Characterization and inhibition of interstrand crosslink repair nuclease SNM1A

Beverlee Buzon, BSc, MSc
McMaster University, beverleebuzon@gmail.com

CHARACTERIZATION AND INHIBITION OF SNM1A

CHARACTERIZATION AND INHIBITION OF INTERSTRAND CROSSLINK
REPAIR NUCLEASE SNM1A

By BEVERLEE BUZON, BSc, MSc

A Thesis Submitted to the School of Graduate Studies
in Partial Fulfillment of the Requirements for
the Doctorate of Philosophy

McMaster University Doctorate of Philosophy (2018) Hamilton, Ontario (Biochemistry
and Biomedical Sciences)

TITLE: Characterization and inhibition of interstrand crosslink repair nuclease SNM1A
AUTHOR: Beverlee Buzon, BSc (Ryerson University), MSc (McMaster University)
SUPERVISOR: Dr. Murray Junop
NUMBER OF PAGES: 213

ABSTRACT

Interstrand cross-links (ICLs) are a type of DNA damage that prevents strand separation required for basic cellular processes. ICL-based anti-cancer therapies exploit the cytotoxic consequences of replication and transcription inhibition, however, they are limited by the ability of the cell to repair DNA crosslinks. The challenge of ICL repair involves coordinating multiple DNA repair pathways to remove damage occurring on both strands of DNA. Participation of factors that are both exclusive and essential to crosslink repair suggests a pathway requirement to process unique structures and/or intermediates arising only in ICL repair. SNM1A is a nuclease required for survival of human cells in response to ICL exposure, but the specific function and role of SNM1A remain unclear. Here we show that, in addition to known 5'-3'exonuclease activity, SNM1A possesses single-strand specific endonuclease activity. Furthermore, SNM1A exhibits translesion nuclease activity on crosslinks which deform the helical backbone, but not non-distorting stable ICLs. We report the identification and characterization of nine small molecules inhibitors of SNM1A, isolated from an *in vitro* high-throughput screen of nearly 4,000 bioactive compounds. Finally, we demonstrate that inhibitors of SNM1A potentiate the cytotoxicity of ICL-inducing agent cisplatin in HeLa cells. The work in this thesis expands the possible roles of SNM1A in ICL repair and lays the groundwork for SNM1A inhibition in ICL sensitization efforts.

screen), Cecilia Murphy (HTS project), Jenny Wong (HTS technician), and Fazmin Nazim (HTS data processing).

Many thanks to everyone in the Biochemistry department at Western, particularly Dr. Lazlo Gyenis and Dr. Caroline Schild-Poulter.

Thanks to all my friends and family for their support, even though I was AWOL for so long. To my parents for their never-ending encouragement and love. And finally, to my husband, Ian Tam. Thank you for waiting for so so long. Mahal kita.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

53BP1	p53-binding protein 1
6-FAM	6-carboxyfluorescein
6-OH DOPA	6-hydroxy-dl-dihydroxyphenylalanine
8-MOP	8-methoxypsoralen
ADE2	phosphoribosylaminoimidazole carboxylase
AlkB	alpha-ketoglutarate-dependent dioxygenase
Alt-EJ	alternative end-joining
APC	anaphase promoting complex
APE1	apurinic/apyrimidinic endonuclease
APIM	AlkB homologue 2 PCNA-interacting motif
AT	ataxia telangiectasia
ATA	aurintricarboxylic acid
ATM	ataxia telangiectasia mutated
ATP7A	copper-transporting P-type ATPase
BCNU	bis-chloroethylnitrosourea
BER	base excision repair
BHQ1	Black Hole Quencher 1®
BLM	Bloom syndrome protein
BRCA1	breast cancer type 1 susceptibility protein
BRCA2	breast cancer type 2 susceptibility protein
BSA	bovine serum albumin
BTR	BLM-TOP3A-RMI
CAK	CDK-activating kinase
CASP	CPSF-73, Artemis, Snm1 and Pso2
CMG	CDC45-MCM-GINS
CFU	colony-forming units
CPSF-73	cleavage and polyadenylation specificity factor 73

CS	Cockayne syndrome
CSA	Cockayne syndrome group a
CSB	Cockayne syndrome group b
CT	camptothecin
Ctf18	chromosome transmission fidelity 18
CtIP	CtBP-interacting protein
CTR1	copper transporter 1
Cy3	cyanine dye 3
DHPG	7-methyl deshydroxypyrollagallol-4-carboxylic acid
DIDS	4,4'-diisothiocyanostilbene-2,2'-sulfonic acid sodium salt
DMEM	Dulbecco's modified eagle medium
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
DNA-PKcs	DNA protein kinase catalytic subunit
dNTP	deoxynucleotide
DSB	double-strand break
dsDNA	double-stranded DNA
DTT	dithiothreitol
EDTA	ethylenediaminetetraacetate
EME1	essential meiotic endonuclease 1
ERCC1	excision repair cross-complementing 1
ERCC4	excision repair cross-complementing 4
FA	Fanconi anemia
FAAP100	Fanconi anemia-associated protein 100
FAAP20	Fanconi anemia-associated protein 20
FAAP24	Fanconi anemia-associated protein 24
6-FAM	carboxyfluorescein

FAN1	Fanconi anemia-associated nuclease 1
FANCA	Fanconi anemia complementing group a
FANCB	Fanconi anemia complementing group b
FANCC	Fanconi anemia complementing group c
FANCD2	Fanconi anemia complementing group d2
FANCE	Fanconi anemia complementing group e
FANCF	Fanconi anemia complementing group f
FANCG	Fanconi anemia complementing group g
FANCI	Fanconi anemia complementing group i
FANCL	Fanconi anemia complementing group l
FANCM	Fanconi anemia complementing group m
FANCQ	Fanconi anemia complementing group q
FPLC	fast protein liquid chromatography
GEN1	flap endonuclease GEN homolog 1
GG-NER	global genome nucleotide excision repair
GSH	glutathione (reduced)
GSSG	glutathione (oxidized)
GST	glutathione s-transferase
H3	histone 3
His	histidine
HJ	Holliday junction
HMG1	high mobility group protein 1
hMutSβ	MSH2-MSH3 heterodimer
MSH2	DNA mismatch repair protein 2
MSH3	DNA mismatch repair protein 3
HR	homologous recombination
HTS	high throughput screening
HU	hydroxyurea

ICL	interstrand crosslink
ID2	FANCI-FANCD2
(Ub)	monoubiquitinated
IMAC	immobilized metal affinity chromatography
IPTG	isopropyl β-d-1-thiogalactopyranoside
IQM	interquartile mean
IRES	internal ribosome entry site
LB	Luria Burtani
LigI	DNA ligase 1
LigIII	DNA ligase 3
MBL	metallo β-lactamase
MEF	mouse embryonic fibroblast
MMC	mitomycin c
MMS	methyl methanesulfonate
MPG	n-methylpurine DNA glycosylase
MRE11	meiotic recombination 11
MRN	Mre11-Rad50-Nbs1 complex
MRP2	multidrug resistance-associated protein 2
MT	metallothionein
MUS81	methansulfonate, UV sensitive protein 81
MutL	DNA mismatch repair protein mutl
MWCO	molecular weight cut-off
Nbs2	nibrin
NER	nucleotide excision repair
NHEJ	non-homologous end-joining
NTA	nitrilotriacetic acid
OD	optical density
p53	cellular tumor antigen p53

PAGE	polyacrylamide gel electrophoresis
PARP1	poly [ADP-ribose] polymerase 1
PCNA	proliferating cell nuclear antigen
PCR	polymerase chain reaction
PDB	pyrrolobenzodiazapine
PIAS1	protein inhibitor of activated stat protein 1
PIKK	phosphatidylinositol 3-kinase-related kinase
PIP	PCNA interacting protein box
PMSF	phenylmethane sulfonyl fluoride
Pol z	polymerase zeta
Pol δ	polymerase delta
Pol ε	polymerase epsilon
Pol η	polymerase eta
Pol ι	polymerase iota
Pol κ	polymerase kappa
Pol ν	polymerase nu
Pol ξ	polymerase xi
PSO2	psoralen sensitive protein 2
RAD18	radiation sensitive protein 18
RAD23	radiation sensitive protein 23
Rad50	radiation sensitive protein 50
RAD51	radiation sensitive protein 51
RAD6	radiation sensitive protein 6
RCF	replication clamp factor
RFU	relative fluorescence units
RIR	replication independent repair
RMI1	RECQ-mediated genome instability protein 1
RMI2	RECQ-mediated genome instability protein 2

RNA	ribonucleic acid
RPA	replication protein a
SCID	severe combined immunodeficiency
SCLC	small cell lung carcinoma
SDS	sodium dodecyl sulfate
SNF	switch/sucrose non-fermentable
SNM1A	sensitive to nitrogen mustard 1a
SSB	single-strand break
ssDNA	single stranded DNA
STAT	signal transducer and activator of transcription
SUMO	small ubiquitin-like modifier
TAPS	n-tris(hydroxymethyl)methyl-3-aminopropanesulfonic acid
TBE	tris/borate/edta
TCEP	tris(2-carboxyethyl)phosphine
TC-NER	transcription-coupled nucleotide excision repair
TEV	tobacco etch virus
TFIIH	transcription factor IIH
TLS	translesion synthesis
TOPIIIA	topoisomerase III alpha
TRF2	telomeric repeat-binding factor 2
Ub	monoubiquitinated
UBC9	ubiquitin conjugating enzyme E2i
UBZ	ubiquitin binding zinc finger
USP1	ubiquitin specific peptidase 1
UAF1	USP1-associated factor 1
UTR	untranslated region
UV	ultraviolet
V(D)J	variable (diversity) joining

VRR-NUC	virus replication repair nuclease
XAB2	XPA binding protein 2
XP	xeroderma pigmentosa
XPA	xeroderma pigmentosum complementation group a
XPB	xeroderma pigmentosum complementation group b
XPC	xeroderma pigmentosum complementation group c
XPD	xeroderma pigmentosum complementation group d
XPF	xeroderma pigmentosa, complementation group f
XPG	xeroderma pigmentosum complementation group g
XRCC3	x-ray repair cross complementing 3
YPD	yeast peptone dextrose
γGCS	gamma glutamylcysteine synthetase
γH2AX	gamma H2A histone family, member x

CHAPTER ONE: INTRODUCTION

1.1   DNA Damage by Interstrand Crosslinks (ICLs)

The integrity of DNA is imperative as DNA encodes the genetic instructions that dictates all cellular processes. As such, damage to DNA is deleterious, if not lethal. Of the types of DNA damage, interstrand crosslinks (ICLs) are amongst the most cytotoxic (Noll et al., 2006). An ICL adduct covalently links opposing strands of duplex DNA, preventing strand separation essential for replication and transcription. As a result, replication fork collapse and subsequent formation of double strand breaks (DSBs) underlie the cytotoxicity of ICLs. As few as twenty ICLs are capable of inducing cell death in mammalian cells if left unrepaired, making ICLs one of the most lethal forms of DNA damage (Lawley and Phillips, 1996).

ICL-inducing agents are bifunctional compounds that give rise to numerous types of DNA lesions, including DNA-protein adducts, mono-adducts, and both intra- and interstrand DNA crosslinks (Noll et al., 2006). Interstrand crosslinks only comprise a small percentage of DNA lesions that occur in response to exposure to ICL-inducing agents. However, these bifunctional crosslinking agents are as much as a thousand-fold more cytotoxic compared to their monofunctional counterparts. This reflects the difficulty that cells have in dealing with DNA simultaneously damaged on both strands and is the basis of the lethality caused by ICL damage (Vogel et al., 1996). Unlike monofunctional DNA damaging agents that typically result in only a single base change mutation, ICL formation results in clastogenic events, including double-strand breaks, hyper-

recombination, chromosomal rearrangements and deletions (Vogel et al., 1996). As such, these structural changes to chromosomal DNA underlie the extreme cytotoxicity exerted by ICL-inducing agents.

1.2   Sources of ICL Damage

Evidence of ICLs and their consequences was first observed during World War I, when ICL damaging agents were deployed as chemical weapons (Goodman et al., 1946). Subsequently, the study of ICLs were focused on effects of ICL damage caused by external agents and paid little attention to other sources of ICL damage. The extent of ICL damage arising endogenously were largely ignored as ICLs formed from endogenous ICL-inducing agents were unstable *in vitro* (Niedernhofer et al., 2003). However, studies have now clearly demonstrated ICL formation in the absence of exogenous ICL-inducing agents, indication that ICL formation occurs endogenously.

1.2.1   Endogenous Sources of ICLs

Within the cell, endogenous sources of ICL-inducing agents (**Figure 1**B) can result from the conversion of metabolic by-products into DNA damaging agents. The most characterized cellular ICL-inducing agents are by-products of lipid peroxidation, in which reactive oxygen species break down lipids into bifunctional compounds (Stone et al., 2008). Malondialdehyde and crotonaldehyde are formed from the oxidation of polyunsaturated fats and arachidonic acid, respectively (Niedernhofer et al., 2003). Crotonaldehyde is also formed as a product of the aldol condensation of acetaldehyde, which is a metabolite of alcohol (Rekoske and Barteau, 2011). Consequently, mice

deficient in aldehyde dehydrogenase and people who suffer from alcoholism show chromosomal instability indicative of ICL damage (Brooks and Theruvathu, 2005; Folmer et al., 2003).

Recent studies have shown that abasic sites within DNA are also able to give rise to interstrand crosslinks. Abasic sites formed during base excision repair (BER) or by spontaneous depurination, are highly reactive. If left unrepaired, abasic sites are capable of crosslinking up to 20% of DNA within 12 hours (Greenberg, 2013). However, since abasic sites are strongly sequestered by BER factors [e.g. apurinic endonuclease 1 (APE1), Poly(ADP-ribose) polymerase 1 (PARP1)], the extent of ICL formation arising from abasic sites is unclear (Admiraal and O'Brien, 2015). Nevertheless, since abasic site formation is one of the most common forms of DNA damage, occurring an estimated ten thousand times a day within a cell, the significance of abasic site conversion to ICL damage endogenously warrants further investigation (Greenberg, 2014).
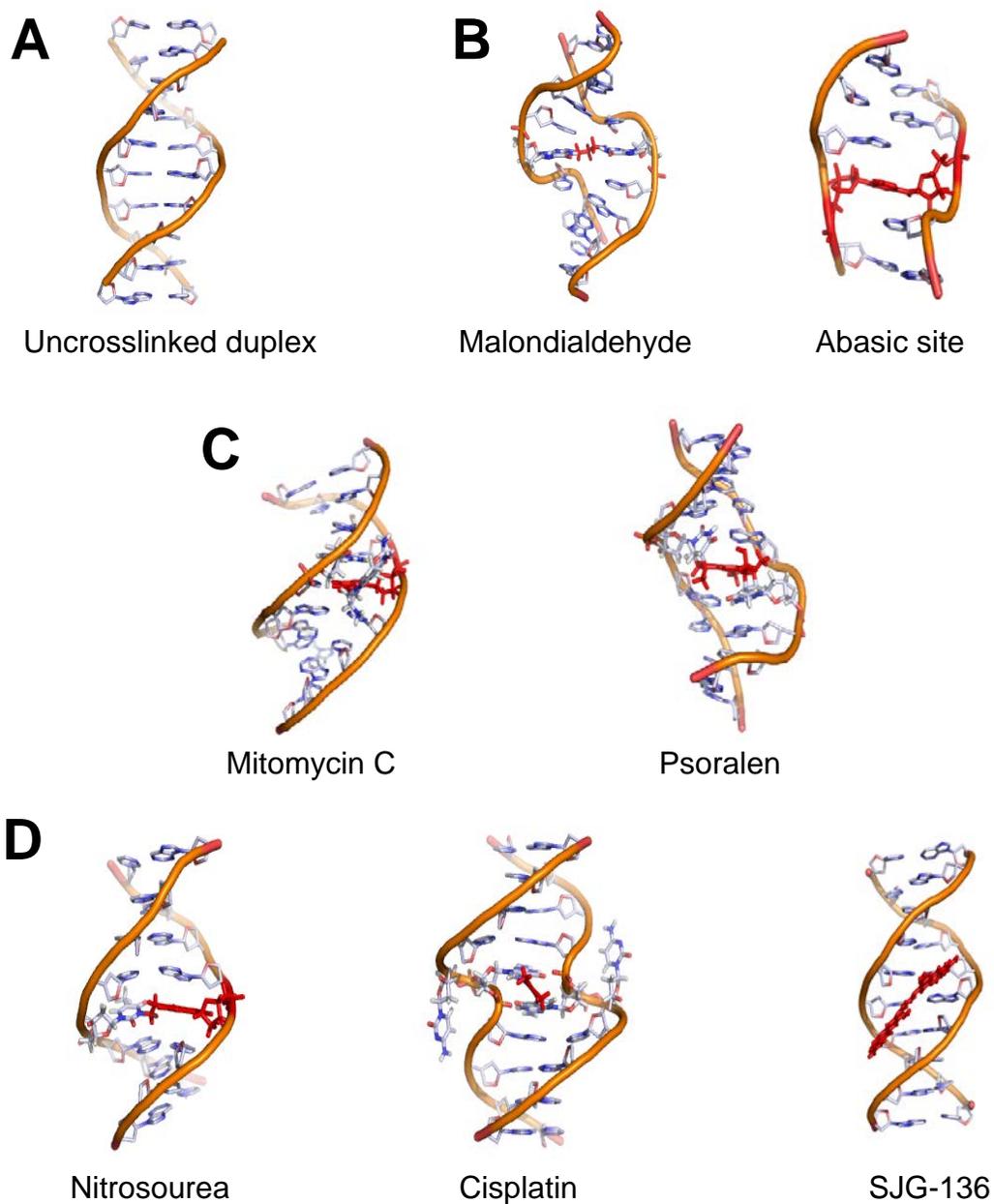
Figure 1: Sources of ICL-inducing agents

(**A**) Duplex DNA. (**B**) Endogenous sources of ICL-inducing agents. (**C**) Natural exogenous sources of ICL-inducing agents. (**D**) Synthetic sources of ICL-inducing agents.

1.2.2 Exogenous Sources of ICL-Inducing Agents

Natural sources of exogenous ICL-inducing agents (**Figure 1**C) have evolved in some plants and bacteria as a defense mechanism against other organisms. Actinobacteria *Streptomyces caespitosus* produces mitomycin C (MMC), a bio-reductive compound requiring intracellular conversion to its reduced cytotoxic form (Paz et al., 2012). The anaerobic nature of tumors result in a reducing environment, which preferentially renders hypoxic tumor cells susceptible to MMC damage compared to normal cells (Verweij and Pinedo, 1990). Therefore, MMC is used against a number of malignancies (e.g. bladder and liver cancers), but development of thrombocytopenia and leukocytopenia off-target effects limit the use of MMC as a main anti-cancer drug (Verweij and Pinedo, 1990).

In addition, plants from the *Apium* family (e.g. carrots, celery, parsley) and certain citrus fruits (e.g. grapefruit, pomelos) produce furocoumarins that include psoralen and its derivatives (Sayre and Dowdy, 2008). Although the average dietary daily intake of psoralen is about 2mg, ICL cytotoxicity is limited since psoralens must be activated by UVA to crosslink DNA (Sayre and Dowdy, 2008). As such, psoralen and its derivative 8-methoxypsoralen (8-MOP) have been successfully used to treat skin conditions including psoriasis and vitiligo, as local application and activation allows for specific targeting (Gupta and Anderson, 1987).

Unlike naturally produced ICL-inducing agents, synthetic ICL-inducing agents (**Figure 1**D) contain reactive bifunctional alkylating groups and therefore do not require significant conversion to exert their cytotoxic effects. Nitrogen mustards crosslink DNA

strands via their chloroethylene functional groups. This class of compounds includes mustard gas whose use during WWI in chemical warfare was responsible for thousands of deaths (Pechura and Rall, 1993). Paradoxically, nitrogen mustards (e.g. melphalan, chlorambucil, and cyclophosphamide) have since saved countless lives as part of common chemotherapy regimens. Related nitrosoureas [e.g. bis-chloroethylnitrosourea (BCNU), lomustine] are lipophilic ICL-inducing agents that readily pass the blood brain barrier and have been used extensively for treatment of brain cancers (Wang et al., 1999).

More recently, crosslinking compounds have been rationally designed to overcome off-target reactions with non-DNA biomolecules. Of particular note, SJG-136 is currently in phase II of clinical trials for a broad range of cancers (Hartley and

Amongst synthetic ICL-inducing agents, platinum-based compounds are the most widely used anti-cancer drugs (Galluzzi et al., 2012). Cisplatin is considered an essential drug by the World Health Organization as it is commonly used to treat the ten most frequently occurring tumors worldwide (WHO Model List of Essential Medicines, 20th edition). Cisplatin is a first-line treatment in many solid mass malignancies, especially efficacious in the treatment of testicular cancers (Galluzzi et al., 2012). Unfortunately, the chemotherapeutic dose is limited by peripheral neurotoxicity. Additional off-target effects include nephrotoxicity, ototoxicity, severe nausea and vomiting (Kelland, 2007). In attempts to reduce side-effects, platinum derivatives (e.g. carboplatin and oxaliplatin) were created (Harrap, 1985). Although these derivatives exhibit reduced off-target toxicity, efficacy is also reduced and therefore cisplatin continues as a mainstay cancer therapy (Kelland, 2007).

More recently, crosslinking compounds have been rationally designed to overcome off-target reactions with non-DNA biomolecules. Of particular note, SJG-136 is currently in phase II of clinical trials for a broad range of cancers (Hartley and

Hochhauser, 2012). SJG-136 contains two reactive pyrrolobenzodiazapine (PDB) moieties and a structure that permits high (nM) affinity interaction specifically with the minor grove of DNA, making it an ideal compound with limited off-target toxicity and overall high efficacy (Hartley et al., 2004a).

1.3   Chemoresistance to ICL-Based Anti-Cancer Therapies

Since ICLs block DNA strand separation and therefore DNA replication, highly proliferative tumor cells are preferentially sensitive to the effects of ICL-inducing agents. Despite initial therapeutic success in response to ICL-based chemotherapy, resistance to ICL-inducing agents frequently occurs and results in therapeutic failure and cancer relapse (Gatti and II, 2005). Chemoresistance is particularly problematic with cisplatin-based anti-cancer regimens (Köberle et al., 2010). Resistance to ICL-inducing agents is generally classified into intrinsic or acquired resistance mechanisms (**Figure 1**A). Acquired resistance emerges after initial treatment, resulting in refractory tumors that no longer respond to initial effective doses of ICL-inducing agent. In ovarian cancers, more than 60% of patient tumours respond well to ICL treatment (Eckstein, 2011). Unfortunately, approximately 80% of these tumors return, resulting in patient relapse within six months (Gotlieb et al., 2007). Worse still, in metastatic ovarian cancer, no additional treatments are available following acquisition of ICL resistance and chemotherapy failure, leaving women with only palliative care options (Cannistra, 2004). Unlike acquired resistance, intrinsic resistance manifests as tumor insensitivity in which there is no initial response to chemotherapeutic treatment. Up to 40% of ovarian, most

non-small lung cell, breast, and prostate cancers do not respond to initial cisplatin chemotherapy (Eckstein, 2011; Teply and Hauke, 2016). Both acquired and intrinsic resistance share common mechanisms mediating chemoresistance.

1.3.1 Mechanisms of ICL-Based Chemoresistance

Exposure of tumor cells to ICL-inducing chemotherapy (cisplatin, MMC, nitrogen mustards) frequently results in acquired chemoresistance. Given the widespread use of cisplatin, much effort has been directed toward understanding the mechanisms underlying resistance. Cisplatin resistance is mediated in two basic ways: reduced accumulation of ICL-inducing agent within the cell, and alterations in DNA damage repair and response (**Figure 1**B).

Cisplatin must be actively transported into the cell, thus downregulation of the copper transporter 1 (CTR1) gene reduces intracellular levels of cisplatin (Holzer and Howell, 2006). Cisplatin can be driven out of the cell via upregulated efflux pumps [e.g. ATP7A (copper-transporting P-type ATPase) and multidrug resistance-associated protein 2 (MRP2)] (Aida et al., 2005; Korita et al., 2010; Safaei et al., 2004). Where cisplatin remains in the cell, glutathione proteins (GSH/γGCS/GST) have been shown to sequester the electrophilic groups of activated cisplatin, while metallothioneins (MTs) bind to the heavy metal core of cisplatin for detoxification (Kasahara et al., 1991; Lewis et al., 1988).
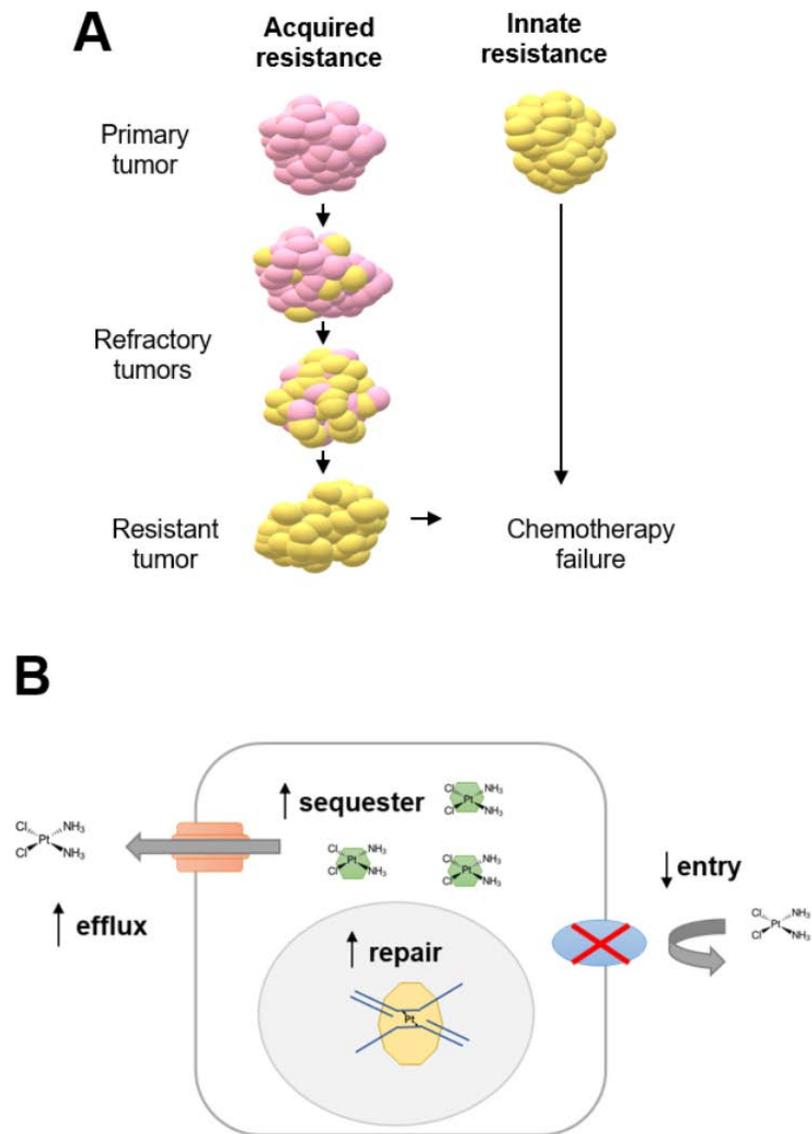
Figure 2: Mechanisms of chemoresistance

(A) Acquired resistance development or innate resistance in of tumors. **(B)** Cellular responses mediating cisplatin resistance.

