
40. Rossi,M.N., Carbone,M., Mostocotto,C., Mancone,C., Tripodi,M., Maione,R. and Amati,P. (2009) Mitochondrial localization of PARP-1 requires interaction with mitofilin and is involved in the maintenance of mitochondrial DNA integrity. *J. Biol. Chem.*, **284**, 31616–31624.
41. Park,Y.U., Jeong,J., Lee,H., Mun,J.Y., Kim,J.H., Lee,J.S., Nguyen,M.D., Han,S.S., Suh,P.G. and Park,S.K. (2010) Disrupted-in-schizophrenia 1 (DISC1) plays essential roles in mitochondria in collaboration with Mitofilin. *Proc. Natl Acad. Sci. USA*, **107**, 17785–17790.
42. Jilani,A., Ramotar,D., Slack,C., Ong,C., Yang,X.M., Scherer,S.W. and Lasko,D.D. (1999) Molecular cloning of the human gene, PNKP, encoding a polynucleotide kinase 3′-phosphatase and evidence for its role in repair of DNA strand breaks caused by oxidative damage. *J. Biol. Chem.*, **274**, 24176–24186.
43. Nilsen,H., Otterlei,M., Haug,T., Solum,K., Nagelhus,T.A., Skorpen,F. and Krokan,H.E. (1997) Nuclear and mitochondrial uracil-DNA glycosylases are generated by alternative splicing and transcription from different positions in the UNG gene. *Nucleic Acids Res.*, **25**, 750–755.
44. Lakshmipathy,U. and Campbell,C. (1999) The human DNA ligase III gene encodes nuclear and mitochondrial proteins. *Mol. Cell Biol.*, **19**, 3869–3876.
45. Sykora,P., Croteau,D.L., Bohr,V.A. and Wilson,D.M. III (2011) Aprataxin localizes to mitochondria and preserves mitochondrial function. *Proc. Natl Acad. Sci. USA*, **108**, 7437–7442.
46. Baker,M.J., Frazier,A.E., Gulbis,J.M. and Ryan,M.T. (2007) Mitochondrial protein-import machinery: correlating structure with function. *Trends Cell Biol.*, **17**, 456–464.
47. Li,M., Zhong,Z., Zhu,J., Xiang,D., Dai,N., Cao,X., Qing,Y., Yang,Z., Xie,J., Li,Z. *et al.* (2010) Identification and characterization of mitochondrial targeting sequence of human apurinic/apyrimidinic endonuclease 1. *J. Biol. Chem.*, **285**, 14871–14881.
48. Bernstein,N.K., Williams,R.S., Rakovszky,M.L., Cui,D., Green,R., Karimi-Busheri,F., Mani,R.S., Galicia,S., Koch,C.A., Cass,C.E. *et al.* (2005) The molecular architecture of the mammalian DNA repair enzyme, polynucleotide kinase. *Mol. Cell*, **17**, 657–670.
49. Meijer,M., Karimi-Busheri,F., Huang,T.Y., Weinfeld,M. and Young,D. (2002) Pnk1, a DNA kinase/phosphatase required for normal response to DNA damage by gamma-radiation or camptothecin in *Schizosaccharomyces pombe*. *J. Biol. Chem.*, **277**, 4050–4055.
50. Chacinska,A., Koehler,C.M., Milenkovic,D., Lithgow,T. and Pfanner,N. (2009) Importing mitochondrial proteins: machineries and mechanisms. *Cell*, **138**, 628–644.
51. Zheng,L., Zhou,M., Guo,Z., Lu,H., Qian,L., Dai,H., Qiu,J., Yakubovskaya,E., Bogenhagen,D.F., Demple,B. *et al.* (2008) Human DNA2 is a mitochondrial nuclease/helicase for efficient processing of DNA replication and repair intermediates. *Mol. Cell*, **32**, 325–336.
52. Shen,J., Gilmore,E.C., Marshall,C.A., Haddadin,M., Reynolds,J.J., Eyaid,W., Bodell,A., Barry,B., Gleason,D., Allen,K. *et al.* (2010) Mutations in PNKP cause microcephaly, seizures and defects in DNA repair. *Nat. Genet.*, **42**, 245–249.
53. Freschauf,G.K., Karimi-Busheri,F., Ulaczyk-Lesanko,A., Mereniuk,T.R., Ahrens,A., Koshy,J.M., Rasouli-Nia,A., Pasarj,P., Holmes,C.F., Rininsland,F. *et al.* (2009) Identification of a small molecule inhibitor of the human DNA repair enzyme polynucleotide kinase/phosphatase. *Cancer Res.*, **69**, 7739–7746.
54. Xie,J., Marusich,M.F., Souda,P., Whitelegge,J. and Capaldi,R.A. (2007) The mitochondrial inner membrane protein mitofilin exists as a complex with SAM50, metaxins 1 and 2, coiled-coil-helix coiled-coil-helix domain-containing protein 3 and 6 and DnaJC11. *FEBS Lett.*, **581**, 3545–3549.
55. Darshi,M., Mendiola,V.L., Mackey,M.R., Murphy,A.N., Koller,A., Perkins,G.A., Ellisman,M.H. and Taylor,S.S. (2011) ChChd3, an inner mitochondrial membrane protein, is essential for maintaining crista integrity and mitochondrial function. *J. Biol. Chem.*, **286**, 2918–2932.