Cisplatin that persists within the cell is able to damage DNA; however, these lesions can be actively repaired, thus expression of some ICL repair factors is highly correlated with cancer treatment response. Extensive research of excision repair cross-complementing 1 (ERCC1) and tumor response to cisplatin suggests that expression of ERCC1 may facilitate chemotherapy treatment planning in some malignancies (McNeil and Melton, 2012). Non-small cell-lung cancers highly express ERCC1 and are innately chemoresistant to cisplatin therapy, while cancers particularly responsive to cisplatin therapy (e.g. testicular cancer) possess low ERCC1 levels (Olaussen et al., 2006; Usanova et al., 2011). Similar cisplatin sensitivity has been observed for essential meiotic endonuclease 1 (EME1) in some cancers (Tomoda et al., 2009). Compared to primary tumors, corresponding refractory tumors show increased expression of ICL repair proteins (Galluzzi et al., 2012). This suggest that, given the heterogeneity of malignant cells within a tumour, cells with increased ICL repair proficiency evade chemotherapy remain, resulting in a subsequent non-responsive tumour (Eckstein, 2011). Thus, expression of proteins, like ERCC1 and EME1, can be used as predictors of tumor response to ICL-based chemotherapy, sparing patients from the significant side effects of chemotherapy when it would otherwise be ineffective (McNeil and Melton, 2012).

1.4 Significance of ICL Repair in Fanconi Anemia Syndrome

While reduced ICL repair proficiency is beneficial for treatment of tumors, defective ICL repair is detrimental in humans. The importance of ICLs arising from

endogenous sources is best exemplified by the autosomal recessive disorder, Fanconi anemia, resulting from mutations within proteins involved in ICL repair. Fanconi anemia syndrome (FA) is a complex disorder with clinical manifestations including bone marrow failure, progressive aplastic anemia, congenital anomalies, as well as high incidence of hematological and squamous cell cancers (Fanconi, 1967). ICL-induced genomic instability within the blood stem cell compartment results in the characteristic hematological defects of FA . Compared to the general population, the risk of developing leukemia is increased by more than 700-fold (Rosenberg et al., 2003). Diagnosis of FA includes analysis of chromosomes following MMC exposure, in which the presence of ICL-specific chromosomal aberrations (e.g. radial formations and isochromatid breaks) is a positive indicator of FA (Blood, 1988). Since all cells of patients with FA, including normal non-replicative cells, lack the ability to properly repair ICL damage, cancer treatments must be carefully considered and conservative approaches must be taken (Kutler et al., 2016). FA, therefore, demonstrates a requirement for repair of ICLs arising from both endogenous and exogenous sources.

1.5 DNA Repair Pathways Involved in ICL Repair

Cells have evolved specialized pathways to combat different types of DNA damage [e.g. ultraviolet (UV) radiation, reactive oxygen species, and ionizing radiation (IR)]. Unlike other types of DNA damage, which have dedicated pathways for their repair, ICL repair instead makes use of factors from different pathways. Indeed, the complexity of ICL repair is highlighted by the necessity to repair damage present on both

strands in the absence of undamaged template for synthesis (Clauson et al., 2013).

Moreover, the incisions required to remove ICL damage are capable of generating DSBs,

endangering the integrity of the genome (Hanada et al., 2006). ICL repair in all

eukaryotes makes selective use of proteins from three different DNA repair pathways:

nucleotide excision repair, translesion synthesis, and homologous recombination (Clauson

et al., 2013). Higher eukaryotes have also evolved a fourth pathway known as the Fanconi

anemia pathway (Deans and West, 2011). Depending on the stage of the cell cycle in

which the ICL is detected, different subsets of proteins are recruited for repair. Therefore,

to fully understand the repair of ICLs, an overview of the four pathways is necessary.

1.5.1   Nucleotide Excision Repair

Bulky adducts on one strand of DNA are formed by exposure to UV irradiation and

monoalkylating agents. Repair of such monoadducts is accomplished by the nucleotide

excision repair pathway (NER). The basic steps of this repair process involve excision

and removal of damage, repair synthesis of the gap, and ligation of the DNA. This

process somewhat differs depending on the point at which the lesion is recognized.

General surveillance of the entire genome for these lesions and their repair is initiated by

XPC-RAD23B in global genome NER (GG-NER) (Sugasawa et al., 1998). Alternatively,

when RNA polymerase stalls upon encountering damage, CSA, CSB, and XAB2 are

recruited to initiate transcription-coupled NER (TC-NER) (Fousteri and Mullenders,

2008). Downstream excision, synthesis and ligation proceed similarly regardless of the

point of lesion recognition. Post damage detection, TFIIH unwinds and opens the duplex

to verify the presence of DNA damaged. TFIIH is a ten-subunit complex consisting of three major domains: the core (containing XPB), the kinase (CAK) and linker (XPD). XPB ATPase activity pries open DNA, after which XPD helicase and ATPase activity translocates and further opens the DNA, respectively (Fuss and Tainer, 2011). Damage is verified by TFIIH stalling at the lesion. Scaffolding protein XPA, ssDNA binding protein RPA, and endonuclease XPG are recruited to form the pre-incision complex (Rechkunova et al., 2011). XPF-ERCC1 is last to arrive, but makes the first incision, acting on the 5' side of the bubble (Fagbemi et al., 2011). Incision on the 5' end generates a 3'OH that is used by the replication machinery for repair synthesis, as well as stimulates RPA to promote XPG cleavage on the opposite 3' side (Fagbemi et al., 2011). An oligonucleotide, spanning about 30 bases, is removed by TFIIH to completely excise the lesion from the duplex (Fuss and Tainer, 2011). In conjunction with RCF, Pol ε or δ, repair synthesis is carried out (Ogi et al., 2010). Pol κ has also been implicated in NER (Ogi et al., 2010). The processes is completed by ligation of the nicked DNA. NER makes highly important contributions to ICL repair by providing factors responsible for incision of crosslinked DNA.
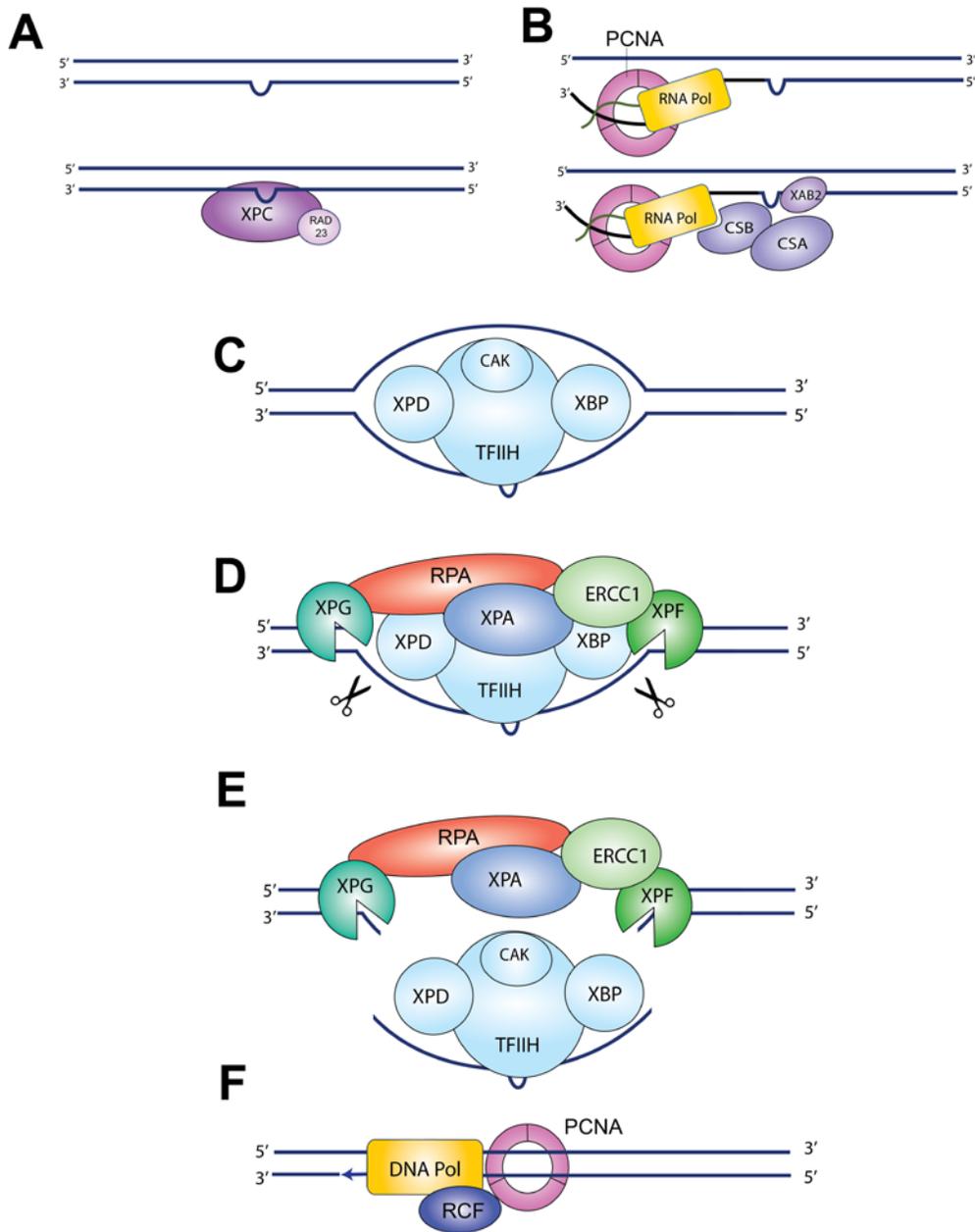
Figure 3: Nucleotide excision repair pathway

(**A**) Global genome recognition of lesion. (**B**) Transcription-coupled recognition of lesion. (**C**) TFIIH-mediated lesion validation. (**D**) Pre-incision complex assembly. (**E**) Coupled-incision and lesion removal. (**F**) Repair synthesis.

1.5.2 Translesion Synthesis

While larger bulky adducts are repaired by NER, other smaller lesions are not detected until the helix has begun unwinding. In this case, it is easier to bypass the lesion and continue synthesis, thereby eliminating the potential for replication fork collapse and possible DSB formation. A polymerase must bypass a lesion and insert a base, at times, in the absence of the second strand. Furthermore, such polymerase must only be utilized in the presence of DNA adducts, otherwise routine use of the polymerase would result in a significant mutation accumulation. Thus, the general mechanism of translesion synthesis (TLS) involves switching the replicative polymerase for a lesion tolerant polymerase(s) to move towards, incorporate a dNTP to bypass, then extend past a lesion.

With the exception of Pol ξ, TLS polymerases generally belong to the Y-family of polymerases, which include Pol η, Pol κ, Pol ι, and REV1(Vaisman and Woodgate, 2017). These Y-family polymerases contain larger active sites to accommodate lesions and it is the structure of the lesion that influences the activity of TLS polymerase. Highly distorting lesions are more easily accessed by a TLS polymerase since the duplex is already destabilized and unpaired, however these lesions are more difficult to bypass. On the other hand, approaching less distorting adducts tends to be more difficult for TLS polymerases, but bypass and extension past the adduct is easier (Roy and Schärer, 2016).

Central to TLS is DNA clamp, PCNA (proliferating cell nuclear antigen), which not only acts as a processivity factor during replication, but also recruits factors involved in DNA replication, repair, and chromatin remodeling (Maga and Hubscher, 2003). Recruitment can be mediated through two PCNA binding motifs, the PNCA-interacting

peptide (PIP)-box or AlkB homolog 2 PCNA interacting motif (APIM), PCNA (Gilljam et al., 2009). Interchange of polymerases also depends on interaction of their conserved ubiquitin zinc finger (UBZ) domains with monoubiquitinated of PCNA, a DNA damage response event mediated by RAD18-RAD6 (Davies et al., 2008). TLS therefore involves several proteins that together, synthesize DNA across an ICL adduct during ICL repair.
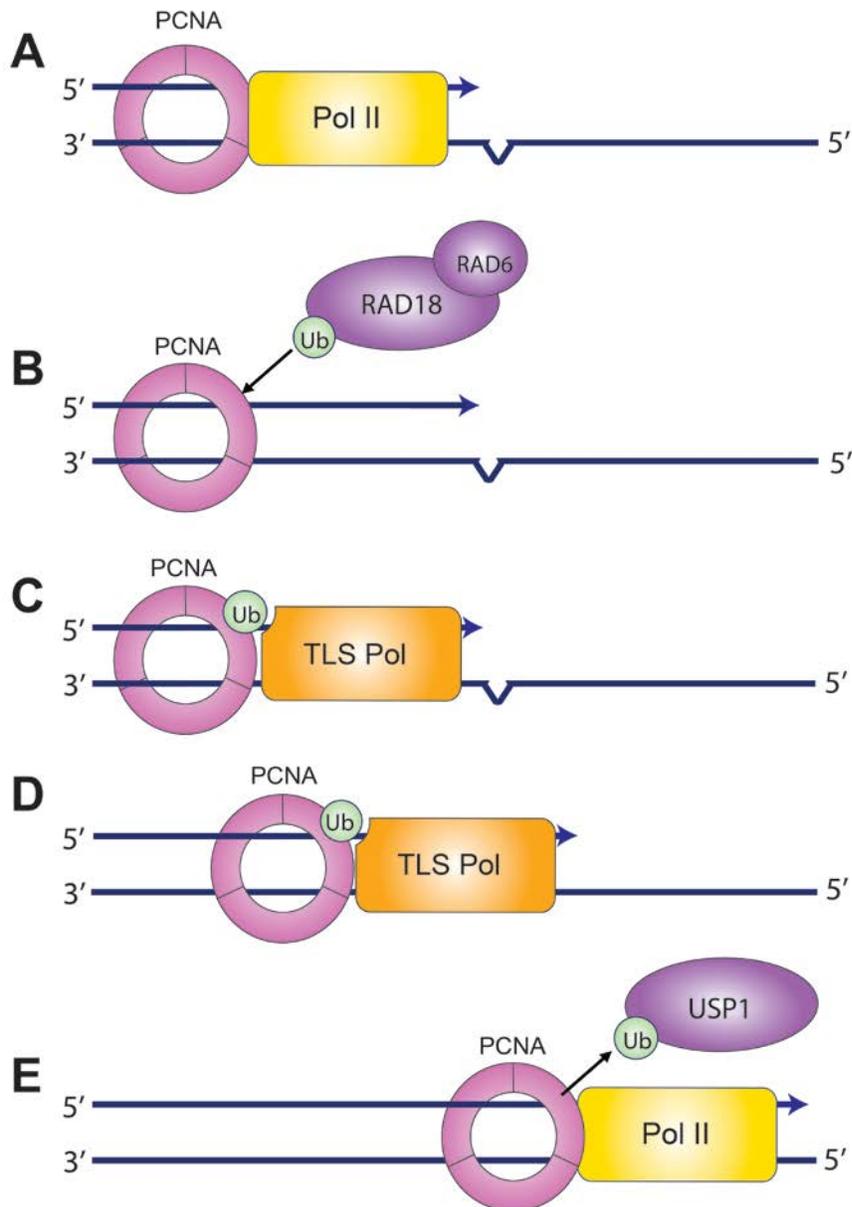
Figure 4. Translesion synthesis repair pathway

(**A**) Lesion detection. (**B**) PCNA monoubiquitination. (**C**) Polymerase switching. (**D**) Lesion bypass synthesis (**E**) Deubiquitination and synthesis continuation.

1.5.3 Homologous-Directed Double-Strand Break Repair

Double strand break (DSB) repair is an essential repair process since a single DSB can be lethal to the cell (Bennett et al., 1993). In mammalian cells, DNA ends are rapidly joined by the non-homologous end joining (NHEJ) repair pathway to avoid DNA deletions, inversions, or translocations (Chang et al., 2017). This process is quick and has been considered to be error prone (Rodgers and McVey, 2016). In contrast, homologous recombination (HR) repairs DNA with high fidelity by making use of a second strand as a template during repair and typically occurs only in the presence of an undamaged sister chromatid (Jasin and Rothstein, 2013). Progression into HR requires generation of ssDNA containing regions of homology that are exchanged to perform error-free repair.

DSBs are initially recognized by Ku, which with 53BP1, inhibits end resection and HR (Shao et al., 2012). In S and G2 phase however, DNA ends are resected by the Mre11 nuclease of the Mre11-Rad50-Nbs2 (MRN) complex to create ssDNA 3' overhangs needed for HR initiation (Symington, 2014; Trujillo et al., 1998). Endonuclease CtIP, which also resects DNA, is recruited by BRCA1 (Yun and Hiom, 2009). BRCA1 also binds PALB2, required for BRCA2 and RAD51 recruitment (Huen et al., 2010). Resected DNA is coated by RPA, which is displaced by BRCA2 to mediates nucleofilament formation by recombinase RAD51 (Jasin and Rothstein, 2013). Strand invasion by RAD51-ssDNA 3'overhang carries out the search for complementary DNA to permit repair synthesis from the 3'OH primer (Holloman, 2011). Extension from the undamaged strand is thus the reason for HR repair accuracy. The Holliday junction (HJ) formed by strand invasion is disassembled by the BTR (BLM-TOPIIIA-RMI1-RMI2)

complex (West et al., 2015). In the absence of the BTR complex, HJ resolution is accomplished by structure-specific endonucleases including GEN1 or SLX1-SLX4 with MUS81-EME1 (West et al., 2015). Subsequent ligation completes the repair process. In ICL repair, HR is needed during replication to repair the DNA breaks generated while removing the adduct for lesion bypass.
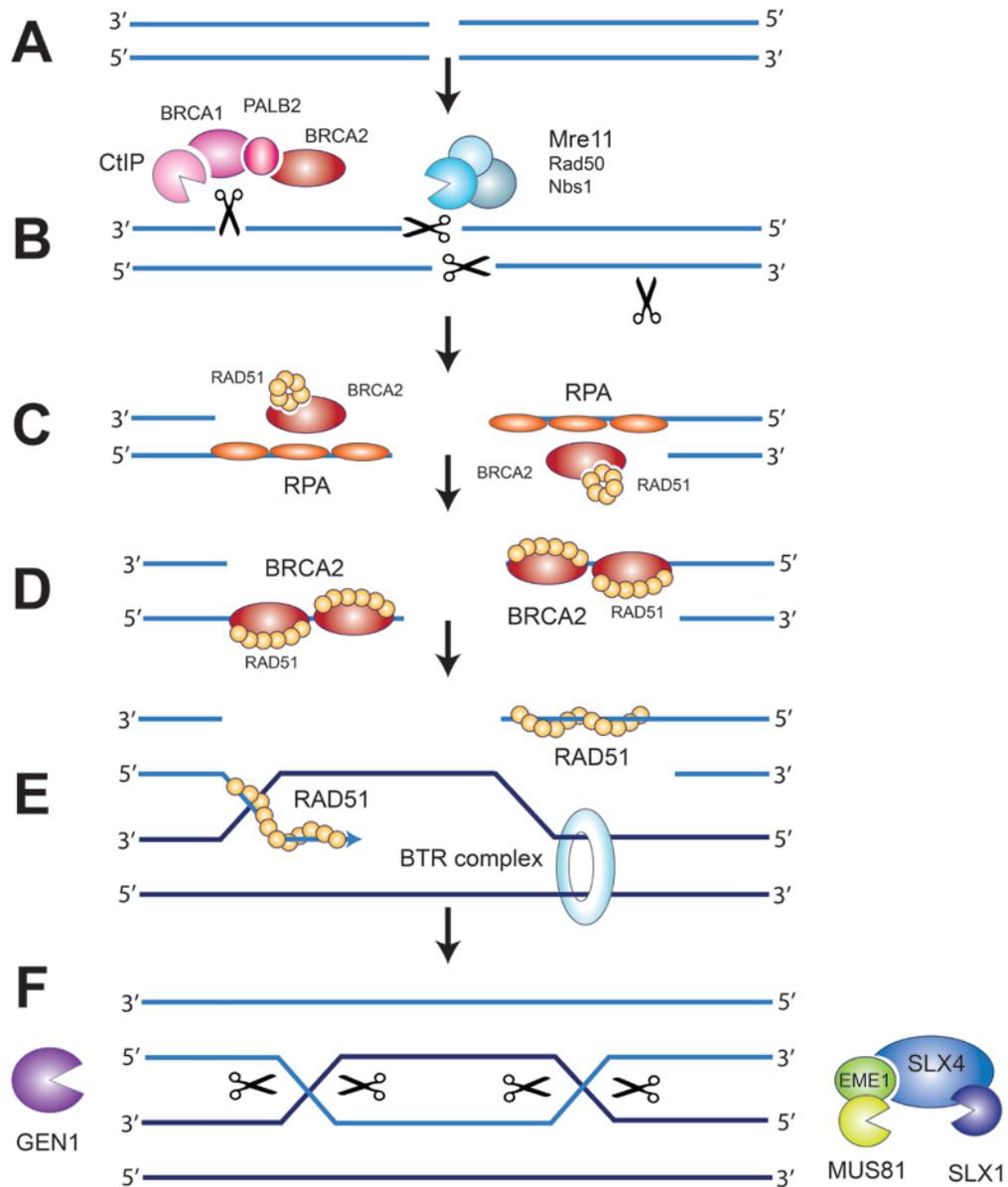
Figure 5: Homologous recombination repair pathway

(A) Double-strand break. **(B)** 5' end incision and resection. **(C)** RPA single-strand binding. **(D)** Nucleofilament formation. **(E)** Strand invasion. **(F)** Holliday junction resolution.

1.5.4 Fanconi Anemia Pathway

The Fanconi anemia pathway was thought to be activated in response ICL damage during replication, but may more generally respond to replication fork stress and collapse. The general mechanism consists of damage recognition, assembly of an E3 ligase for monoubiquitination of proteins to recruit structure-specific endonucleases. Recognition of the collapsed forked by FANCM recruits nine core FA factors, FANCA, FANCB, FANCC, FANCE, FANCF, FANCG, FAAP20, FAAP100 and FANCL (Niedernhofer, 2007). The assembled complex uses the E3 ligase activity of FANCL to perform the key FA event: monoubiquitination of heterodimer FANCD2-FANCI (ID2 complex) (Gregory et al., 2003). This monoubiquitinated ID2 complex, ID2(Ub), not only interacts with DNA more strongly, but also serves to recruit structure-specific endonucleases either directly, or though binding of the SLX4 scaffolding protein (Cybulski and Howlett, 2011). ID2(Ub) and SLX4 likely license and direct incision, resection, and HJ resolution by coordinating nucleases during repair. Deubiquitination is also a crucial event in the FA pathway, where disruption of USP1-UAF1 deubiquitinase results in ICL repair defects despite elevated levels of monoubiquitination (Oestergaard et al., 2007). During replication, FA functions to recognize ICL damaged forks and coordinates nucleases via monoubiquitination.
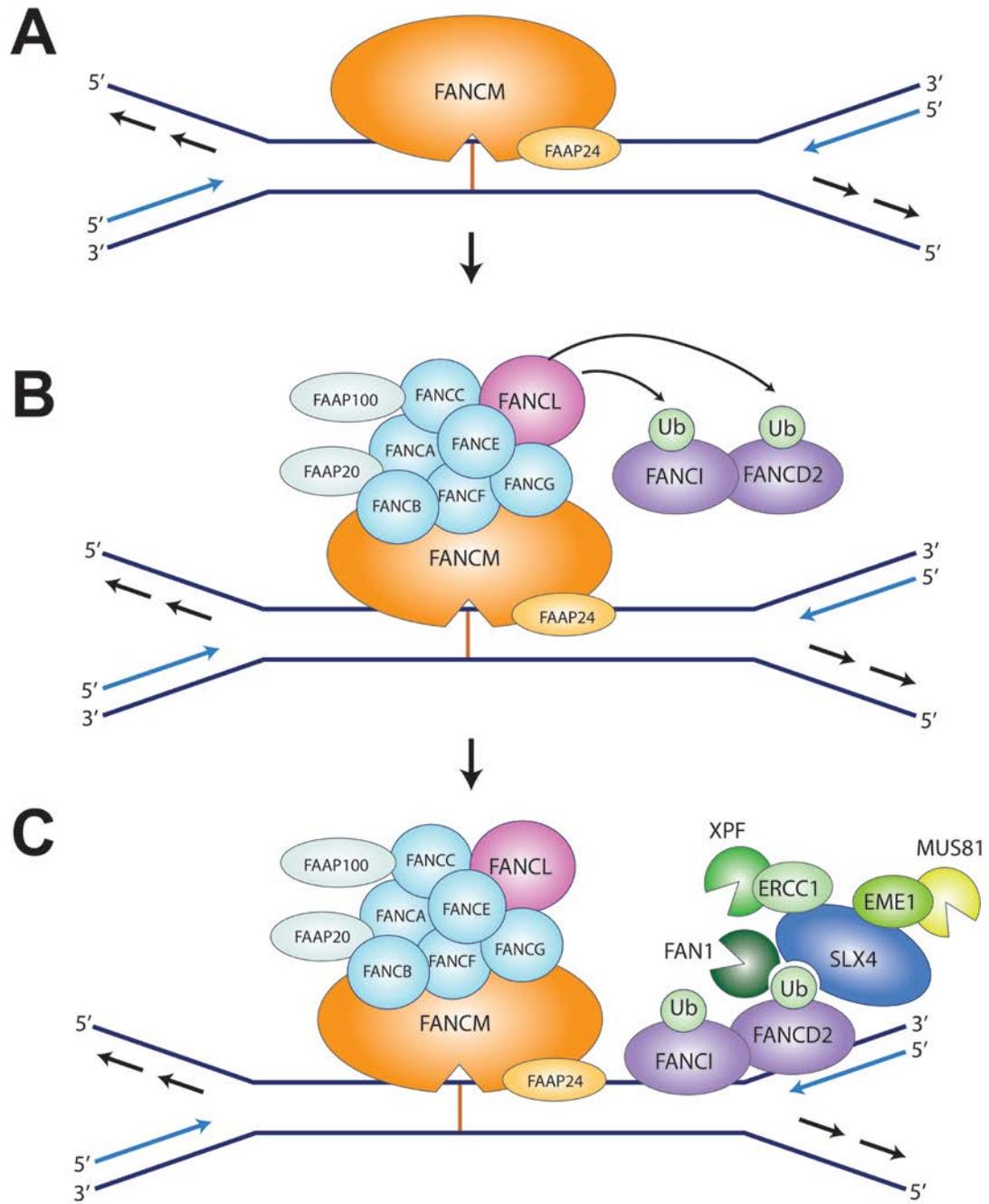
Figure 6: Fanconi anemia repair pathway

(A) Damage recognition. **(B)** FA complex assembly and FANCD2-FANCI monoubiquitination **(C)** Nuclease recruitment.

1.6   Interstrand Crosslink Repair Pathways

Although all ICL-inducing agents covalently link opposing strands of DNA, not all ICLs distort the helical backbone in the same manner. These differences in structural aberrations have implications in how and when an adduct is recognized and repaired (Dronkert and Kanaar, 2001). Highly distorting lesions are more likely to be detected during all stages of the cell cycle given that they are easily recognized within the duplex, similar to to GG-NER (Smeaton et al., 2008). ICL repair can also occur during transcription when the cell is forced to handle the ICL blocking strand separation. Less distorting ICLs may escape detection during most of the cell cycle, however a crosslinked replication fork is unambiguously damaged and forces a DNA repair response. The general mechanism for ICL repair requires adduct excision to permit bypass during repair synthesis and restoration of a DSB created during replication (if necessary). Therefore, NER excision, TLS repair synthesis, HR repair, all coordinated by the FA pathway are required to fully restore DNA damage from ICLs. Unique to ICL repair however, is the requirement to generate a suitable template from DNA damaged on *both* sides. To accomplish this, incisions are required flanking the DNA in a process referred to as "unhooking". This unhooked substrate is further "trimmed" by resection to facilitate the ICL damaged nucleotide to be "flipped out", or moved away from the gap for lesion bypass and repair synthesis (Noll et al., 2006).

1.6.1 Replication-Dependent ICL Repair

Since ICLs are an absolute impediment to strand synthesis, the dominant ICL repair pathway is considered to be replication-dependent. As such, ICLs are thought to be most commonly detected by the blocked replisome, although this bias towards replication-dependent repair may instead reflect an inability to monitor the extent of ICL damage and repair in other phases of the cell cycle. Using Xenopus extracts, it was found that two converged replication forks signal for ICL repair. Both leading strands loaded with Cdc45-Mcm-GINS (CMG) helicase complexes converge at the ICL (Zhang et al., 2015). Complex on one leading strand at the fork dissociates, allowing polymerase progression towards the crosslink. Recognition of the stalled replication fork by the FA complex and validation of the presence of an ICL results in monoubiquitination of the ID2 complex to recruit structure-specific endonucleases for unhooking. XPF-ERCC1 has been shown to be responsible for making the initial incision, where an incision indicative of unhooking is absent with depletion of XPF-ERCC1 in Xenopus extracts (Klein Douwel et al., 2014). In the event that XPF-ERCC1 is unable to make this incision, it is thought that MUS81-EME1 instead induces the ICL-dependent DSB, possibly in attempts to rescue the collapsed replication fork (Wang et al., 2011a). The nuclease responsible for making the second incision to unhook the ICL and precise location of incisions remains unclear (Zhang and Walter, 2014a). For TLS to efficiently bypass the ICL, trimming of DNA flanking the adduct is required to increase access by lesion tolerant polymerase. Exonuclease SNM1A may be employed for such a task, particularly since this nuclease

reportedly has activity which can accommodate, like TLS polymerases, and digest damaged DNA (Wang et al., 2011a). Once the ICL is flipped out, it is thought that PCNA switches replicative polymerase for REV1, which inserts a base across the adducted nucleotide. The strand is further extended by Pol ξ (Budzowska et al., 2015). Since the sister chromatid is present during replication, the unhooked DNA or DSB is repaired via HR. As such, not only do mutations in FA factors cause the development of FA syndrome, some HR factor mutations have also been shown to result in the same disorder (Haynes et al., 2015).

1.6.2   Replication-Independent ICL Repair

While ICLs are an absolute block to replication requiring immediate resolution, mechanisms for repair must exist for non-replicating cells. Replication-independent repair (RIR), also referred to as FA-independent repair, is less understood, but has been shown that NER is intimately involved in RIR (Williams et al., 2013). Participation of NER factors specific to TC-NER or GG-NER (e.g. CSA/CSB or XPC, respectively) suggest at what point the ICLs are recognized and processed. Transcription machinery encountering and ICL adduct activates RIR in a manner similar to TC-NER. Additionally, similar to GG-NER, distortion of the DNA helix may enable direct detection of damaged DNA. *In vitro,* studies have shown direct binding of RPA and High Mobility Group protein 1 (HMG1) to cisplatin ICLs (Park and Lippard, 2011; Patrick et al., 2008). XPC-hHR23B with RPA-XPA, and mismatch repair protein hMutSβ, have been shown to bind to psoralen crosslinks (Zhao et al., 2009).

The central excision event of NER by XPF-ERCC1 is thought to be required for unhooking of ICLs in RIR. While XPG is required to make the second incision in NER, XPG depletion does not result in significant ICL hypersensitivity (Wood, 2010). This suggests another nuclease is required to complete unhooking in RIR. Similar to replication-dependent ICL repair, the nuclease which makes the other incision is not known (Williams et al., 2012). Monoubiquitinated PCNA recruits either error-prone TLS Pol ζ or Pol κ to bypass the ICL adduct and complete the repair of one strand of damaged DNA (Ho & Scharer, 2010). To fully remove the unhooked ICL lesion, the monoadducted DNA is repaired via NER.

1.7   Structure-Specific Endonucleases in ICL Repair

As discussed, ICL repair requires several coordinated nuclease events to excise and trim DNA surrounding the adduct. Since both strands of DNA are crosslinked and simultaneously damaged, nucleases are fundamental in the repair of ICLs. It is therefore not surprising that many nucleases have been implicated in ICL repair, particularly given the complexity of repair spanning all phases of the cell cycle. While it is possible that each distinct step of repair is carried out by a single nuclease, it is more probably that several nucleases are redundant for the same step or used for multiple steps. Most likely, the repair of ICLs does not follow a single pathway, but instead responds with several coordinated attempts to remove the lesion. Some nucleases appear to participate exclusively in ICL repair, while others are involved in the repair of other DNA lesions. Nonetheless, all of them collectively form a set of DNA scissors which the cell uses for

survival against the destructive effects of ICLs.

One of the unresolved questions in ICL repair is which nucleases are required to unhook the ICL. Strand incision is a crucial event, as it commits the cell to the repair process. While there are several good candidates able to cleave 3'flaps, the second nuclease that fully unhooks the ICL is unknown. Implicated nucleases for unhooking in mammalian ICL repair include XPF-ERCC1, MUS81-EME1, SLX1-SLX4, FAN1, and DNA β-CASP nuclease, including SNM1A.

1.7.1 XPF-ERCC1

The best studied nuclease in ICL repair is the heterodimer XPF-ERCC1. This structure-specific endonuclease makes the first of two incisions required to excise damage in both GG-NER and TC-NER. Mutations in human XPF-ERCC1 result in several genetic disorders. Xeroderma pigmentosa (XP) is characterized by an extreme sensitivity to UV while Cockayne syndrome (CS) results in neurodegeneration, poor growth, photosensitivity and premature aging (Kashiyama et al., 2013). Mutation in XPF (also known as FANCQ) also causes FA (Bogliolo et al., 2013). $ERCC1^{-/-}$ or $XPF^{-/-}$ mice have similar defects as XP and CS, but die within less than a month after birth (Kirschner and Melton, 2010). XPF-ERCC1 appears to also act in DSB repair by removing non-homologous 3' ends before end joining (Ahmad et al., 2008). In DT40 chicken cells, disruption of XPF-ERCC1 results in increased chromosomal breaks in response to camptothecin (CPT), a topoisomerase inhibitor which induces spontaneous homologous recombination (Kikuchi et al., 2013). Thus XPF-ERCC1 may play a role in homologous

recombination as well.

Unlike other nucleases involved in ICL repair, depletion of XPF-ERCC1 produces significant hypersensitivity in response to a broad number of ICL-inducing agents. Given the numerous roles XPF-ERCC1 plays in other DNA repair pathways, it is possible that either XPF-ERCC1 is involved in more than one step in ICL repair, or that XPF-ERCC1 is required to repair other types of damage (e.g. monoadducts or intrastrand crosslinks) formed by ICL-inducing agents. XPF-ERCC1 is pivotal to ICL repair and it is unsurprising that expression of ERCC1 or XPF negatively correlates with response to ICL-based chemotherapy (McNeil and Melton, 2012). XPF-ERCC1 expression is correlated in clinical outcomes in head and neck squamous, testicular, and bladder cancers (Kirschner and Melton, 2010). As such, it may be used as a biomarker for ICL-based chemotherapy response and relapse.

XPF-ERCC1 is part of the PD-(D/E)XK superfamily containing the conserved VERKX$_3$D active site within the ERCC4 domain of XPF (Hanada et al., 2006; Kosinski et al., 2005). XPF-ERCC1 has been shown to act on the 3' end of a single to double-strand DNA junction (Nishino and Morikawa, 2002). In NER, XPF-ERCC1 cuts the 5' end of the bubble formed to remove the damaged oligonucleotide. Its role in DSB repair involves removal of 3'flaps of non-homologous DNA (Ahmad et al., 2008). Recent work using Xenopus extracts has demonstrated that immunodepletion of only XPF-ERCC1, and no other endonucleases, disrupts formation of an unhooked substrate in plasmids containing a single ICL (Klein Douwel et al., 2014). This indicates that XPF-ERCC1 is required for incision to initiate replication-dependent repair. Furthermore, this activity

requires FANCD2 monoubiquination, a key event in FA-dependent repair (Klein Douwel et al., 2014). Although XPF-ERCC1 preferentially cleaves the 3' end of splayed arm, bubble and stem-loop DNA structures *in vitro*, it has more recently reported to be able to cut within duplex DNA 5' to an ICL in the presence of RPA (Abdullah et al., 2017; De Laat et al., 1998). Therefore, the position of XPF cleavage in ICL repair is unclear.

1.7.2 MUS81-EME1

Methanesulfonate, UV-sensitive protein 81 (MUS81) and essential mitotic endonuclease 1 (EME1) are best known for their role in HJ resolution and replication restart. Cells defective in MUS81-EME1 result in chromosomal breaks, fusions, and tri-radials (Hanada et al., 2006). While MUS81-EME1 is a poor HJ resolvase compared GEN1, MUS81-EME1 has robust activity on nicked HJs (Wyatt et al., 2013). Thus, in conjunction with endonuclease SLX1-SLX4, MUS81-EME1 asymmetrically cuts a HJ, producing both crossover and non-crossover events during mitosis and G2. In addition to mild hypersensitivity to ICL-inducing agents, MUS81-EME1 mutants are also hypersensitive to agents that induce replication fork stalling and collapse, including UV damage, methyl methanesulfonate (MMS), hydroxyurea (HU), and CPT (Sarbajna and West, 2014). In the absence of MUS81-EME1, cells accumulate DSBs in response to damage, suggesting that replication fork cleavage by MUS81-EME1 may subsequently feed into HR repair (Sarbajna and West, 2014; Wang et al., 2011a). MUS81 directly interacts with SLX4 recruited by FANCD2(Ub), as well as FANCA, suggesting a function in replication-dependent ICL repair (Benitez et al., 2014; Wyatt et al., 2017).

MUS81, also a PD-(D/E)XK family member, shares the same ERCC4 domain and VERKX$_3$D active site as XPF (Hanada et al., 2006; Kosinski et al., 2005). Although both MUS81 and XPF can carry out similar functional activities, they differ in their efficiency to act on various DNA substrates (Ciccia et al., 2008). It has been reported that in the absence of MUS81-EME1, with or without XPF-ERCC1, formation of DSBs does not occur during S phase (Wang et al., 2011a). Moreover, γH2AX signaling is absent, suggesting that MUS81-EME1 may represent a later response to ICL damage. It is unclear if the defect lies in the role of MUS81-EME1 in replication fork restart or if it occurs later during homologous recombination.

1.7.3 SLX1-SLX4

SLX1 is required, in partnership with SLX4, to make the first nick for MUS81-EME1 in HJ resolution (Wyatt et al., 2013). Mice lacking SLX1 show moderate sensitivity to ICL-inducing agents, suggesting a role in ICL repair (Castor et al., 2013). SLX1 is part of the RecQ helicase family, containing a UvrC intron-endonuclease domain, and its nuclease activity is stimulated by SLX4 interaction (Zhang and Walter, 2014b). Unlike XPF-ERCC1 and MUS81-EME1, SLX1-SLX4 has more promiscuous endonuclease activity, cutting on both sides of a splayed arm and flap substrate, with slight preference for 5' flap cleavage (Fricke and Brill, 2003). This activity thus makes SLX1-SLX4 a good candidate for cleavage of the 5' side of an ICL (Zhang and Walter, 2014b). At this time, studies on ICL hypersensitivity of SLX1-SLX4 mutants in cells are lacking.

1.7.4   FAN1

Fanconi anemia associated nuclease 1 (FAN1) of the ICL repair Fanconi anemia pathway demonstrates endonuclease activity on 5' flap structures and replication forks (Jin and Cho, 2017). Depletion of FAN1 results in hypersensitivity to ICL-inducing agents, resulting in chromosomal instability (Yoshikiyo et al., 2010). However, microdeletions in the FAN1 gene do not result in the FA phenotype (Trujillo et al., 2012). Mutations in FAN1, however, have been shown to cause karyomegalic interstitial nephritis (Zhou et al., 2012). FAN1 defects are additive, not epistatic, with FA pathway defects suggesting that FAN1 is not strictly required for ICL repair (Yoshikiyo et al., 2010).

The PD-(D/E)XK active site of FAN1 lies within the VRR-NUC (virus replication repair nuclease) domain, and is the only known human protein containing this VRR-NUC domain (MacKay et al., 2010). Biochemically, FAN1 possesses 5'flap endonuclease activity and 5'-3' phosphate-dependent exonuclease activity (Liu et al., 2010; MacKay et al., 2010). Recent crystal structures show that the phosphate binding pocket and catalytic metal binding site are spaced approximately four nucleotides apart within the active site of FAN1, resulting in a distributive nuclease mechanism that has been observed to bypass an ICL. Depending on placement of the ICL and context of the duplex region, FAN1 is able to create a nick past the ICL to bypass an ICL as a translesion nuclease.

1.8   β -CASP Nuclease Family

SNM1A belongs to the β-CASP family of nucleases comprised, but not limited to CPSF-73, Artemis, SNM1 and Pso2. Nucleases of the β-CASP family share a metallo β-lactamase (MBL) domain containing the HxHxDH catalytic site and β-CASP homology insertion cassette (Callebaut et al., 2002). Found across all domains of life, these nucleases share 5'phosphate-dependent nuclease activity and often endonuclease activity (Fernandez et al., 2011). Of the β-CASP nucleases which cleave DNA, SNM1A, SNM1B, and SNM1C are human paralogs originating from yeast ortholog SNM1(or Pso2) (Callebaut et al., 2002). All DNA β-CASP nucleases have been implicated in some aspect of DNA repair.

1.8.1   Pso2

Identified nearly 40 years ago, *S. cerevisiae pso2* mutants demonstrate diminished survival in the presence of ICL-inducing agents, but not any other DNA damaging agents (Henriques and Moustacchi, 1980). This suggested that PSO2 is one of the few genes in budding yeast uniquely involved in ICL repair, as other phenotypic defects have not been observed. Biochemical activity of Pso2 was not reported for another 25 years, however it is clear now that Pso2 has robust 5'-3' exonuclease activity that is dependent on a 5' phosphate (Li et al., 2005). Further characterization in our lab demonstrated that Pso2 also possess hairpin-opening endonuclease activity, with the ability to act on hairpin intermediates of transposable elements within cells (Tiefenbach and Junop, 2012). With regards to ICL repair, it has been suggested that this hairpin-opening activity may play a
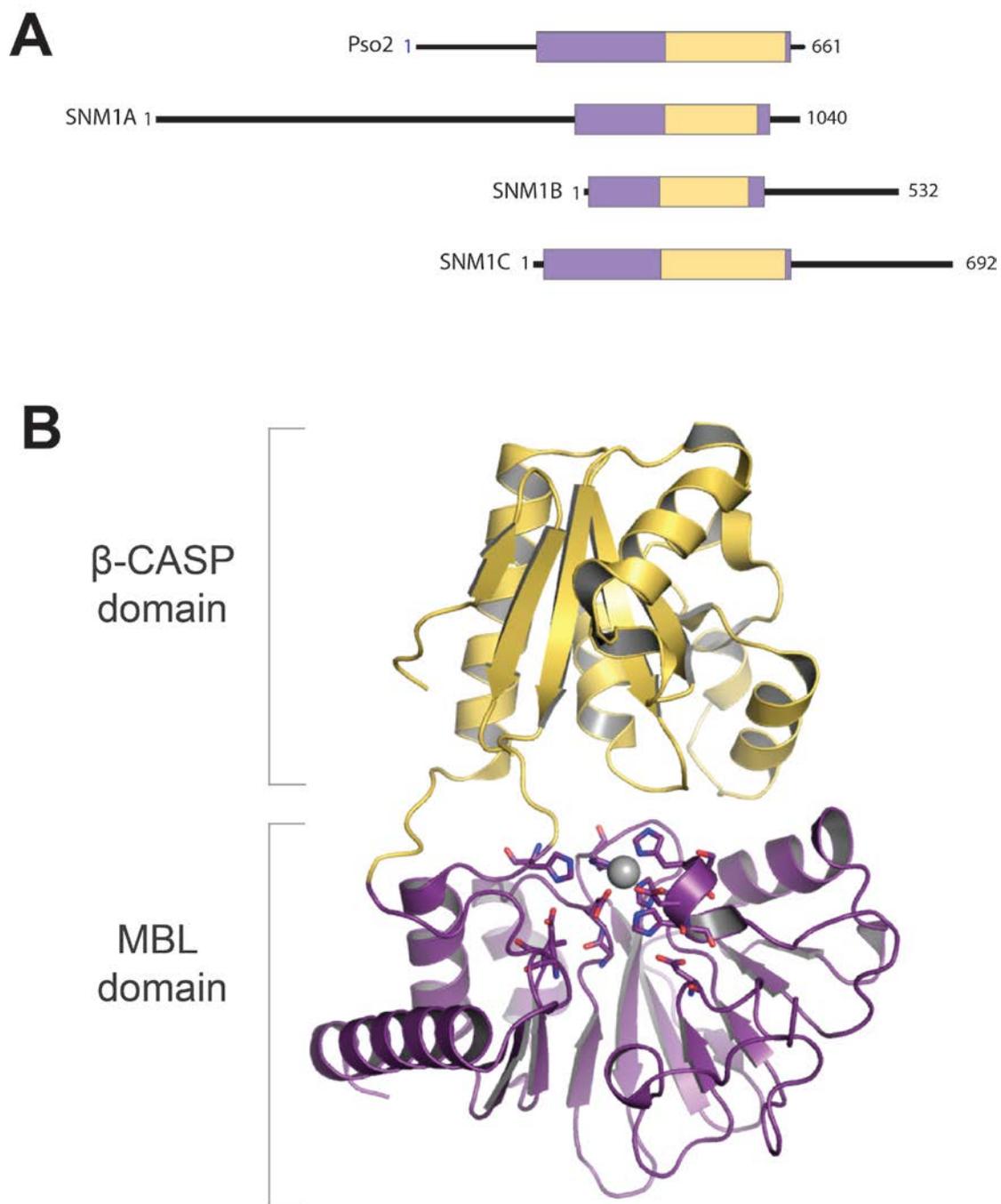
Figure 7: β-CASP nuclease family

(**A**) Alignment of DNA β-CASP nucleases. (**B**) Metallo beta-lactamase and β-CASP subdomains (PDB ID: 4B87).

role in cruciform resolution that may form during ICL repair, however the extent of this cruciform formation during repair is unclear (Brendel et al., 2003). Given that Pso2 has exonuclease and endonuclease activity, it remains possible that it might function in multiple steps of ICL repair.

1.8.2   SNM1B

SNM1B (also known as Apollo) is best known for its role in telomere maintenance, where it promotes t-loops formation and protection of chromosomal DNA from DSB repair. SNM1B directly interacts with TRF2, a component of the Sheltrin complex (Demuth et al., 2008). Splice variants of SNM1B that lack the TRF2-interacting region result in Hoyerall-Hreidarsson syndrome, characterized by bone marrow failure, immunodeficiency, and premature aging (Touzot et al., 2010). In addition to telomeric dysfunctions, depletion of SNM1B results in hypersensitivities to MMC, cisplatin, and IR in mice, DT40 chicken, and human cells (Demuth et al., 2004; Ishiai et al., 2004). SNM1B has been implicated in the Fanconi anemia pathway of ICL repair via its interaction with SLX4 and FANCD2 (Bae et al., 2008; Mason and Sekiguchi, 2011). Additionally, SNM1B interacts with other nucleases required for DNA repair, including MRE11 and MUS81(Bae et al., 2008). Given the central role of FANCD2 in nuclease coordination, the hypersensitivity of SNM1B may suggest a role downstream of unhooking. In the absence of SNM1B, DSBs do not accumulate, whereas wild-type cells result in DSB formation in response to MMC. Indeed, loss of SNM1B results in impaired HR repair in response to MMC as RAD51 and BRCA1 filament formation is disrupted,

suggesting nuclease activity of SNM1B may be required for HR. Biochemically, SNM1B possess 5'phosphate-dependent 5'-3' exonuclease activity, preferentially acting on dsDNA (Sengerová et al., 2012). Endonuclease activity has not been reported for SNM1B. Given this, it is possible that SNM1B is required for DNA resection to create 3'overhangs for HR strand invasion, consistent with its postulated role in generating 3'overhangs for telomere protection. However, the specific mechanism of SNM1B in the described pathways remains unclear.

1.8.3   SNM1C

SNM1C (also referred to as Artemis) is the best well-characterized β-CASP DNA nuclease. SNM1C is required for opening DNA hairpins formed during in V(D)J recombination (Niewolik et al., 2006). Thus, mutations in SNM1C result in severe combined immunodeficiency (SCID). Such SCID patients are highly sensitive to IR, indicating impairments in DSB repair, specifically NHEJ (Buckley et al., 1997). Indeed, SNM1C interacts with NHEJ kinase DNA-PKcs to process DNA ends prior to ligation. Phosphorylation by and association with DNA-PKcs relieves autoinhibition by SNM1C to permit structure-specific endonuclease activity on a variety of DNA substrates (Goodarzi et al., 2006). Like other β-CASP family members, SNM1C has been reported to possess 5'-3' exonuclease activity, however, this activity is much weaker and has been suggested to be a result of contaminating nuclease activity since mutations to the endonuclease catalytic residues of SNM1C do not abolish exonuclease activity (Pannicke et al., 2004a; Pawelczak and Turchi, 2010a). However, since the exonuclease activity

requires a 5'phosphate and is abolished in catalytic mutants expressed in insect cells, it is likely that exonuclease activity is intrinsic to SNM1C (Li et al., 2014). The β-CASP domain of SNM1C (1-385) has been shown to suffice for V(D)J recombination, but it is insufficient to repair DNA damage by IR (Poinsignon et al., 2004a). This may be due to its requirement for phosphorylation and association with DNA-PKcs. It is thought that the C-terminus of SNM1C acts as a negative regulator of SNM1C (Niewolik et al., 2017). Truncation of SNM1C permits hairpin-opening activity, however it is not clear if hairpin-opening activity is required for the repair of double strand breaks during NHEJ.

1.8.4   MBL/β-CASP RNases

In addition to DNA substrates, the MBL/β-CASP domain has also been shown to process RNA. Cleavage and polyadenylation specificity factor 73 (CPSF-73) is a 5' mRNA processing nuclease (Mandel et al., 2006). CPSF-73[1-460] has both endonuclease and 5'-3' exonuclease activity, sufficient for mRNA processing *in vitro* (Mandel et al., 2006). Similarly, RNaseJ possesses the same RNA processing activities, however a 5'phosphate is not strictly required for exonuclease activity .

1.8.5   Structure of MBL/β-CASP Nuclease Domain

Members of the β-CASP family contain a conserved catalytic core encompassing a highly conserved metallo β-lactamase (MBL) active site and homology domain shared amongst the family (Cattell et al., 2010). The larger MBL superfamily hydrolyzes a wide variety of substrates, most having an ester linkage and negative charge (Dominski, 2008). This includes the β-lactam ring of antibiotics like penicillin, which are hydrolyzed by β-

lactamases that mediate antibiotic resistance (Carfi et al., 1995). The MBL fold is comprised of a four-layered β-sandwich flanked by four α-helices, with zinc binding residues of the active site at the edge of the sandwich (Carfi et al., 1995). Despite the sequence diversity among MBL-containing proteins, the entire superfamily contains the highly conserved HxHxDH motif which coordinate zinc ions within the active site required for hydrolysis (Callebaut et al., 2002). Four additional motifs facilitate metal coordination either directly, or indirectly by positioning the residues that directly interact with metal ions (Callebaut et al., 2002).

The β-CASP domain appears to be an insertion into the MBL domain, as the full MBL domain requires additional secondary structural elements C-terminal of the β-CASP domain (Fernandez et al., 2011). As two separate domains, it is thought that the CASP domain acts as an exo-site which interacts with DNA to clamp over the MBL active site for hydrolysis, collectively considered the MBL/β-CASP domain (Callebaut et al., 2002). From crystal structures of RNA-bound RNaseJ (PDB ID: 4XWW), the phosphate backbone of the RNA substrate primarily makes ionic contacts with the β-CASP domain (Zhao et al., 2015). Proximal to the active site, a hydrophobic residue of the β-CASP domain intercalates between two bases to stabilize substrate interaction and position the scissile phosphate for catalysis (Zhao et al., 2014). Finally, a phosphate binding pocket was identified for RNaseJ adjacent to the active site, supporting the biochemical requirement for a 5'phosphate for exonuclease activity, a characteristic feature of β-CASP nucleases (Zhao et al., 2015). Despite an impressive amount of structural data for the β-CASP family, definitive evidence of the precise mechanism of catalysis is still lacking.

1.9 SNM1A

1.9.1 SNM1A and Genomic Instability

Unlike SNM1B and SNM1C, diseases or disorders arising from mutations in SNM1A have yet to be identified, although SNM1A has been associated with cancer predisposition. A polymorphism within SNM1A has been shown to increase the risk of small cell lung carcinoma (SCLC) (Kohno et al., 2006). SNM1A expression levels have also been shown to be associated with the incidence of ovarian cancer (Wang et al., 2016).

Mice models of SNM1A deletion illustrate the role of SNM1A in protecting against genomic instability. $SNM1A^{-/-}$ mice are viable and fertile, however, manifest defects of genomic instability. Partial deletion of exon 2 results in $SNM1A^{-/-}$ mice with normal lifespans, but MEFs of these mice displayed defects of ICL repair, including MMC hypersensitivity and increased radial formation in response to ICL damage (Hemphill et al., 2008a). Disruption of exons 2-7 of SNM1A decrease lifespan in mice, but exhibited gender differences in mortality. Resulting mortality in female mice was due to tumorigenesis for female mice (suggesting cell cycle regulation defects), whereas male mice died from bacterial infections of the mandibular and preputial glands, suggesting SNM1A may be also involved in immunity in addition to cell cycle regulation (Ahkter et al., 2005). Proliferation defects of $SNM1A^{-/-}$ MEFs have been observed, however, of cells that proliferated, there was only a mild hypersensitive to MMC. Finally, $SNM1A^{-/-}$ mice generated by Dronkart et al. were found to be hypersensitive to high doses of injected

MMC (Dronkert et al., 2000). Again, a gender difference was observed, in which the same dose resulted in earlier and more severe lethality in female mice. As expected, these MEFs were hypersensitive to MMC, however, they were not hypersensitive to 8MOP+UVA, cisplatin, or melphalan. The basis of gender-dependent defects in *SNM1A*$^{-/-}$ mice, as well as the variability in ICL hypersensitivity, remains unclear.

1.9.2   Cellular Roles of SNM1A

1.9.2.1 SNM1A and ICL Repair

Research investigating SNM1A has mostly focused on its role in ICL repair. It was speculated that SNM1A is involved in ICL repair given its homology to yeast ICL repair nuclease Pso2. Complementation studies in *S. cerevisiae* demonstrated hypersensitivity of *pso2Δ* in response to cisplatin and MMC was only rescued with SNM1A, but not SNM1B or SNM1C expression, suggesting the functional mammalian homologue of Pso2 in ICL repair is SNM1A(Hazrati et al., 2008).

In DT40 chicken cells, *SNM1A*$^{-/-}$ cells have been shown to be hypersensitive to cisplatin and MMC, but not IR, suggesting SNM1A plays an important function in response to ICL damage (Ishiai et al., 2004). Further epistasis analysis with XRCC3 and FANCC disruptants suggests that SNM1A has a function distinct from both HR and FA pathways, respectively (Ishiai et al., 2004). SNM1A is reportedly non-epistatic with TLS factor RAD18 in DT40 cells, contrary to work that demonstrated recruitment of SNM1A in human cells requires RAD18-dependent monoubiquitination of PCNA (Ishiai et al., 2004; Yang et al., 2010). Indeed, SNM1A interacts with monoubiquitinated PCNA via its

PIP-box and UBZ domains in order to properly localize to sites of ICL damage (Yang et al., 2010).

SNM1A localization to sites of DNA damage is also mediated by its interaction with protein inhibitor of activated STAT protein 1 (PIAS1), a SUMO E3 ligase (Ishiai et al., 2004). Sumoylation is a key post-translational modification, known to modulate proteins in response to DNA damage (Jackson and Durocher, 2013). Interestingly, SLX4-UBC9 also possesses SUMO E3 ligase activity, acting on both itself and on XPF (Guervilly et al., 2015). It is unknown if SLX4 activity regulates SNM1A activity. Mutations within conserved metal coordinating residues of SNM1A results in decreased binding to PIAS1 and diminished survival after exposure to cisplatin (Ishiai et al., 2004).

Of particular note, siRNA depletion of SNM1A renders HeLa cells as hypersensitive to MMC and SJG-136 as siRNA depletion of ERCC1 (Wang et al., 2011). Moreover, SNM1A depletion results in greater hypersensitivity than MUS81 depletion, but less than SLX4 depletion, suggesting that SNM1A and ERCC1 work together in a pathway distinct from MUS81 (Wang et al., 2011). This is supported by evidence that SNM1A and/or depletion of ERCC1 leads to MUS81-induced DSBs in response to ICL damage (Wang et al., 2011).

1.9.2.2 SNM1A in DNA Damage Response and Cell Cycle Maintenance

In addition to involvement in ICL repair, SNM1A has also been implicated in cell cycle regulation in the presence of other DNA damaging agents. Although SNM1A inactivating mutants are not hypersensitive to IR damage, damage by IR modulates the localization of SNM1A in the cell. SNM1A forms a nuclear body within the cell which

disperses into distinct foci upon DNA damage ICL-inducing agents and IR (Richie et al., 2002). SNM1A colocalizes with early DSB response factors (e.g. 53BP1, γH2AX, MRE11 and to a lesser extent, BRCA1) which mediate DSB repair pathway choice into HR or NHEJ in response to IR damage (Richie et al., 2002). Of these proteins, SNM1A has been shown to directly associate with 53BP1 *in vitro* (Richie et al., 2002). Although SNM1A foci colocalize with 53BP1 before and after IR damage, SNM1A foci formation occurs in the absence of 53BP1 and vice versa (Richie et al., 2002). Foci formation does, however, depend on ataxia telangiectasia mutated (ATM), which has been demonstrated to phosphorylate SNM1A. SNM1A does not localize to UV-laser micro-irradiation in ATM depleted HT1080 cells or Ataxia Telangiectasia (AT) cells suggesting that SNM1A is involved DNA damage signaling after activation of the DNA damage checkpoint (Akhter and Legerski, 2008). This is supported by evidence that downstream target activation of genomic guardian p53 by both 53BP1 and ATM signaling is diminished in cells deficient in SNM1A (Akhter and Legerski, 2008). Also, cells depleted of SNM1A do not arrest in G1 after IR damage (Akhter and Legerski, 2008). Thus, SNM1A may help regulate the G1 checkpoint arrest by signaling to p53, resulting in either the arrest of cell cycle progression or apoptosis initiation.

SNM1A has been shown to regulate an early mitotic checkpoint induced by spindle poisons. *SNM1A$^{-/-}$* MEFs have increased tetraploidy and increase micronuclei in response to spindle poisons (e.g. nocodazole, colchicine, and taxol) (Akhter et al., 2004). At the cellular level, chromosomes do not condense, and cells exit the G1 phase prematurely as shown by supernumerary centrosomes, prolonged histone H3

phosphorylation, and early cyclin A diminishment (Akhter et al., 2004). This aberrant G1 checkpoint is likely mediated through the interaction of SNM1A with the anaphase promoting complex (APC). It is thought that SNM1A negatively targets APC to arrest cell cycle progression upon exposure to mitotic stress prior to condensation of the chromosome.

There is evidence that SNM1A may also be involved in initiating apoptosis in response to extensive etoposide damage. Etoposide-induced DSBs are either repaired by NHEJ or trigger apoptosis (Abe et al., 2008). *SNM1A$^{-/-}$* DT40 cells show diminished apoptosis and defective apoptotic DNA fragmentation, implicating SNM1A in cell death signaling (Hosono et al., 2011). However, the milder defect in DNA fragmentation of SNM1A compared to SNM1B and SNM1C suggests that SNM1A may play a less prominent role in this regard.

Given the crucial role that SNM1A has in response to DNA damage as well as cell cycle regulation, expression of SNM1A is highly regulated throughout the cell cycle. SNM1A has a 900bp 5' untranslated region (UTR) that contains an internal ribosome entry site (IRES) to promote 5'cap independent translation of SNM1A (Zhang et al., 2002). Moreover, translation is upregulated during mitosis, increasing in late prophase to early telophase compared to the rest of the cell cycle (Zhang et al., 2002). It is thought that IRES-mediated translation is important for cell-fate decisions, particularly in response to damage (Komar and Hatzoglou, 2014). It is likely that cap-independent translation of SNM1A, and thus SNM1A, is important for mitosis.
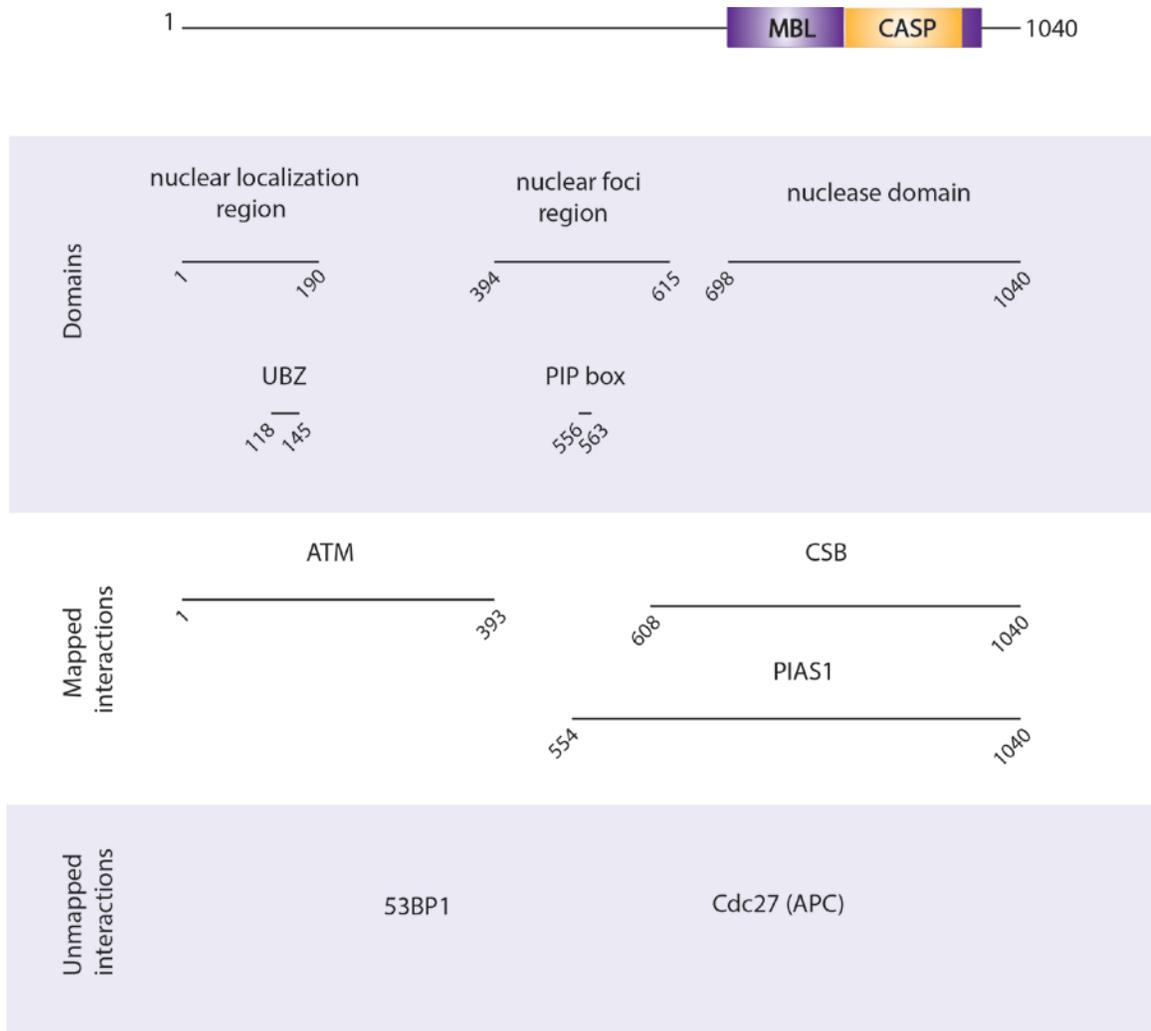
Figure 8: SNM1A domain organization and protein interactions

1.9.3   Structure of SNM1A

While the structure of the MBL/β-CASP domain of SNM1A has been determined, the molecular mechanism underlying catalytic activity is unclear. Firstly, the MBL domain usually coordinates two zinc ions (Dominski, 2007). Indeed, residues within the active site of SNM1A suggest coordination of a second metal, however, within the crystal structure of SNM1A, only a single zinc was observed (Allerston et al., 2015). Interestingly, low concentrations of EDTA stimulated nuclease activity but was inhibitory at higher concentrations, suggesting that another metal may loosely occupy this site, perhaps blocking the correct metal necessary for catalysis (Sengerová et al., 2012). It is unclear if the bound zinc is structurally important or if it participates in catalysis. As expected, mutations within the HxHxDH signature (like SNM1A[D736A]) result in either severely diminished or abolished nuclease activity (Hejna et al., 2007). Mutations in conserved residues of the β-CASP domain in SNM1A have been shown to lead to protein mislocalization (Ishiai et al., 2004).

A key feature observed from the crystal structure of SNM1A is the presence of localized regions of positive and negative charge that are absent in other β-CASP nucleases. The region of SNM1A harbouring a large amount of positive surface charge was hypothesized to encompass a DNA binding groove (Allerston et al., 2015). Extensive mutagenesis revealed that while nuclease activity on small oligo substrates remained essentially the same, digestion of plasmid DNA was impaired or abolished. How this positive patch contributes to the function of SNM1A is not fully understood, but it may be

more important for processivity than catalysis.

1.9.4 SNM1A Nuclease Activities

1.9.4.1 5'-3' Exonuclease Activity

SNM1A has 5'-3' non-processive exonuclease activity that is strictly dependent on the presence of a 5'phosphate (Hejna et al., 2007). While all truncations possessing the MBL/β-CASP domain possess ssDNA exonuclease activity, the extent of this activity, as well as the ability to process dsDNA, seems to differ depending on the length of the N-terminus. Full-length SNM1A purified from yeast and insect cells exhibits weak exonuclease activity on a 15nt substrate and essentially no activity when shortened to 10nt (Hazrati et al., 2008b; Hejna et al., 2007). In contrast, N-terminally truncations SNM1A$^{N\Delta608}$ and SNM1A$^{N\Delta676}$ expressed in bacteria has increased exonuclease activity, able to digest dsDNA to about 5nt and ssDNA to 1nt (Allerston et al., 2015; Sengerová et al., 2012). This truncation has been reported to have a k_{cat} of 2 min^{-1} and 14 min^{-1}, with a K_M of almost 100nM and less than 5nM, for dsDNA and ssDNA oligos, respectively (Sengerová et al., 2012). Surprisingly, these SNM1A truncations are capable of rapidly digesting kilobases of plasmid DNA suggesting that this truncation is highly processive on large duplex substrate despite work demonstrating non-processive exonuclease activity of SNM1A on small substrates (Allerston et al., 2015; Sengerová et al., 2012). If extensive resection of DNA is important for ICL repair is unclear, but it is possible that such activity is used during apoptotic DNA fragmentation.

1.9.4.2 Endonuclease and Translesional Nuclease Activity

Full-length SNM1A has been tested for the presence of endonuclease activity on DNA loops and flaps, however, no endonuclease activity was observed (Hazrati et al., 2008b; Hejna et al., 2007). This finding may not be surprising given that SNM1C must be activated by DNA-PKcs phosphorylation for structure-specific endonuclease activity (Goodarzi et al., 2006). Interestingly, truncation of SNM1C was found to demonstrate constitutive endonuclease activity in the absence of DNA-PKcs (Poinsignon et al., 2004b). SNM1A can complement a *pso2Δ* for ICL repair therefore the two proteins are thought to function similarly in repair (Hazrati et al., 2008b). Surprisingly, the lack of observed endonuclease activity in SNM1A is in direct contrast with the nuclease activity of Pso2 (Tiefenbach and Junop, 2012b).

Although all reported SNM1A truncations lack structure-specific endonuclease and have comparatively weaker exonuclease activity on dsDNA, SNM1A strikingly possesses nuclease activity able to traverse an SJG-136 ICL (Sengerová et al., 2011; Wang et al., 2011a). It is unknown if SNM1A possesses lesion bypass nuclease activity with other types of ICL adducts. The fact that all truncations are able to digest DNA crosslinked with SJG-136 suggests translesion nuclease activity is an important function of SNM1A.

1.10 Thesis Hypothesis, Questions, Objectives

It is evident that SNM1A has a role in ICL repair, however the specific function of SNM1A is unclear. Indeed, many details of ICL repair are still emerging, but given the

critical role of XPF-ERCC1 in ICL unhooking, as well as the epistatic relationship between XPF-ERCC1 and SNM1A, SNM1A may play a pivotal role in ICL repair. Our current understanding of how SNM1A functions may be skewed by existing reports that do not reflect the entire capacity of SNM1A activity.

Given the relationship between SNM1A and Pso2, as well as the activation requirement for endonuclease activity of SNM1C, it is not unreasonable to hypothesize that SNM1A may also possess endonuclease activity. Observing SNM1A endonuclease activity would be significant because it could provide a bridge in the gap of understanding of what nuclease carries out the second incision event during ICL unhooking.

It has been disputed which endonucleases are required for ICL unhooking. While other endonucleases have been implicated in human ICL repair, none have demonstrated the severe defect of SNM1A depletion, comparable to XPF-ERCC1 in response to ICL exposure. Even though SNM1A and ERCC1 have been shown to be epistatic in human cells, the details of the events carried out at the ICL are unknown. Furthermore, models of unhooking in which SNM1A acts as a translesion exonuclease after incision by XPF-ERCC1 do not agree with the presently understood biochemical functions of SNM1A and XPF-ERCC1.

What is clear is that SNM1A plays a key role in human ICL repair. This makes SNM1A an attractive target for inhibition to prevent the repair of ICLs generated during ICL-based chemotherapy treatment. Unlike XPF-ERCC1, which participates in repair of several types of DNA damage, SNM1A is known only to function specifically in ICL repair. Therefore, inhibition of SNM1A may be more specific to ICL repair. Based on

current understanding, we postulate that inhibition of SNM1A nuclease activity may sensitize cells to ICL-inducing agents to hinder the repair capacity of ICL-tolerant cells.

The first aim of this thesis is to comprehensively investigate the nuclease activities of SNM1A, specifically directed at answering if SNM1A possesses structure-specific endonuclease activity. Functional analysis will also be directed at characterization of the translesion nuclease activity of SNM1A. These studies are required to establish the abilities of SNM1A before the role of SNM1A in ICL can be defined.

The second aim is to find small molecules that inhibit SNM1A nuclease activity/activities. Detailed studies on SNM1A inhibition will help provide some insight into how these compounds might affect SNM1A activity both outside and inside the cell. Using SNM1A as a target to sensitize cells to ICL-inducing damage, the goal within this thesis is to reduce the capacity of the cell to repair crosslink damage. In doing so, this work may be first steps in laying basic groundwork addressing the clinically significant problem of chemoresistance.

.

CHAPTER TWO: MATERIALS AND METHODS

1.1 SNM1A Construct Cloning

SNM1A constructs were generate using polymerase chain reaction (PCR) and template from either cDNA or *Escherichia coli* (*E. coli*) codon-optimized template DNA. Amplified segments of SNM1A open reading frame were cloned into pDONR201 using the Gateway cloning system (Invitrogen). For protein expression in *E. coli*, SNM1A was sub-cloned into pDEST-544-ccdB (gift from Dominic Esposito, Addgene plasmid #11519) containing a N-terminal His_6 NusA fusion and TEV protease cleavage site. For protein complementation in yeast, SNM1A was sub-cloned into yeast expression vector pAG423GAL-ccdB, a gift from Susan Lindquist (Addgene plasmid # 14149). *E. coli* strain TOP10 (Invitrogen) was used for plasmid amplification. Plasmids were transformed into chemically competent bacteria via heat shock. All plasmids and constructs generated in this study were verified by DNA sequencing (MOBIX Lab).

1.2 SNM1A and Pso2 Protein Expression and Purification

SNM1A constructs were expressed in Star pRARE pLysS *E. coli* (Invitrogen), and induced at 0.700 OD_{600} with 1mM IPTG at 25°C overnight. Cells were resuspended in Nickel A buffer (50mM Tris pH 7.5, 500mM NaCl, 30mM imidazole, 0.5mM TCEP, 0.01% TritonX100, 10% glycerol) with protease inhibitors (3uM aprotinin, 1uM pepstatin A, 1mM benzamidine, 1uM leupeptin, 1mM PMSF) and lysed with three passes though a cell disruptor at 10,000 psi. The lysate was clarified by centrifugation at 48,000xg for 40

minutes and filtered. The sample was loaded onto a HisTrap HP nickel-chelating column (GE Healthcare) and step eluted with Nickel A buffer containing 210mM imidazole. The sample was diluted to 300mM NaCl using QA Buffer (50mM Tris pH 8.5, 0.5mM TCEP, 10% glycerol) and loaded onto a Q-sepharose HP column (GE Healthcare). Protein was then eluted with 400mM NaCl. TEV protease was added at 5:1 SNM1A to TEV to cleave the His_6-NusA fusion protein overnight. The cleaved sample was diluted into SA buffer (50mM Tris pH 7.5, 0.5mM TCEP, 10% glycerol) prior to loading onto an SP-HP sepharose column (GE Healthcare). Protein was eluted with a linear gradient from 300mM to 1M NaCl. SNM1A containing fractions were pooled and concentrated with a 10kDa MWCO centricon (Corning) Samples were flash frozen in liquid nitrogen and stored at -80°C.

Pso2 was expressed in BL21 pRARE pLysS *E. coli*, and induced at 0.500 OD_{600} with 1mM IPTG overnight at 16°C. Cells were resuspended in NiA buffer (50mM Tris pH 7, 500mM NaCl), 3mM β-mercaptoethanol, 1% TritonX100, and protease inhibitors (3uM aprotinin, 1uM pepstatin A, 1mM benzamidine, 1uM leupeptin, 1mM PMSF), and then lysed with four passes through a French press at 10,000 psi. The lysate was clarified by centrifugation at 48,000xg for 40 minutes and filtered. The sample was loaded onto a HisTrap HP nickel-chelating column (GE Healthcare), washed with NiA buffer and eluted at NiA buffer containing 210mM imidazole. The sample was then diluted to 100mM NaCl using QA buffer (40mM Tris pH 7, 5mM DTT) and loaded onto a Q-sepharose HP column (GE Healthcare). Pso2 was eluted with a linear gradient from 100mM to 500mM NaCl. Pso2-containing fractions were pooled and concentrated with a 30kDa MWCO

centricon (Corning). Samples were flash frozen in liquid nitrogen and stored at -80°C in 10% (v/v) glycerol.

1.3 Endonuclease and Translesion Nuclease Assays

1.3.1 Preparation of Structure-Specific Oligonucleotides

Oligonucleotides containing a 6-FAM or Cy3 fluorescent label (BioBasic) were first purified using 20% denaturing PAGE. Purified DNA oligonucleotides were annealed in 10mM magnesium chloride, 10mM Tris pH 7.5, and 100mM NaCl. DNA substrate requiring intramolecular interactions were annealed at 1µM and heated to 90°C for 10 minutes, then flash cooled on ice for 2-3 minutes. Substrates requiring intermolecular interactions were annealed at 10mM using 1.5 molar equivalents of non-labeled complementary oligonucleotides and heated to 90°C for 10 minutes before slowly cooling to room temperature. Annealed DNA was futher purified using 20% native PAGE. Final substrates were resuspended to a final concentration of 1µM.

1.3.2 Preparation of Site-Specific ICL Oligonucleotides

1.3.2.1 Substrate Crosslinking with SJG-136

Annealed DNA (10µM) containing a single "GATC" site was incubated in 25mM triethanolamine, 1mM Na_2EDTA, pH 7.2 with a 40-fold excess of methanol-dissolved SJG-136 at 37°C overnight. Crosslinked DNA was ethanol precipitated and resuspended into formamide buffer (98% formamide, 10mM EDTA) for denaturing PAGE purification. Crosslinked DNA was resolved with 20% denaturing PAGE, electrophoresed

at 850V. DNA was detected at 526nm using the GelDoc-EZ (Bio-Rad) and bands were excised to separate crosslinked and uncrosslinked DNA. Crosslinked DNA was eluted at room temperature, ethanol precipitated overnight and then resuspended to 1µM.

1.3.2.2 Substrate Crosslinking with Cisplatin

Cisplatin was first activated by reacting 0.9 molar equivalents of silver nitrate with cisplatin and incubated for 37°C for 15 minutes in the dark. The reaction was centrifuged and the supernatant containing activated cisplatin was used to create a monoadduct of the labelled top DNA strand. The labeled top strand containing a single "GC" was reacted with 3 molar equivalents of activated cisplatin in 10mM sodium perchlorate at 37°C for 15 minutes. The complementary strand was then annealed at room temperature with 400mM NaCl for 24 hours. The samples were buffer exchanged into 100mM sodium perchlorate and crosslinked at 37°C for 48 hours. After ethanol precipitating, samples were resuspended into formamide buffer and crosslinked DNA was separated from uncrosslinked DNA using 20% denaturing PAGE. DNA was detected at 526nm using a Typhoon Imager (Amersham) and bands were excised to separate crosslinked and uncrosslinked DNA. Crosslinked DNA was eluted at 37°C, ethanol precipitated, and resuspended to 1µM.

1.3.2.3 Substrate Crosslinking with Psoralen

Annealed DNA containing a single "AT" site was reacted in psoralen crosslinking reaction buffer (50mM NaCl, 5mM Tris pH 7.5, 0.2mM EDTA) with equimolar concentrations of 8-methoxypsoralen (8-MOP) dissolved in methanol for 1 hour at 360nm

on ice. After ethanol precipitation and sample resuspension in formamide buffer, crosslinked DNA was separated from uncrosslinked DNA using 20% denaturing PAGE in the dark. DNA was detected at 526nm using a Typhoon Imager (Amersham) and bands were excised to separate crosslinked and uncrosslinked DNA. Crosslinked DNA was eluted at 37°C overnight, ethanol precipitated, and resuspended to 1µM.

1.3.3   Exonuclease and Endonuclease Assays

SNM1A nuclease assays were performed in 50mM Tris acetate pH 7.2, 75mM potassium acetate, 10mM magnesium chloride, 1mM DTT, 100ug/uL BSA with approximately 0.1µM of fluorescently-labeled DNA. Pso2 reactions were performed with 50mM sodium chloride, 10mM Tris pH 7.9, 10mM magnesium chloride, and 1mM DTT. Reactions were initiated by the addition of protein and incubated at 37°C for various amounts of time. Reactions were stopped with formamide buffer and products were separated with 20% denaturing PAGE at 850V for 26.8 cm gels, or 300V for 6.7cm gels, in 0.5X TBE (Tris/Borate/EDTA) buffer. DNA was detected at 526nm using a Typhoon Imager (GE Healthcare) or ChemiDoc-XRS (Bio-Rad) and quantification was performed using ImageJ or ImageLab (Bio-Rad), respectively.

1.3.4   Translesion Nuclease Assays

Reactions were performed as indicated above. SJG-136 substrates were resolved as crosslinked DNA using denaturing PAGE electrophoresed at room temperature to ensure crosslinked DNA remained intact. For analysis of digestion of SJG-136 substrates, products were heat denatured by boiling for 10 minutes in formamide buffer.

Uncrosslinked DNA was resolved as indicated above. Cisplatin and psoralen substrates were also resolved as above except gels were protected from light to avoid uncrosslinking DNA, particularly in the case of light-sensitive psoralen crosslinked substrates.

1.4 SNM1A Inhibition High-Throughput Screening (HTS) Campaign

1.4.1 HTS Assay Buffer for SNM1A Reactions

All nuclease reactions were performed at room temperature in buffer containing 50mM N-[Tris(hydroxymethyl)methyl]-3-aminopropanesulfonic acid (TAPS) pH 9.1, 10mM magnesium acetate, 75mM potassium acetate, 1mM DTT, and 100ug/ml of BSA. Unless indicated, the substrate used was [5'PO$_4$] AGC [dT-6FAM] A [dT BHQ1] GGTTCGATCAAG [3'OH]. Reactions conditions (including SNM1A and DNA concentrations) are as described below.

1.4.2 K$_M$ Determination for SNM1A HTS Substrate

Reactions containing 25nM SNM1A and DNA ranging from 12.5-3200nM were prepared in black, flat-bottom 384-well plates (Corning) on ice. Reactions were initiated with the addition of DNA. Fluorescence was measured every minute for 2 hours at 26°C at 526nm using the BioTek Synergy 4 Hybrid Microplate Reader. Initial velocity for each curve was calculated and plotted against corresponding substrate concentrations. Product formation, expressed as % fluorescence using the equation below, was used to determine kinetic parameters using GraphPad Prism (v6).

% fluorescence = fluorescence$_p$ – fluorescence$_n$,

where p represents fluorescence SNM1A products and n represents fluorescence of the negative control.

1.4.3   SNM1A Activity Unit Standardization for HTS Assay

Reactions containing 5nM SNM1A and 43.8nM substrate were prepared in black, flat-bottom 384-well plates on ice. Reactions were initiated with the addition of DNA. Fluorescence was measured every minute for 2 hours at 26°C at 526nm using the BioTek Synergy 4 Hybrid Microplate Reader. Full digestion of substrate in 40 minutes defined 1 unit of SNM1A activity.

1.4.4   Z' Factor Determination for HTS Assay

Z' determination for the HTS assay for SNM1A inhibitors was performed using a Biomek FX workstation (Beckman Coulter) equipped with a BioRAPTR (Beckman Coulter) liquid dispensing system. Buffer (10µL) was added to all wells of a black 384-well plate. DMSO (0.5µL), used to dissolve compounds for the HTS assay, was dispensed into each well. Zinc acetate (0.5µL, final concentration 1mM) or water was added to the wells. Buffer containing 1 unit of SNM1A was added (29µL). Reactions (total 40µL) were incubated for 40 minutes at 26°C. DNA (10µL, 42.8nM) was dispensed and fluorescence was immediately measured with the Envision Plate Reader (Perkin Elmer) at 535nm. Final endpoint measurements were taken after 3 hours. Relative fluorescence units (RFU) measurements were used to generate the Z-factor as defined by the following equation:

$$Z' = 1 - \frac{3(\sigma_H + \sigma_L)}{|\mu_H - \mu_L|}$$

where σ represents the mean, μ represents the standard deviation, H represents the high activity control and L represents the zinc-containing low control.

1.4.5 Compound Preparation for HTS Campaign

The McMaster Bioactives Set, compiled by the Center for Microbial Chemical Biology (sourced from Prestwick, BIOMOL, Sigma, and MicroSource), was used for high throughput screening. Compound preparation was performed using the Biomek FX workstation (Beckman Coulter) equipped with a BioRAPTR (Beckman Coulter) liquid dispensing system. 96-well plates containing compounds dissolved in DMSO were dispensed in duplicate into adjacent wells of a black 384-well microplates containing buffer only. For pilot and primary screening, 0.5µL of 1mM compound in DMSO (1% DMSO, 10µM final) was dispensed into 10µL buffer using long pin tools. For concentration-response screening, compounds ranging from 2.4 to 2500nM were dispensed as described above.

1.4.6 Primary Fluorescence-Based HTS Campaign

The primary HTS screen for SNM1A inhibitors was performed using a Biomek FX workstation (Beckman Coulter) equipped with a BioRAPTR (Beckman Coulter) liquid dispensing system. Buffer (10µL) was added to all wells of a black 384-well plate. Compounds were dispensed as described above. Zinc acetate (0.5µL, final concentration 1mM) or water was added to 36 wells to standardize plate to plate variation. Buffer

(29μL) containing 1 unit of SNM1A cooled to 4 °C (Peltier) then was added. Reactions (total 40μL) were incubated for 40 minutes at 26 °C to replicate incubation time of compounds with SNM1A. Addition of DNA (10μL, 42.8nM) was dispensed and fluorescence was immediately measured with the Envision Plate Reader (Perkin Elmer) at 535nm. Midpoint and endpoint of the reactions were measured after 40 minutes and 3 hours, respectively.

1.4.7 Control-Based Normalization of HTS Data

The percent activity of SNM1A in response to each compound was calculated from the measured RFU using the equation:

$$\% \text{ Activity} = \left(\frac{S-L}{H-L}\right) \times 100$$

where S represents the measured sample value and H and L represent the mean of the high-activity and low-activity controls of each plate, respectively.

1.4.8 Interquartile Method Normalization of HTS Data

Quartiles refer to the division of a rank-ordered data set into four equal parts. The interquartile mean is the mean of the middle two quartiles or the middle 50% of rank-ordered data. Sample data was normalized on a per plate basis by dividing every well of the plate by the interquartile mean of the plate according to equation:

$$\text{IQM Normalization activity} = \left(\frac{S}{\mu_{iq}}\right)$$

where S represents the measured sample value and μ_{iq} represents the mean (μ) of the interquartile (iq) data of the plate.

The value of μ_{iq} was found for each plate by sorting the data in ascending order and finding the mean of the interquartile (middle 50%) data. Sorting the data facilitated the creation of rank-ordered plots. The trimmed mean of the middle 50% of the data (as a quality control check for each plate) was compared with the plate median, and any plates where the two values differed by more than 15% were flagged as potentially problematic.

1.5 Gel-Based Inhibitor Analysis

1.5.1 Gel-Based Secondary Screen

SNM1A reactions containing of 3nM SNM1A and 1% DMSO-dissolved compound (6.25 and 25µM final) were prepared in 1.7mL tubes and incubated at room temperature for 40 minutes. Reactions were initiated by addition of 5'phosphate 3'labelled ssDNA (43nM), incubated at 37°C for 3 hours. Reactions were stopped with the addition of formamide loading buffer (95% formamide, 10mM EDTA). Products were separated using 20% denaturing PAGE and detected at 526nm using the Typhoon Imager (GE Healthcare).

1.5.2 Gel-Based Nuclease Specificity Assay

Nuclease specificity assays were performed with 0.5-1µL protein (RecJ, Lambda Exonuclease, and Pso2 exonucleases; Mung Bean endonuclease, T7 Endo, and S1 nuclease endonucleases) in their respective buffers. Reactions were initiated by addition

of 5'phosphate 3'labelled ssDNA (43nM), incubated at 37°C for 3 hours, and stopped with the addition of formamide loading buffer (95% formamide, 10mM EDTA). Products were separated using 20% denaturing PAGE and detected at 526nm using the Typhoon Imager (GE Healthcare).

1.5.3 Gel-Based Inhibitor Characterization

1.5.4 Reaction Conditions

All reactions were performed at 37°C in buffer containing 50mM Tris acetate pH 7.2, 10mM magnesium acetate, 75mM potassium acetate, 1mM DTT, and 100ug/ml of BSA. Unless indicated, exonuclease activity was measured using the 5P1F substrate and endonuclease activity was measured using a gap substrate. All gels were resolved with 23% denaturing PAGE and imaged with the ChemiDoc XRS (Bio-Rad) at 526nm for 2 seconds.

1.5.5 Concentration Response Assay

Compounds demonstrating inhibition of nuclease activity with secondary gel-based screening were further tested for the full range of concentration response using the gel-based screen. Each compound ranging from 3µM to 5mM was diluted 2-fold in 100% DMSO. Reactions containing SNM1A (0.2nM) and 0.5µL of DMSO dissolved compound (5% DMSO final) were incubated for 20 minutes at room temperature. Reactions were initiated by addition of 10 picomoles of DNA and incubated at 37°C for 1 hour. Reactions were stopped with the addition of formamide loading buffer and DNA analyzed as described above.

1.5.6 Time Course Assay

To determine the concentration required for full substrate digestion at 60 minutes, 2μM of SNM1A was diluted from 20- to 1200-fold and a time course assay from 2 to 64 minutes was performed. Master mix containing diluted SNM1A was aliquoted and initiated with addition of 100nM DNA. Product formation, as a percentage of total substrate was calculated based on:

$$\text{Reaction Progression} = \frac{\% \text{ Product in reaction} - \% \text{ Product in control}}{100}$$

where a time point reflecting reaction progression of 20% was used for K_M determination.

1.5.7 K_M Determination

Reactions containing SNM1A (0.2nM for exonuclease activity and 200nM for endonuclease activity) were prepared in a 1.7mL volume. Reactions were prewarmed for 2 minutes to 37°C. DNA, ranging from 20-1000nM, was added to initiate reactions. Exonuclease reactions were incubated for 3 minutes and endonuclease reactions for 15 minutes. K_M reaction velocities were determined using:

$$\text{Velocity} \left(\frac{[\text{DNA}]}{\text{sec}} \right) = \frac{\text{Reaction Progression} * [\text{DNA}]}{\text{Time Incubated (sec)}}$$

Triplicate reaction velocities were curve-fitted using Michaelis-Menten kinetics on GraphPad PRISM 6.0.

1.5.8 IC_{50} Determination

Reactions containing SNM1A (0.2nM for exonuclease activity and 200nM for

endonuclease activity) and inhibitor in DMSO (30nM-50µM) were prepared in 1.7mL tubes and incubated for 20 minutes at room temperature. DNA (109nM exonuclease substrate or 28nM endonuclease substrate) was added to initiate reactions. Exonuclease and endonuclease reactions proceeded at 37°C for 150 and 60 minutes, respectively.

1.5.9 Mode of Action Assays

DNA dose-dependent reactions were incubated with inhibitors at their respective IC_{50} values. Velocities were determined using concentrations of DNA ranging from 20nM to 4µM. Reactions containing Reactive Blue 2 and ATA were incubated at 37°C for 40 minutes, while those containing theaflavin digallate and 4,4'-diisothiocyanostilbene-2,2'-sufonic acid sodium salt (DIDS) proceeded for 15 minutes. Reaction times differed since increased maximum velocities of competitive inhibitors required longer incubation times.

1.6 Cell-Based Assays

1.6.1 AlamarBlue® Cell Viability Assay

Cells in 100ul of DMEM were seeded and grown in 96-well black tissue-culture treated plates. To determine cell viability, 11µL of alamarBlue® (Invitrogen) was added directly to the media and incubated at 37°C. Plates were removed from the light for three hours before fluorescence was measured at 590nm.

1.6.2 IC$_{10/50}$ Determination

HeLa cells were seeded at 3500 cells/well in 96-well tissue culture treated black plates in 100 µL of DMEM media. After 24 hours, media was removed and fresh media containing cisplatin, MMC, or gemcitabine (solubilized in water) were added to the cells. Plates were incubated for 72 hours at 37°C, after which cell viability was measured using alamarBlue® as described above.

1.6.3 Compound Testing

HeLa cells were seeded at 3500 cells/well in 96-well tissue culture-treated black plates in 100 µL of DMEM media. After 24 hours, media was removed and fresh media containing the compound(s) of interest was added. Compounds from the McMaster Bioactives set were solubilized in 100% DMSO, and assayed at a concentration of 25µM, with a final concentration of 1% DMSO. Cisplatin, MMC, and gemcitabine were solubilized in water to the appropriate concentration. Untreated controls were included on all plates as a reference. Plates were incubated for 72 hours before measuring cell viability by alamarBlue®.

1.6.4 Yeast Transformations

Yeast was transformed with SNM1A expression vectors, by culturing a single isolate of *S. cerevisiae* in yeast peptone dextrose (YPD) media overnight at 30° C. Cultured yeast (1mL) was pelleted at 1000 x g before addition of ~1ug plasmid DNA, 20ug single-stranded herring DNA, and transformation buffer (40% polyethylene glycol 4000, 100mM lithium acetate, 10mM Tris pH 7.5, 1mM EDTA pH 8). The mixture was

incubated at room temperature overnight. Pelleted cells were plated on selective dropout base minimal media agar with 2% glucose. Transformants appeared after incubation at 30°C for three days.

1.6.5 Transposase Assay

Complementation vectors were transformed into MYO37 budding yeast (MAT**a** *ura3Δ0 his3Δ0 ade2ΔHpa*I-*Pfl*M1 *trp1::hisG pso2::kanMX4)* containing transposition plasmid pWL201, a kind gift from Clifford Weil. Cultures were grown in histidine- and uracil-lacking minimal media with 2% raffinose for 24 hours at 30°C. Transposase and complement expression was induced by addition of 2% galactose then grown for another 24 hours at 30°C. Cells were serially diluted in PBS and plated on histidine/uracil-lacking agar ± adenine. Plates were incubated at 30°C for three days and colonies on adenine-containing agar were enumerated to determine the total number of cells. Colonies on adenine-lacking plates were incubated for three weeks before enumeration to determine reversion frequency. Relative survival was calculated by dividing colony forming units (CFUs) from adenine-lacking plates by the CFUs from adenine-containing plates, and then normalized to PSO2 complementation.

CHAPTER THREE: RESULTS

3.1   Endonuclease Activity of SNM1A

3.1.1   Endonuclease Activity in Yeast

Prior work from our lab had demonstrated that Pso2 opens DNA hairpin structures *in vitro* and *in vivo* (Tiefenbach and Junop, 2012). We reasoned that since SNM1A complementation had been shown to rescue ICL hypersensitivity defects in *pso2Δ* yeast, SNM1A complementation might reveal similar hairpin-opening activity exhibited by Pso2 in yeast, despite the apparent absence of SNM1A endonuclease activity *in vitro* (Hazrati et al., 2008a; Hejna et al., 2007; Sengerová et al., 2012). To test this possibility, we used a yeast assay that had been previously used to demonstrate Pso2 hairpin-opening activity (Tiefenbach and Junop, 2012a; Yu et al., 2004). Plasmids, shown in **Figure 9**A were transformed into *pso2Δade2Δ* auxotrophic yeast. The transposition plasmid harbours an inducible transposase gene and an ADE2 gene disrupted by a transposable element. Transposition results in hairpin-capped plasmid ends, requiring resolution by the endonuclease activity of Pso2. Only complementation by a competent hairpin-opening endonuclease can permit restoration of the ADE2 gene.

Using this yeast-based assay, full-length and N-terminal truncations of SNM1A were examined for endonuclease activity. We tested full-length SNM1A and SNM1A$^{N\Delta608}$, which had been reported to lack endonuclease activity *in vitro*, as well as SNM1A$^{N\Delta698}$, consisting of only the MBL/β-CASP domain (Hejna et al., 2007; Sengerová et al., 2012).

When normalized to PSO2 reversion frequency, reversion of all SNM1A constructs were found to be greater than complementation with the empty vector, thus suggesting SNM1A is capable of endonuclease activity in yeast (**Figure 9**B). We noted that complementation with SNM1A resulted in more revertants than yeast complemented with PSO2. However, the apparent decreased reversion frequency of PSO2 complementation may reflect Pso2 regulation in the cell and not differences in endonuclease activity between SNM1A and Pso2.

We mutated the aspartic acid of the conserved catalytic HxHxDH motif (D736A) to determine whether the increased reversion from SNM1A complementation was a result of SNM1A nuclease activity. Indeed, the catalytically inactive mutants decreased reversion frequency by about 50% as compared to corresponding wild-type constructs (**Figure 9**C). Although the difference between mutant and wild-type SNM1A was not statistically different, this is likely due to reduced, but not abolished, activity of the mutant. Furthermore, the GAL4 promoter allows for high protein expression and may produce more protein than the number of hairpin intermediates of the transposable element on the plasmid. Nonetheless, since SNM1A rescued the *pso2Δ* defects in transposon hairpin-capped end resolution, it would appear that SNM1A possesses DNA hairpin-opening endonuclease activity similar to Pso2.
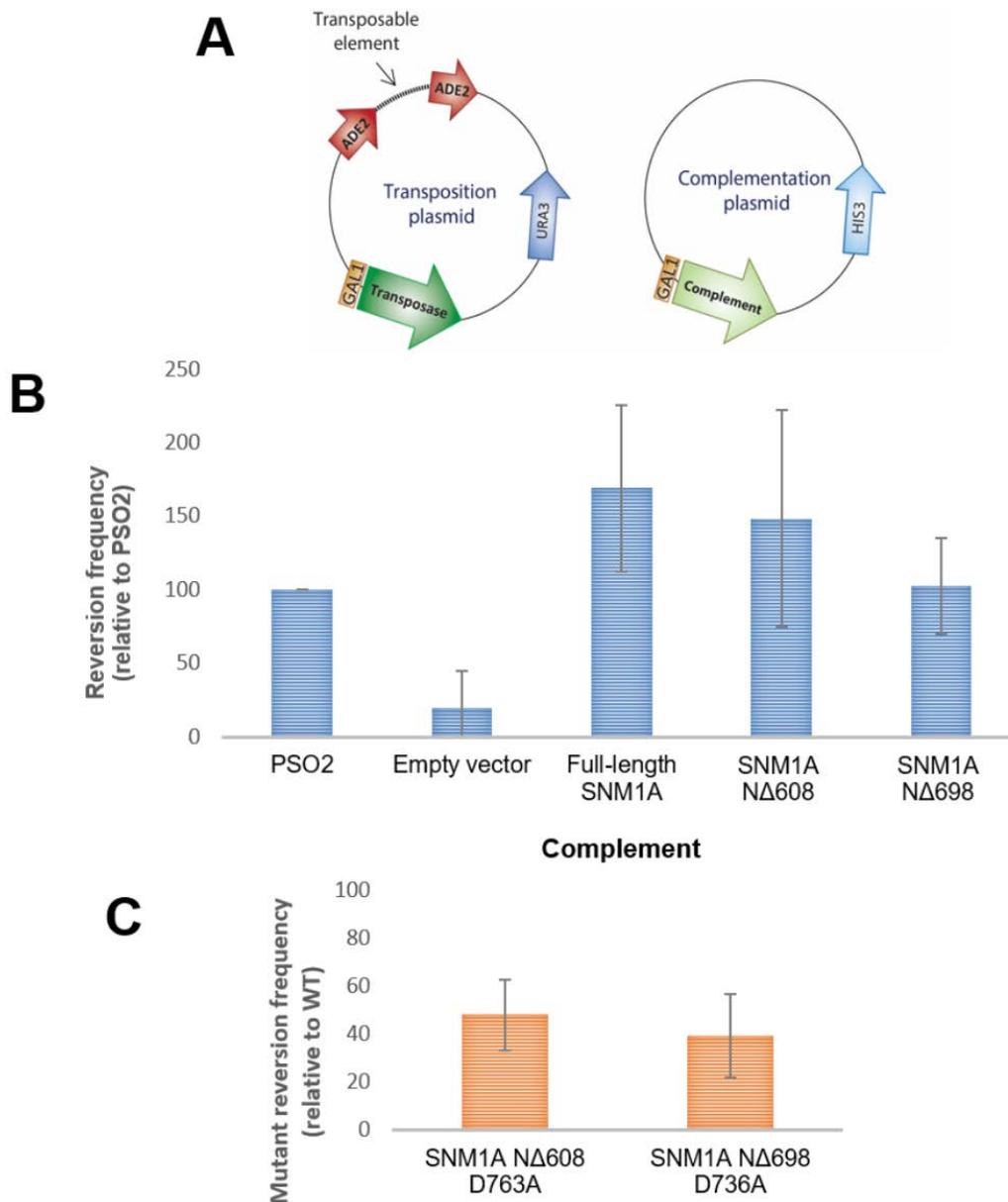
Figure 9: Analysis of SNM1A hairpin-opening activity in yeast.

(**A**) Schematic representation of plasmids used in the hairpin-opening transposase assay in *pso2ΔadeΔ* yeast. (**B**) Reversion frequency of SNM1A constructs indicative of hairpin-opening activity. (**C**) Comparison of reversion frequency of catalytic-dead SNM1A mutants relative to wild-type counterparts.

3.1.2   Hairpin-Opening Activity of SNM1A *In Vitro*

Although SNM1A hairpin-opening activity was observed in yeast, it remained unclear why SNM1A endonuclease activity had not observed previously *in vitro*. It is possible that SNM1A behaves like SNM1C, which is autoinhibited by its non-conserved C-terminus (Goodarzi et al., 2006; Pawelczak and Turchi, 2010b; Poinsignon et al., 2004a). Co-incubation with DNA-PKcs kinase or truncation to the MBL/β-CASP catalytic domain is required for SNM1C endonuclease activity *in vitro*. Thus, to study possible endonuclease activity of SNM1A, we expressed and purified only the MBL/β-CASP domain of SNM1A (residues 698-1040), which is analogous to the SNM1C truncation utilized for SNM1C biochemical characterization. As expected, purified SNM1A$^{N\Delta698}$ (herein referred to as solely SNM1A) had robust exonuclease on single strand DNA (ssDNA) containing a 5' phosphate (5'P), but not on the 5'hydroxyl (5'OH) substrate, indicating that truncation to the catalytic core of SNM1A was sufficient for exonuclease activity (**Figure 10**A, left and middle). Since testing endonuclease activity using a substrate lacking a 5'P can not be degraded by exonuclease activity, we probed a DNA hairpin structure with a terminal 5'OH, where the 5'OH would ensure that any product observed was due to endonuclease activity of SNM1A. As seen in **Figure 10**A, SNM1A exonuclease-"independent" products were observed, suggesting possible SNM1A endonuclease activity. Products of the hairpin-opening assay were subsequently separated using long denaturing PAGE for greater product resolution. Mung Bean
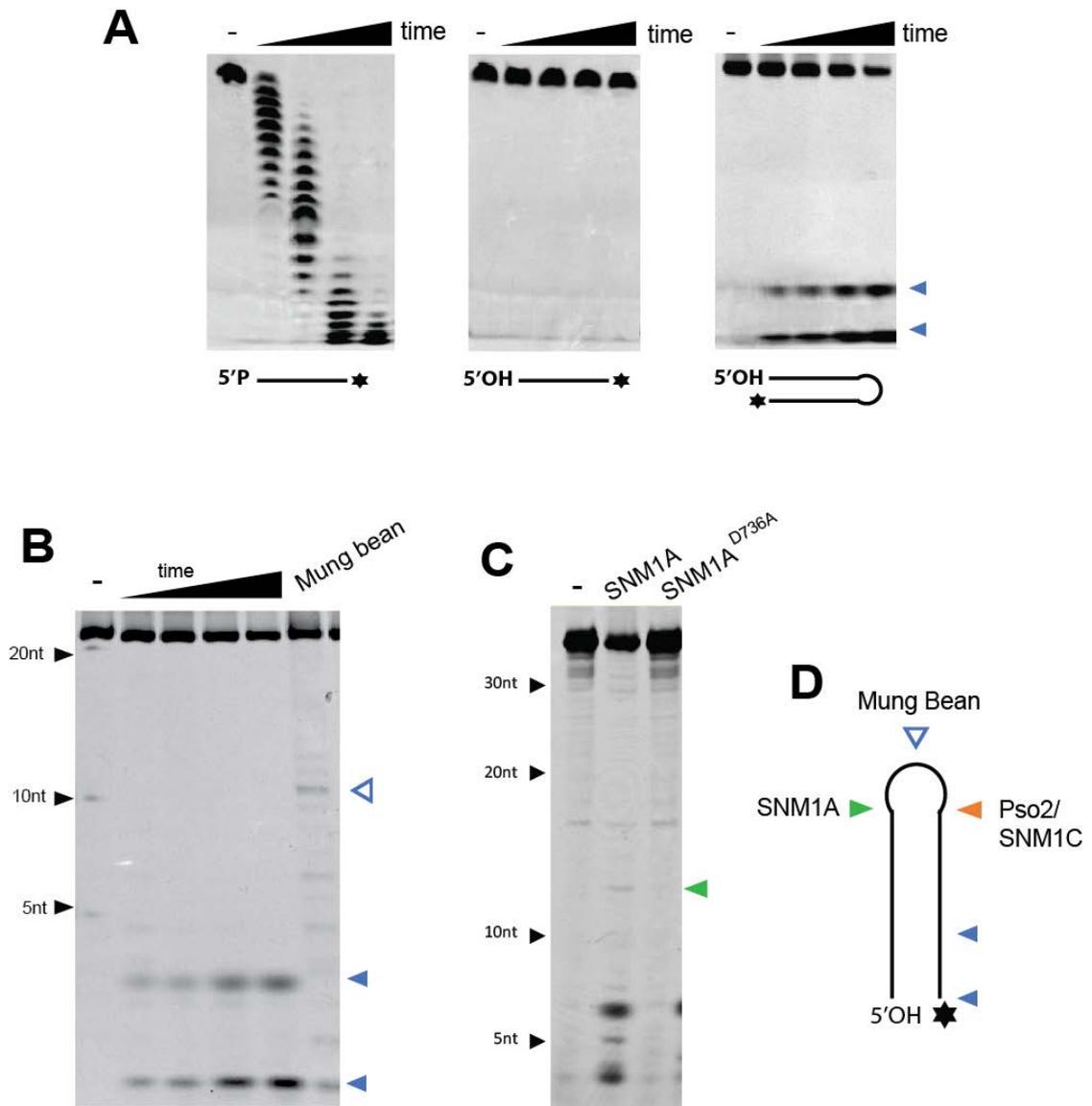
Figure 10: Hairpin-opening nuclease activity of SNM1A *in vitro.*

(**A**) Time course assay of SNM1A nuclease activity on single-strand and hairpin DNA containing a 3' label. (**B**) Comparison of SNM1A endonuclease products to Mung Bean nuclease. (**C**). Nuclease assay of SNM1A activity on 5'F substrate (D) Summary of endonuclease events on hairpin substrate.

cleavage at the apex of the hairpin (**Figure 10**B, open blue triangle) confirmed the hairpin structure. Unexpectedly, SNM1A products (**Figure 10**B, solid blue triangles) were much shorter than anticipated. It was possible that smaller products were generated from exonuclease digestion of contaminating nicked substrate not visible on the gel. Alternatively, SNM1A endonuclease activity within a 5'OH substrate would liberate a 5'P for exonuclease digestion. To eliminate this possibility, a substrate capable of being degraded by only endonuclease activity was required.

A 5'label would accomplish this task, however, ^{32}P labelling was not feasible due to the substrate being subject to 5'phosphate-dependent exonuclease processing. A 5'fluorophore (5'F) label was therefore used. Not only did the fluorophore block SNM1A exonuclease activity, but it also captured initial endonuclease cleavage since the products were not subject to downstream exonuclease activity. Using a 5'F labeled hairpin substrate, digestion by wild-type SNM1A revealed a band (**Figure 10**C, green triangle) that was absent in the negative control and catalytically inactive SNM1A reaction. Interestingly, the endonuclease cut site was two nucleotides 5' of the hairpin apex, contrary to hairpin cleavage by Pso2 and SNM1C (**Figure 10**D, orange triangle), known to occur two nucleotides 3' of the apex. Nonetheless, hairpin-opening activity of SNM1A was clearly observed using a 5'F labeled substrate, suggesting that SNM1A has endonuclease activity.

3.1.3   Structure-Specific Endonuclease Activity of SNM1A *In Vitro*

Since SNM1A has hairpin-opening endonuclease activity, we wondered if SNM1A acts on other DNA structures. SNM1C possesses structure-specific endonuclease activity, therefore we tested similar substrates upon which SNM1C was shown to act (Ma et al., 2005). DNA structures including overhangs, flaps, gaps, and loops were tested for endonuclease activity to determine structural requirements for cleavage. As seen in (**Figure 11**A), SNM1A cut the overhang, pseudo-Y and flap substrate on either the 5' or 3' side, suggesting that SNM1A can access DNA for cleavage, regardless of polarity. Activity observed on closed DNA loops and bubbles further suggests that a free ssDNA end is not required for SNM1A incision. The gap substrate generated the most distinct cleavage pattern, where two main products occurred due to cleavage closer to the 3' end of the unpaired region.

Since, wild-type SNM1A, but not SNM1A^{D736A} cleaved DNA (**Figure 11**A), this implied that observed endonuclease activity was not due to a contaminating nuclease. Additionally, Pso2, which was tested in parallel for comparative analysis, showed similar cleavage patterns. The fact that endonuclease products of SNM1A and Pso2 were similar and that the SNM1A catalytic mutant failed to generate these products supports the claim that observed endonuclease activity was a result of SNM1A and not a contaminating nuclease.
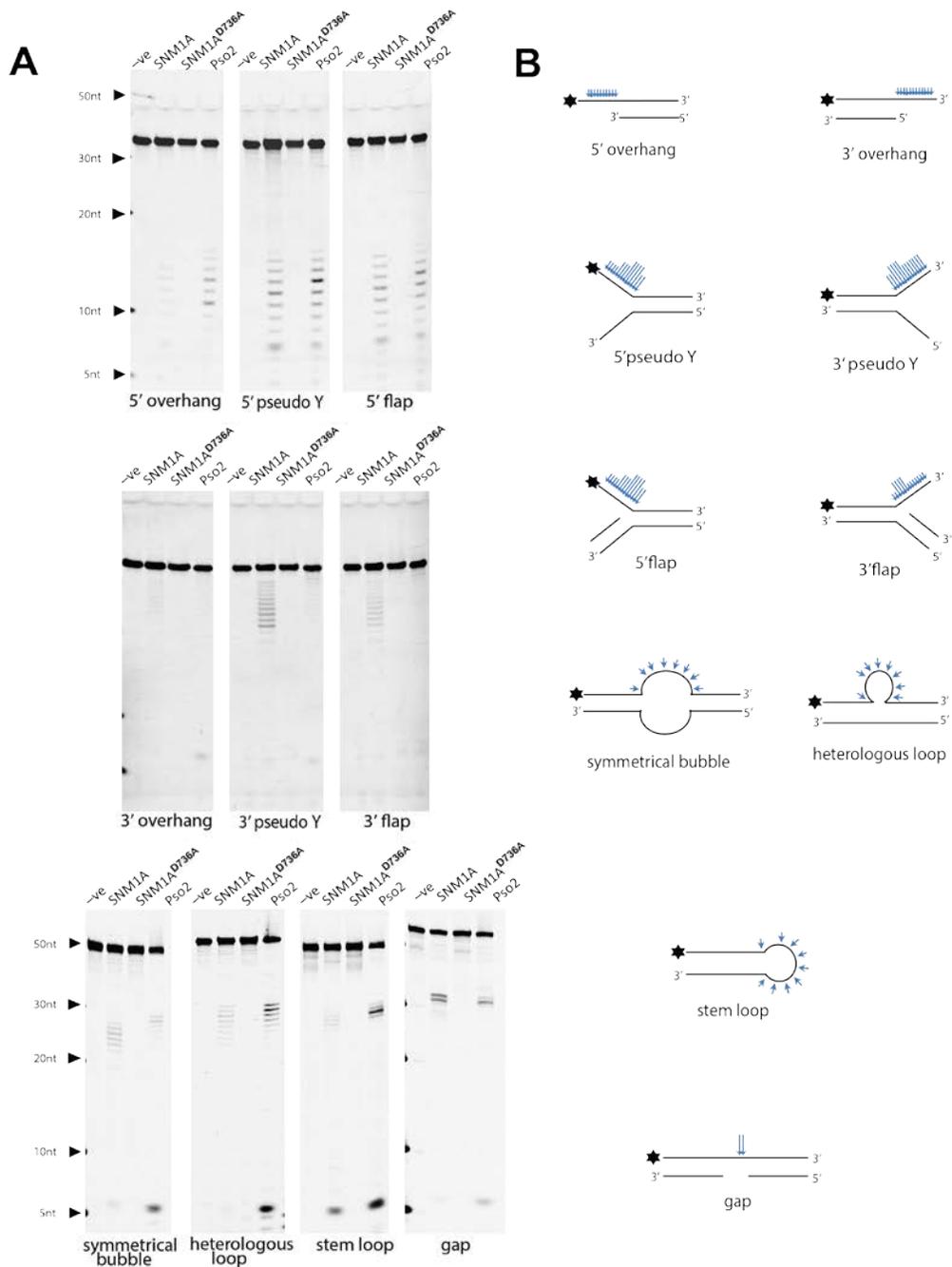
Figure 11: Analysis of SNM1A structure-specific endonuclease activity.

(A) SNM1A endonuclease activity assays *in vitro*. **(B)** Schematic summary of SNM1A endonuclease cleavage events. Long, medium, and short arrows indicate the relative efficiency of cleavage at each position.

3.1.4   Single-Strand Specific Nuclease Activity of SNM1A *In Vitro*

From the DNA structures tested in **Figure 11**, cleavage was notably observed in regions of unpaired DNA. We therefore asked if ssDNA was sufficient for SNM1A endonuclease activity. Three 5' fluorophore-labelled ssDNA substrates composed of only single base nucleotides (polyT, polyA, polyC) were tested for SNM1A endonuclease activity. Substrates were also probed with Lambda exonuclease, a 5'-3' exonuclease which preferentially acts on DNA with a 5' terminal phosphate, to exclude the possibility that products were the result of random substrate processing by a 5'-3' exonuclease.

SNM1A nicked the DNA between nearly all the nucleotides of the polyT substrate, as evident by the intensifying laddering pattern persisting over time (**Figure 12**A). While this laddering pattern could suggest 3'-5' exonuclease activity, these products did not shorten towards the 5' fluorophore over time, indicating this nuclease activity could not be the result of 3'-5' exonuclease contaminating nuclease. Three nucleotides 3' to the 5'F were noted to be absent, possibly due to steric hindrance by the fluorophore. Interestingly, we found that SNM1A showed moderate endonuclease activity on the polyC substrate and essentially no activity on the polyA substrate, as compared in **Figure 12**B. We therefore wondered if the endonuclease activity of SNM1A might be base-specific. We created a polyAC substrate containing alternating adenine and cytosine triplets. Products of SNM1A from the polyAC substrate generated an alternating pattern with three bands present then three bands absent, suggesting preferential activity on pyrimidine-containing nucleotides. It is also possible that single-stranded pyrimidines adopt a more open conformation that preferentially exposes the phosphate backbone for

cleavage. Indeed, the additional imidazole ring of the purine base can result in stronger base stacking. Furthermore, thymine has two polar keto groups that may increase exposure of the phosphodiester bond for endonuclease attack. Therefore, fewer interactions amongst the bases could increase the ability for SNM1A engagement and cleavage, explaining the increased activity of SNM1A on the polyT substrate.
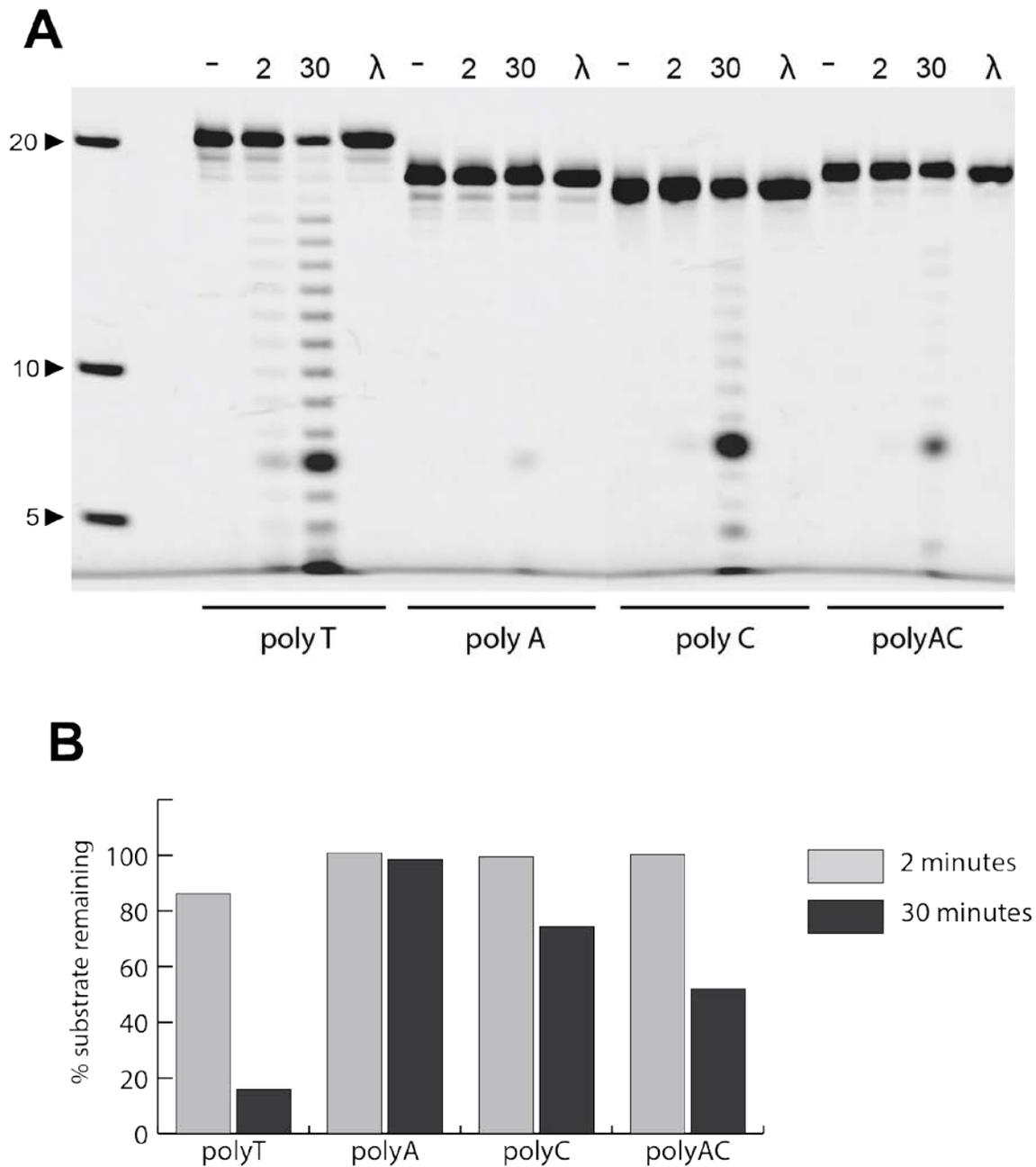
Figure 12: Analysis of SNM1A single-stranded endonuclease activity.

(A) Time course assay of SNM1A nuclease activity on ssDNA. **(B)** Quantification of substrate remaining in (A).

3.1.5   SNM1A Endonuclease Activity Examining DNA Gap Structure

Analysis of SNM1A activity on a gap substrate (**Figure 11**A) surprisingly resulted in only two products instead of the ten possible products available within the unpaired region of the substrate. Moreover, the nucleotides cleaved within purine nucleotides, counter to observations in the prior experiment. The basis of this distinctive endonuclease activity by SNM1A was further explored by creating a second gap substrate. The original gap substrate which contained ten-random unpaired nucleotides (10N) spanning the gap was changed to a gap containing ten thymidine residues (10T). When changed to 10T, SNM1A was able to cut between each nucleotide within of the gap (**Figure 13**A and **Figure 13**B). Secondary structure prediction of the 10N sequence identified a possible hairpin structure. Three DNA structures were further generated to investigate the possible structure of the 10N gap and the structure specificity of SNM1A: a gap spanning three nucleotides (3T), a hairpin containing an apex of four unpaired nucleotides (4T), and an extrusion exchanging the unpaired gap for a hairpin structure. Interestingly, SNM1A only acted on the extrusion, cutting between a GC sequence on either side of the base of the hairpin. These findings suggest that the structure adopted by the original gap substrate may form a cruciform-like substrate, bending the DNA for preferential cleavage (**Figure 13**C). While further work would be required to fully elucidate all the structural requirements for SNM1A cleavage *in vitro*, the occurrence of distinct DNA product was well suited for endonuclease activity validation and product quantification for characterization.
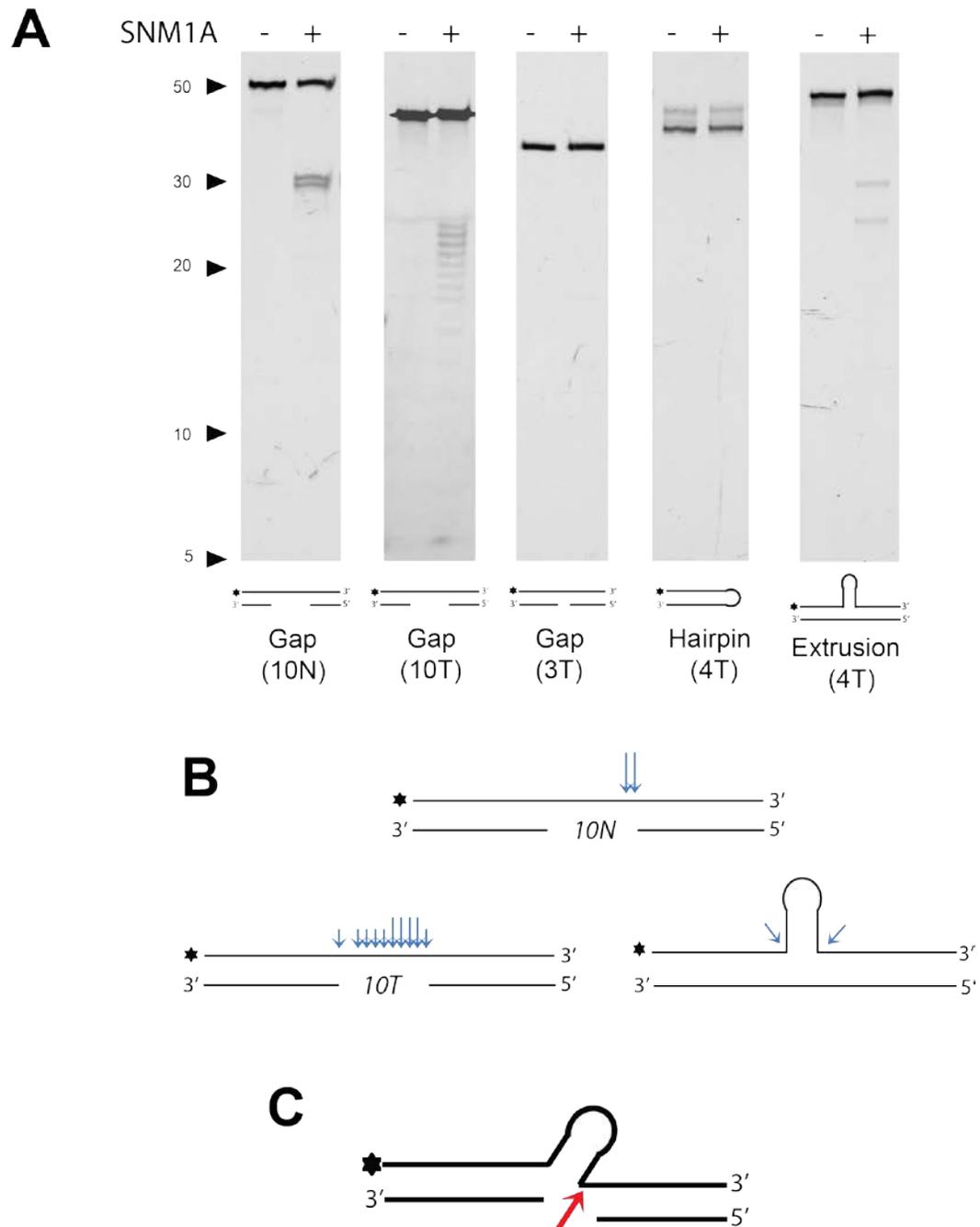
Figure 13: Analysis SNM1A endonuclease activity on gap substrate derivatives

(A) Nuclease assay of SNM1A on DNA derivatives of the of gap substrate. **(B)** Schematic of SNM1A cleavage (in blue arrows) on substrates. **(C)** Possible structure and cleavage (red arrow) of the 10N gap substrate.

3.1.6 Co-Purification of SNM1A Exonuclease and Endonuclease Activity

Since SNM1A had previously been reported to lack endonuclease activity, validation of the observed SNM1A endonuclease activity reported above was necessary. Unlike SNM1A, a single point mutation within the active site of SNM1C did not appear to abolish exonuclease activity (Pannicke et al., 2004). Furthermore, depending on the method of purification, the two nuclease activities did not copurify (Pawelczak and Turchi, 2010). To that end, wild-type SNM1A was purified and fractions collected from ion-exchange chromatography were tested for both 5'phosphate-dependent exonuclease and gap-specific endonuclease activity (**Figure 14**, **Figure 15**, **Figure 16**).

Exonuclease activity was detected in all fractions containing SNM1A, however, endonuclease activity was not observed until after TEV cleavage of the N-terminal NusA solubility fusion (**Figure 15**C). Following the removal of the NusA fusion by TEV digestion, endonuclease and exonuclease activity were observed, suggesting that both activities were from the same protein (**Figure 16**). Western blot analysis with an antibody specific to the C-terminus of SNM1A confirmed the presence of SNM1A in all fractions with exonuclease and endonuclease activities (**Figure 16**B).
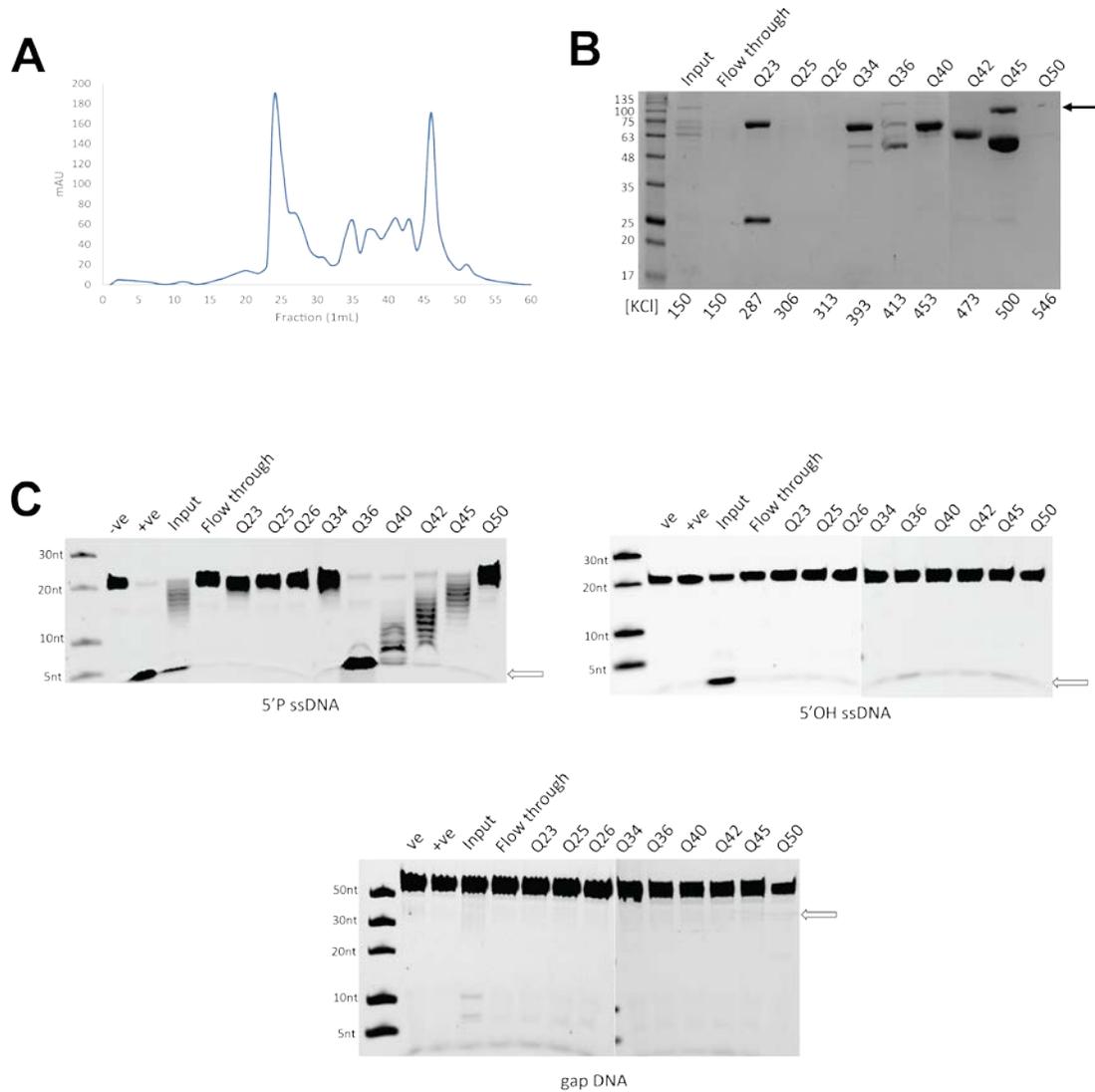
Figure 14: Analysis of nuclease activity of SNM1A from anion exchange purification

(A) Chromatogram of SNM1A purification using Q-sepharose (Q-seph) anion exchange column. (B) SDS-PAGE analysis of Q-seph fractions from UV peaks in (A). (C) Nuclease activity analysis of Q-seph fractions from (B).
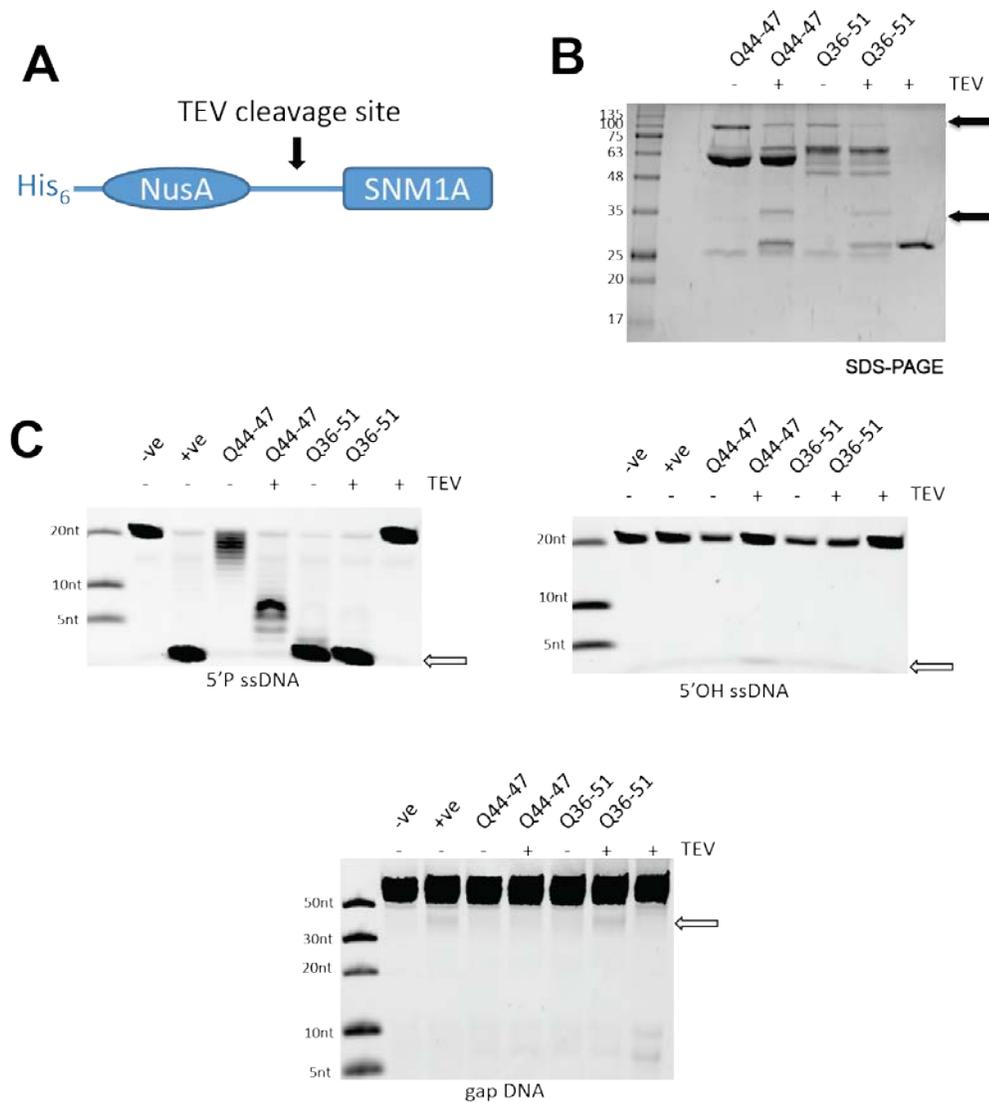
Figure 15: Analysis of nuclease activity of SNM1A after TEV cleavage

(**A**) Schematic of TEV cleavage site of SNM1A construct. (**B**) SDS-PAGE analysis of SNM1A-NusA TEV cleavage from pooled Q-seph fractions from Fig 15. (**C**) Nuclease activity analysis of samples from (C).
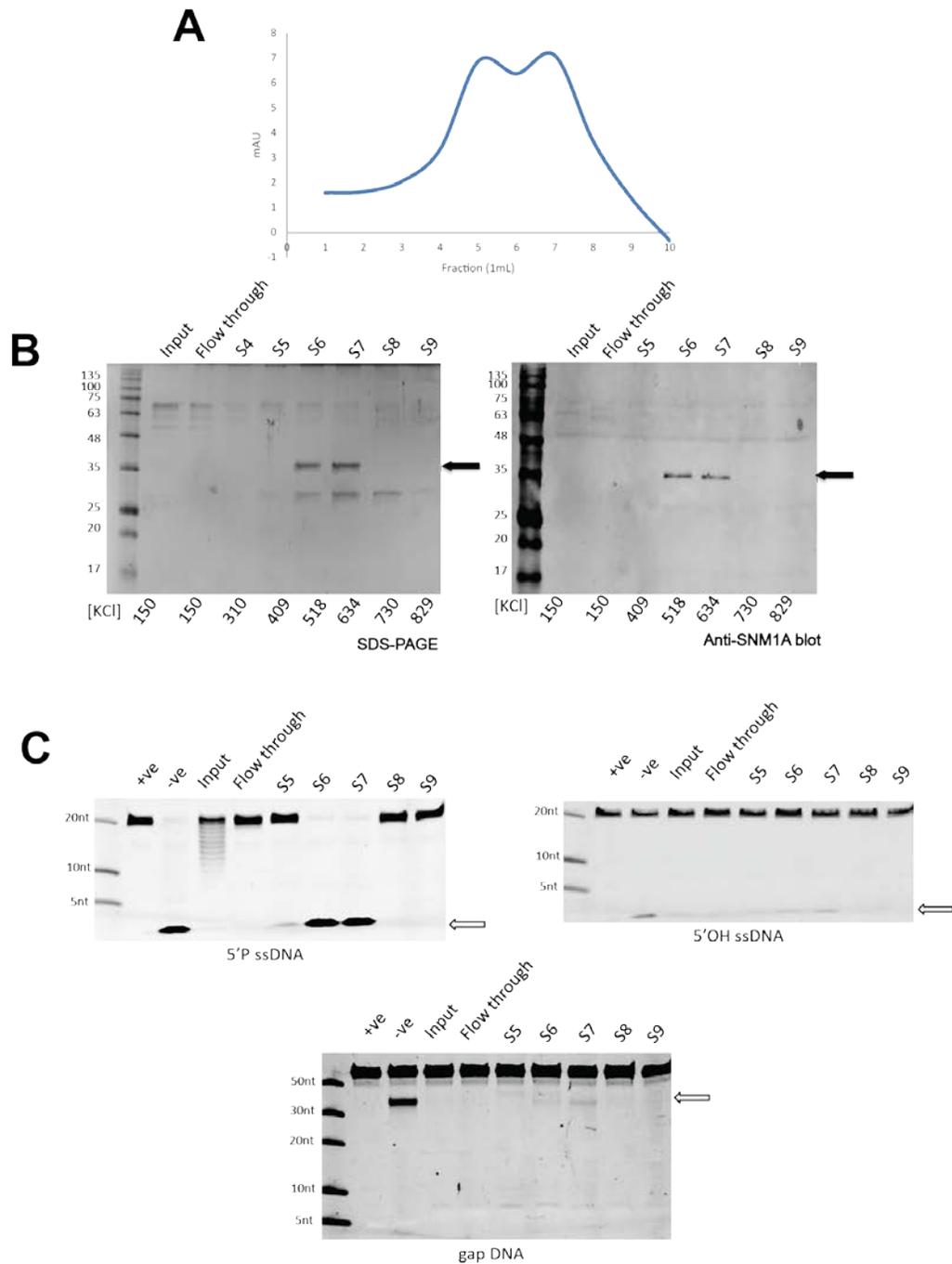
Figure 16: Analysis of nuclease activity from SNM1A cation exchange purification

 (A) Chromatogram of SNM1A purification using a SP sepharose (S-seph) cation exchange column. **(B)** SDS-PAGE and western blot analysis of S-seph fractions from (A). **(C)** Nuclease activity analysis of S-seph fractions from (B).

Endonuclease activity had not been observed with SNM1A$^{N\Delta608}$ (residues 608-1040), encompassing the same MBL/β-CASP domain, but with an addition 90 amino acids N-terminal to the catalytic core (Sengerová et al., 2012). We wondered if endonuclease activity appeared absent because their 3' labelled substrates were acted on by both endonuclease and exonuclease activity, or if the addition of the 90 N-terminal residues altered the endonuclease activity. Using the same purification protocol used for SNM1A$^{N\Delta698}$, SNM1A$^{N\Delta608}$ was expressed and purified. Again, fractions from the final cation exchange column were tested for exonuclease and endonuclease activity (**Figure 17**A and **Figure 17**B). Like SNM1A$^{N\Delta698}$, SNM1A$^{N\Delta608}$ exhibited robust exonuclease activity that co-purified with endonuclease activity when fractions were assayed immediately post-purification. Storage of purified SNM1A$^{N\Delta608}$ were flash frozen for subsequent nuclease assays in the same manner as purified SNM1A$^{N\Delta698}$. However, essentially no endonuclease and limited exonuclease activity were observed in reactions with frozen aliquots of SNM1A$^{N\Delta608}$ (**Figure 17**C). This sensitivity to freeze/thaw may explain why endonuclease activity of SNM1A was not observed by others. Nonetheless, these results strongly suggest that the nuclease domain of SNM1A possesses not only known exonuclease activity, but also previously unreported endonuclease activity.
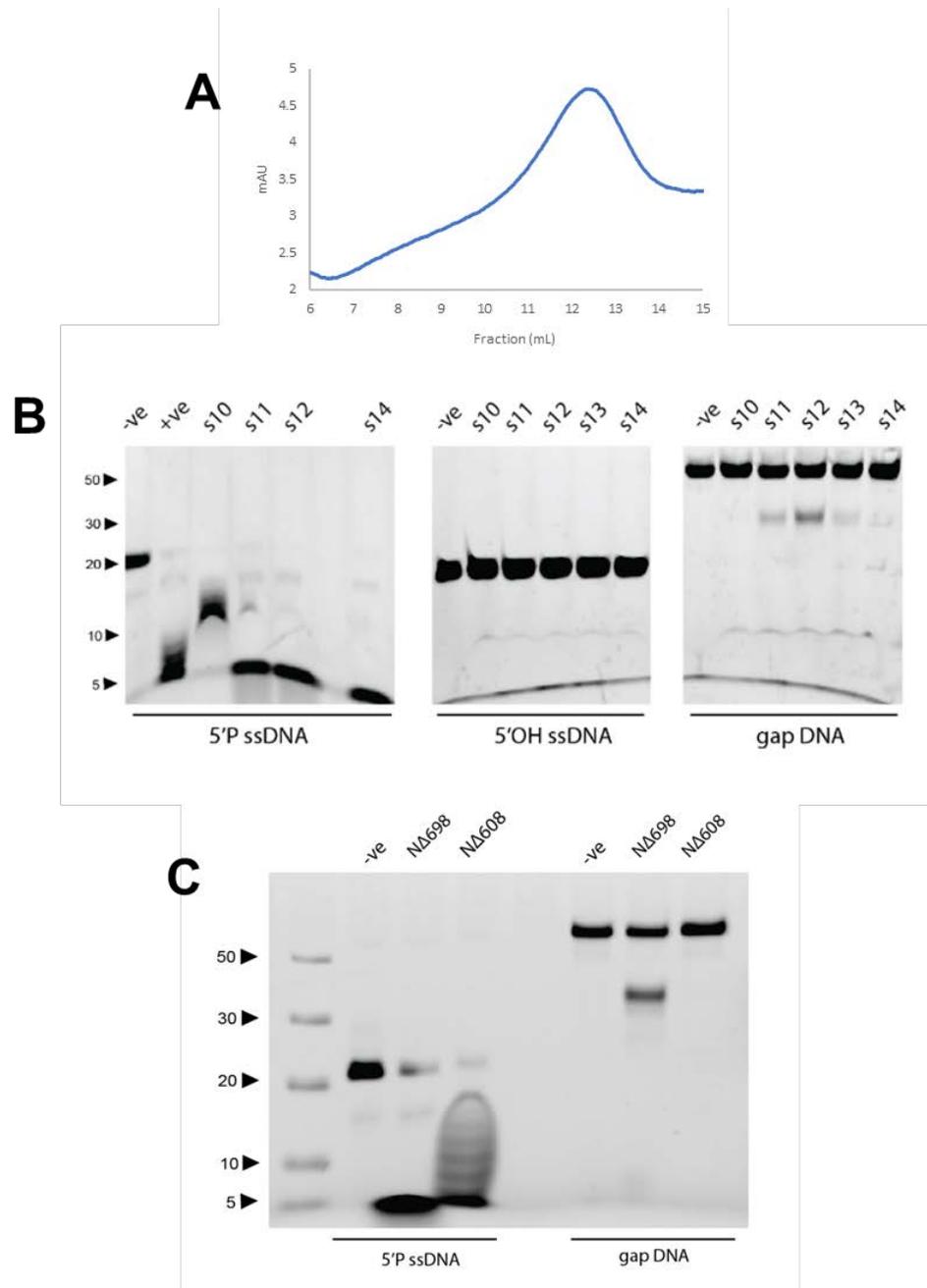
Figure 17: Analysis of SNM1A$^{N\Delta608}$ nuclease activity

(**A**) Chromatogram of SNM1A purification using S-seph column. (**B**) Nuclease activity analysis of S-seph fractions from (A). (**C**) Comparison of SNM1A$^{N\Delta698}$ and SNM1A$^{N\Delta608}$ nuclease activity after flash freezing.

3.1.7 Comparison of SNM1A Nuclease Activity Kinetics

To determine the kinetic parameters of the nuclease activities of SNM1A, quantification of bands in a gel-based assay was carried out. For exonuclease activity, a substrate able to measure a single nuclease event was used to simplify quantification. A ssDNA oligonucleotide containing a 5'P and a fluorescein label conjugated to the first base was used to monitor single turnover events (**Figure 18**A). The product was measured and analyzed using a Michaelis-Menten non-linear regression model, resulting in a K_M of 109nM, k_{cat} of 530ms^{-1}, and catalytic efficiency of 4.7nM^{-1}s^{-1} (**Figure 18**B).

Similarly, the kinetic parameters of SNM1A endonuclease activity were determined using the gap substrate (**Figure 18**C). Although this gap substrate resolves as two products on long denaturing PAGE (**Figure 12**A and **Figure 14**A), the single band observed using short denaturing PAGE (**Figure 18**C) was used for quantification. The endonuclease activity of SNM1A had a K_M of 28nM, a k_{cat} of 0.14 ms^{-1}, and catalytic efficiency of 0.05 nM^{-1}s^{-1} (**Figure 18**D).
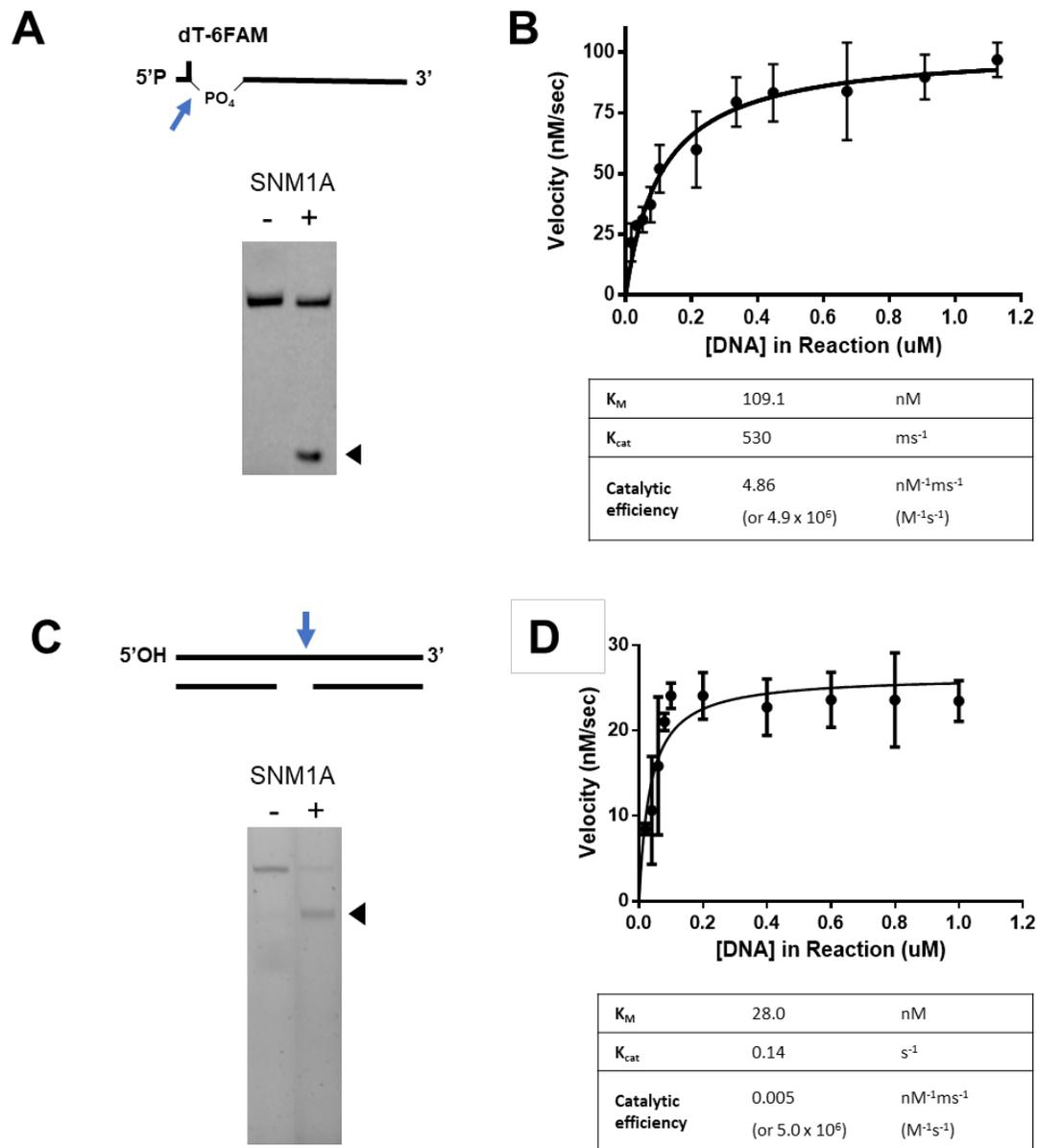
Figure 18: SNM1A K_M and k_{cat} determination

(A) Schematic of exonuclease substrate and reaction product for quantification of SNM1A digestion. **(B)** Michaelis-Menten curves and kinetic parameters of exonuclease activity. **(C)** Schematic of endonuclease substrate and reaction product for quantification of SNM1A cleavage. **(D)** Michaelis-Menten curves and kinetic parameters of endonuclease activity.

Compared to kinetic parameters previously reported for SNM1A$^{N\Delta608}$, the K_M and k_{cat} of SNM1A$^{N\Delta698}$ was 20x greater and 2x less, respectively. Thus, our SNM1A construct had an apparent 10-fold decrease in catalytic efficiency. However, it is likely that this reflects differences in substrates, reaction conditions, and method of quantification. On the other hand, the same conditions and method of quantification allow for comparison of exonuclease and endonuclease activities of the MBL/β-CASP domain of SNM1A. While SNM1A had greater affinity for the gap substrate compared to the ssDNA substrate, the significant difference in k_{cat} resulted in a catalytic efficiency three orders of magnitude greater for the exonuclease activity of SNM1A compared to its endonuclease activity. Given the thousand-fold difference in catalytic efficiency of SNM1A exonuclease to endonuclease activity, it seems reasonable that endonuclease activity of SNM1A was previously unobserved.

3.2   Translesion Nuclease Activity of SNM1A

3.2.1   SNM1A Translesional Nuclease Activity on SJG-136 ICL Substrates

Previous work demonstrated that SNM1A is able to traverse an ICL adduct, specifically an SJG-136 ICL (Sengerová et al., 2012; Wang et al., 2011). SJG-136 is composed of a dimer of two PBD groups that strongly bind into the minor groove of DNA and covalently bond the N7 of guanine via an aminal linkage. The high affinity of the PBD dimer for DNA substantially increases yields of crosslinked oligonucleotide substrates compared to other crosslinking agents (Hartley et al., 2004b). Additionally, heat reversibility of SJG-136 facilitates the analysis of strand cleavage (Wang et al., 2011a). Prior studies have shown that on SJG-126 crosslinked oligonucleotides, SNM1A is able to digest up to the ICL (gray arrows) and then bypass the lesion (as depicted in **Figure 19**A). It was thought that SNM1A stalls after the crosslink (blue arrow), before proceeding to digest DNA (white arrows). It was unclear from these studies why product, presumed to be the result of SNM1A stalling, accumulated three nucleotides (+3) past the covalent linkage. This specific +3 product was interesting, as it is reminiscent of hairpin cleavage by Pso2 and SNM1C (orange arrow). Given the similarity of the hairpin and resected ICL substrate, we wondered if hairpin opening endonuclease activity of SNM1A was the mechanism underlying translesion nuclease activity.

To test the hypothesis that an ICL intermediate might be endonucleolytically cleaved as a hairpin, we created a substrate resembling a resected hairpin-like ICL intermediate as shown in **Figure 19**B. To block exonuclease activity, we replaced the

5'end of a "trimmed" ICL from a 5'P with a 5'OH and incubated the substrate with SNM1A, Pso2 and their respective catalytic mutants (**Figure 19**C). Lambda exonuclease was used as a control since it is a 5'-3' exonuclease with preferential activity on substrates with a 5' phosphate to ensure the presence of the ICL. After reaction completion, products were heat-treated to reverse the SJG-136 crosslinked DNA and analyzed by denaturing PAGE.

Uncrosslinked duplex substrate containing a 5'P (lane 1 of each panel) was digested by wild-type SNM1A and Pso2, as expected. On the 5'P crosslinked substrate with the ICL, both SNM1A and Pso2 appeared to demonstrate translesion nuclease activity, where products were observed smaller than DNA covalently linked by SJG-136 (at position 0) indicative of digestion past the ICL (lane 2 of each panel). The main product of Pso2 digestion past the crosslink was at the +3 position, consistent with reports of nuclease stalling. With SNM1A, products beyond the +3 position were absent under the conditions tested. Strikingly, however, Lambda nuclease also generated products on the 5'P ICL substrate past the crosslink. This was entirely unanticipated as Lambda exonuclease is not known to participate in ICL repair and has not been reported to bypass DNA adducts.
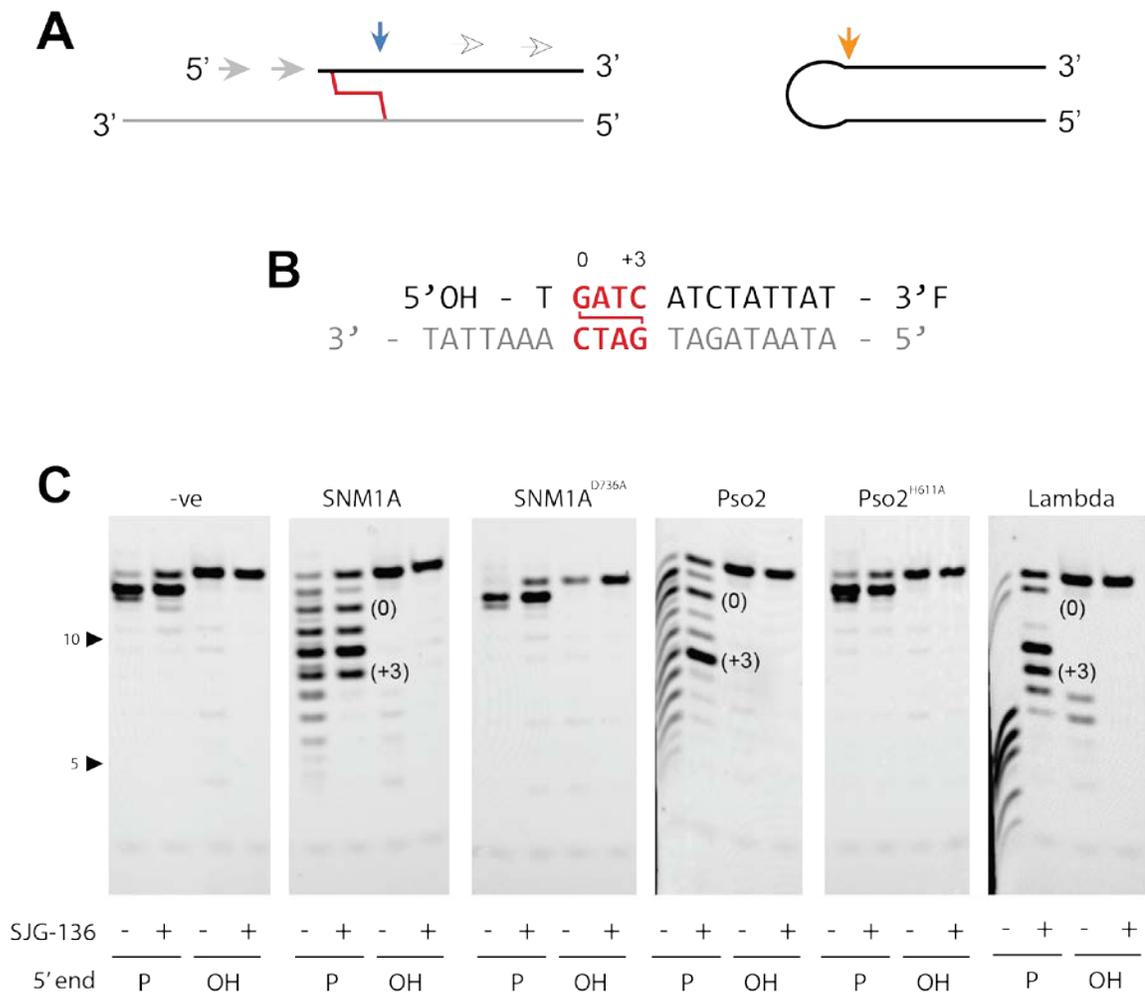
Figure 19: Possible hairpin-like opening activity on SJG-136 crosslinks

(A) Comparison of previously observed SNM1A product (in blue) on SJG-136 crosslinked DNA (left) and cleavage position (in orange) of hairpin DNA by Pso2 (right). **(B)** Sequence of ICL substrate with single SJG-136 sequence (in red). **(C)** Nuclease assays comparing SNM1A and Pso2 activity on substrates with or without an SJG-136 crosslink and 5'P or 5'OH. All DNA was subject to heat treatment to reverse SJG-136 crosslink prior to gel electrophoresis.

It remained possible that translesion nuclease activity of Lambda exonuclease was an artifact of how our ICL substrates were created. In studies where SNM1A showed translesion nuclease activity on a substrate resembling a trimmed SJG-136 ICL intermediate, substrates were created by first crosslinking fully duplexed DNA. The crosslinked substrate was then nicked 5' to the ICL to create a trimmed ICL substrate (Wang et al., 2011a). We instead annealed "pretrimmed", or complementary DNA one base pair 5' to the ICL (top strand in **Figure 19**B) prior to crosslinking the substrate. To eliminate possible differences in substrate preparation, fully duplexed DNA was annealed and crosslinked with SJG-136 (**Figure 20**A). To ensure that the substrate was indeed crosslinked, we analyzed reactions containing the intact ICL (**Figure 20**B, left) and heated reactions containing the heat reversed ICLs (**Figure 20**B, middle), in addition to duplex DNA (**Figure 20**B, right). Analysis of products with the intact ICL reveals both the labelled and unlabelled crosslinked strands together. In addition, this analysis demonstrates the purity of the crosslinked substrate. To minimize crosslink reversal, voltage used for denaturing PAGE was reduced to maintain gels at room temperature. Under these modified conditions, approximately 2% of the crosslinked substrate degraded to uncrosslinked substrate (**Figure 20**B, lane 1). Heat treatment of reactions facilitates analysis of cleavage products of the labelled strand, independent of the crosslink. Finally, the products of SNM1A activity on duplex DNA shows exonuclease activity unimpeded by the ICL.

Using the fully duplexed ICL substrate in **Figure 20**A, most of the products accumulated at the crosslink (position 0) in **Figure 20**B. A band 3+ to the ICL (+3) and

smaller products arising from digestion 3' to the ICL were also observed, consistent with reported activity of SNM1A on this ICL (**Figure 20**B, lanes 6 and 7). We again used Lambda exonuclease to probe the nature of the crosslinked substrate. Digestion by Lambda exonuclease resulted in product accumulation one nucleotide up to the crosslink (-1 position). A second product corresponding to resection of the unlabeled crosslinked strand (**Figure 20**B, lane 4) was also observed, however heat denaturation of reactions with Lambda exonuclease showed that digestion of this unlabeled strand did not interfere with analysis (**Figure 20**B, lane 8). Again, smaller products indicative of exonuclease activity past the ICL were observed (**Figure 20**B, lanes 4 and 8), implying that either the crosslink was not stable or that Lambda exonuclease also possesses translesional nuclease activity.

We noted that DNA 3' of the SJG-136 adduct had a T_m lower than the temperature at which the nuclease assays were performed. A longer ICL, differing only in length of duplex DNA 3' of the ICL, was created in attempts to keep DNA 3' of the ICL annealed during the entire reaction. The short ICL substrate (**Figure 20**A) had duplex DNA 3' of the ICL with a T_m of less than 25°C, whereas the T_m of the long ICL substrate (**Figure 20**C) was greater than 45°C. Like the short ICL substrate, nearly all products of the long ICL accumulated at the crosslink (**Figure 20**D). A very faint +3nt band was also seen just below the ICL as well (**Figure 20**D, lane 7). However, no additional bands 3' of this "stall point" were detected in SNM1A reactions with the long ICL substrate. Lambda exonuclease still generated smaller products indicative of exonuclease activity past the ICL (**Figure 20**C, lane 8).
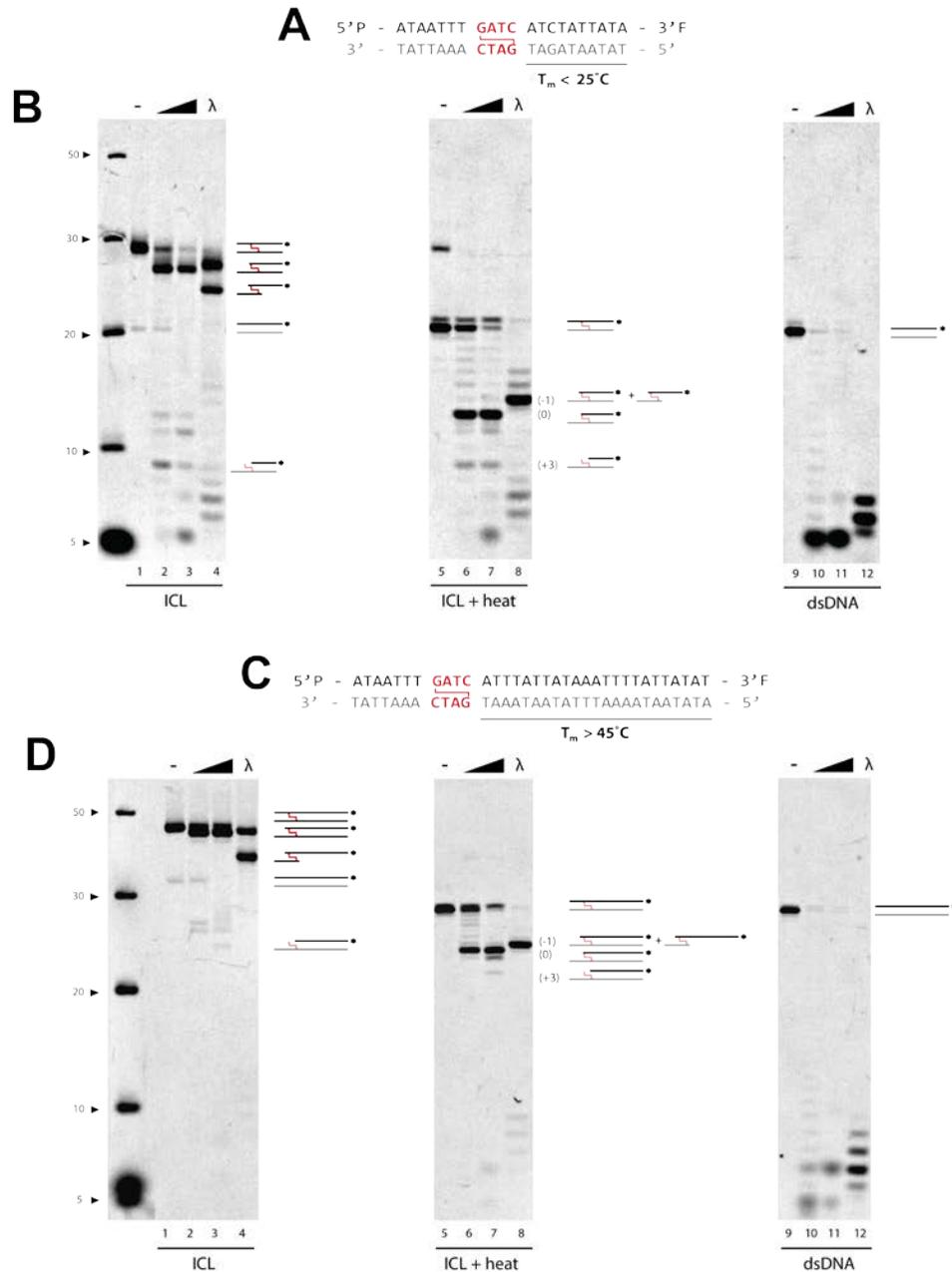
Figure 20: Analysis of translesional nuclease activity on stable duplex DNA.

(A) Short SJG-136 ICL substrate and T_m of DNA 3' to crosslink. **(B)** Nuclease assays of SNM1A and Lambda exonuclease using substrate in (A). **(C)** Long SJG-136 ICL substrate and associated T_m of DNA 3' to crosslink. **(D)** Nuclease assays of SNM1A and Lambda exonuclease using substrate in (C). Uncrosslinked DNA was heat treated to reverse SJG-136 crosslink prior to gel electrophoresis.

To further stabilize duplex DNA downstream of the ICL, we created GC1, a GC-rich substrate (**Figure 21**A, top). As expected, both SNM1A and Lambda exonuclease yielded product at position 0 of the GATC crosslink (**Figure 21**C, left). Strikingly, an additional major band was observed, preceding a pair of cytosines (CC). Furthermore, minor products were observed past the ICL corresponded to guanines 3' to the ICL. With the exception of the band at position 0, all other bands were unexpected, as SJG-136 is thought to be a sequence-specific crosslinking agent (Rahman et al., 2009).

To further investigate the nature of this species, the GATC crosslink sequence was modified (in blue) such that only one strand could be monoadducted for GC2 substrate (**Figure 21**A, bottom). The CC sequence in GC1 was also changed to GG in GC2 (in purple). Although GC2 did not crosslink as efficiently as GC1, a majority of GC2 duplex DNA reacted with SJG-136 and purified as an ICL substrate, even in the absence of the reported GATC crosslinking sequence (**Figure 21**B).

GC2 yielded products of similar intensity at most, but not all, guanines for both SNM1A and Lambda exonuclease (**Figure 21**C, right). Product accumulation at GAAA was expectedly diminished, but still present. These findings suggest that the crosslinking sequence, which contains the reacting guanine on the opposing strand, enhances but is not required by SJG-136 for DNA adducting. Additionally, smaller products corresponding to digestion up to a guanine suggests that SJG-136 also forms monoadducts, impeding both SNM1A and Lambda exonuclease. Surprisingly, changing the CC sequence to GG surprisingly decreased the number of products at these positions, despite the fact that the

covalent linkage occurs through the guanine. Thus, inhibited nuclease progression past the CC sequence suggests that SJG-136 has a significant steric impact on nuclease activity.

Therefore, we suggest that SNM1A and Lambda exonuclease rapidly digest DNA to the crosslink, but are obstructed, as evident by the accumulation of product preceding the lesion. The aminal bond linking SJG-136 to the N7 guanine either undergoes spontaneous hydrolysis or is absent, but progression is sterically hindered by the PBD moiety, particularly in stable duplex DNA. Release of the covalent bond on the opposing strand relieves steric hindrance to allow full bypass of the nuclease, regardless of the exonuclease used. While it is possible that these results may also suggest that Lambda exonuclease possesses translesion nuclease activity, this is unlikely since DNA must enter a central channel before entry into the buried active site as seen in the crystal structure of substrate-bound Lambda exonuclease (PDB ID:3S1P) (Mitsis and Kwagh, 1999). Collectively, these results do not support SNM1A translesion nuclease activity on SJG-136 crosslinks.
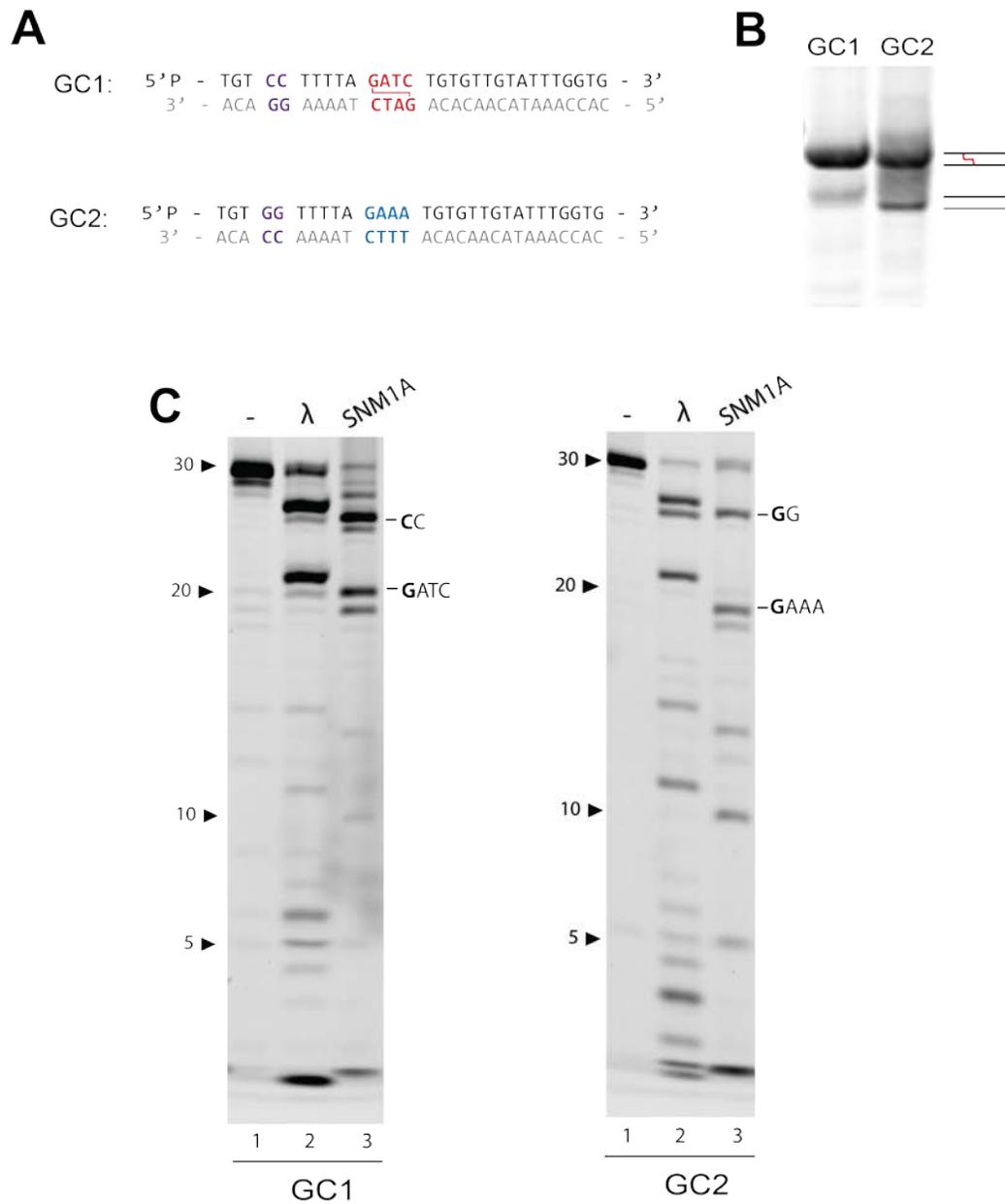
Figure 21: Analysis of nuclease activity on SJG-136 adducts.

(A) GC1 substrate (left) with single SJG-136 crosslinking sequence (in red) and GC2 (right) without crosslinking sequence (in blue). Purple indicates additional difference between GC1 and GC2. (B) SJG-136 crosslinking reaction products using (A). (C) Nuclease assay with substrates in (A). All DNA was subject to heat treatment to reverse SJG-136 crosslink prior to gel electrophoresis.

3.2.2   SNM1A Activity on Psoralen and Cisplatin ICL Substrates

Although translesion nuclease activity by SNM1A was not observed on an SJG-136 ICL, we wondered if SNM1A might have translesion nuclease activity on other ICLs adducts. Unlike SJG-136, ICLs formed by cisplatin and psoralen result in considerable distortion to the DNA backbone. Since the extent of distortion of an ICL can affect recognition and translesion bypass by TLS polymerases, we wondered if similar effects might also affect translesion nuclease activity.

An oligonucleotide substrate was designed with a 5'P and a single AT site at which psoralen forms crosslinks (**Figure 22**A). Methoxy-substituted psoralen (8MOP) was used as the crosslinking agent due to its increased solubility compared to unsubstituted psoralen. As shown in **Figure 22**B, wild-type SNM1A and Pso2 almost fully processed DNA substrates with and without an 8MOP ICL. By contrast, Lambda exonuclease digest uncrosslinked DNA, but was impeded by an 8MOP crosslink, showing no product at the bottom of the gel with the ICL containing substrate. Product accumulation by Lambda exonuclease indicated that the crosslink was indeed stable. Ability of SNM1A and Pso2 to fully degrade substrate in spite of a stable psoralen crosslink suggests both nucleases are competent translesion nucleases on an 8MOP ICL. Further time course analysis demonstrated rapid digestion of the 8MOP ICL by SNM1A, where the intermediate at the crosslink (**Figure 22**C, lane 2) was fully turned over by 10 minutes.
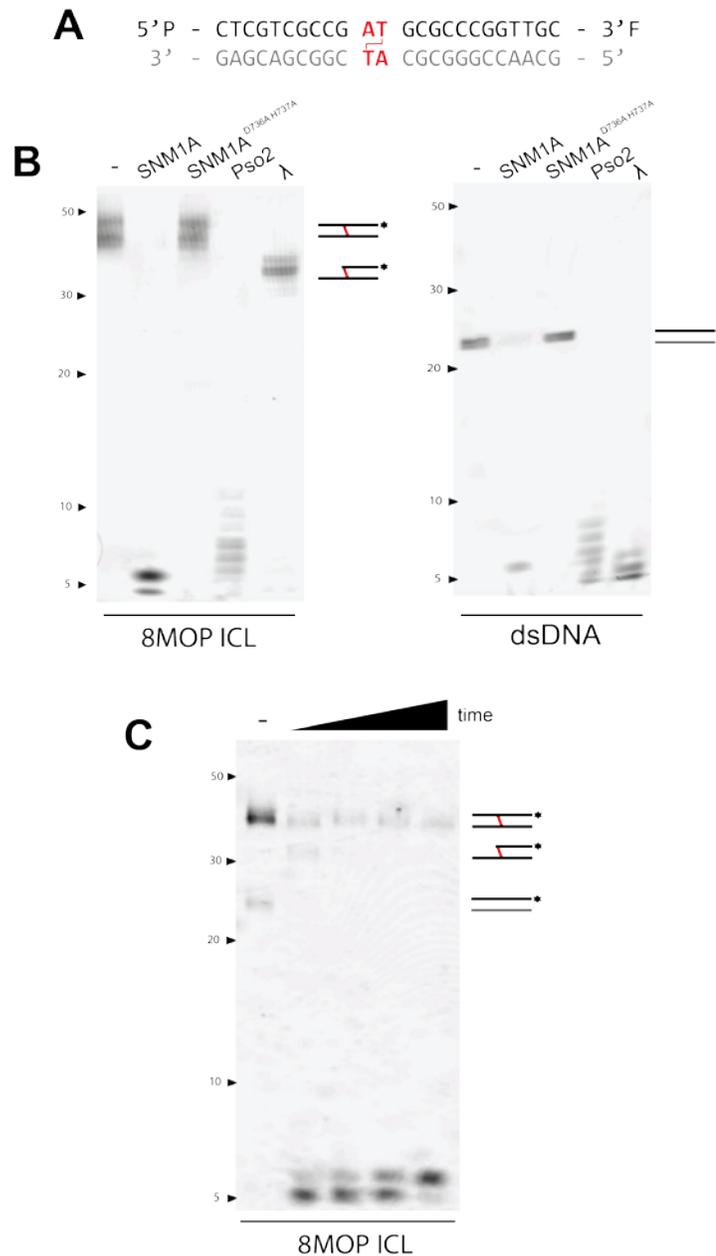
Figure 22: Translesional nuclease activity on psoralen crosslinked DNA

(A) Substrate containing single 8-methoxypsoralen (8MOP) crosslink (in bold). **(B)** Nuclease assay on 8MOP crosslinked (left) and duplex only (right) DNA. **(C)** Time course assay of SNM1A nuclease activity on 8MOP crosslinked substrate.

Likewise, a cisplatin ICL substrate was designed, as shown in **Figure 23**A. Digestion of this substrate by SNM1A resulted in half of the product accumulating at the crosslink within 5 minutes (**Figure 23**B, lane 2), and less than 10% after 60 minutes (**Figure 23**B, lane 5). On the same substrate, nearly all product of Lambda exonuclease persisted at the crosslink (**Figure 23**C, lanes 2-5). Products towards the bottom of the gel can be accounted for as digestion products of the contaminating uncrosslinked DNA (**Figure 23**C, lane 1). Since Lambda exonuclease could not progress past the site of the cisplatin ICL, it confirms the site and stability of the crosslink. These results suggest that SNM1A translesion nuclease activity on cisplatin ICLs is less robust than exonuclease activity on undamaged duplex DNA. Although SNM1A appears to have translesion exonuclease activity on both psoralen and cisplatin adducts (**Figure 23**B), the results are less clear since there was a time dependent loss of total DNA (perhaps due to aggregation of SNM1A over time).
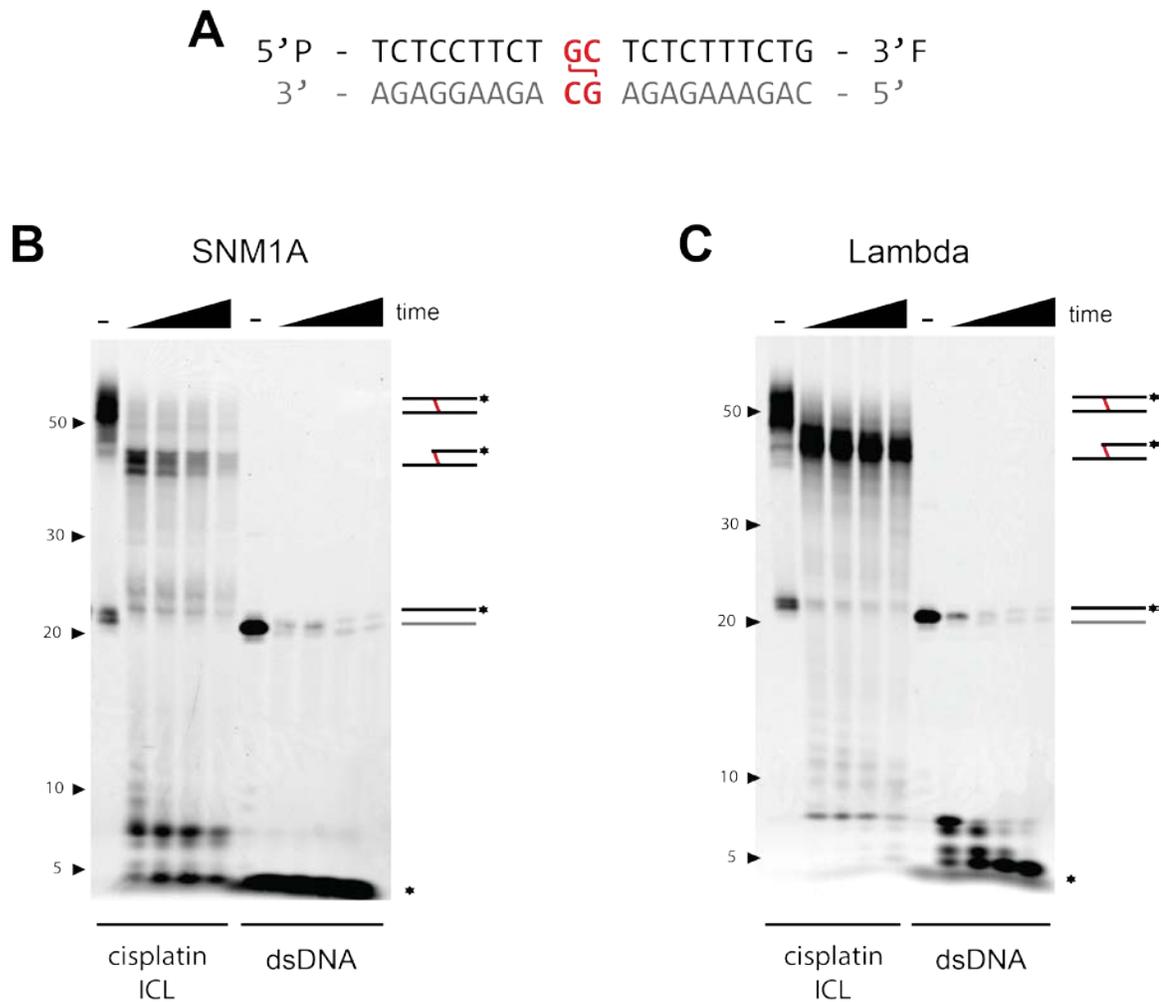
Figure 23: Translesion nuclease activity on cisplatin crosslinked DNA

(A) Substrate containing a single cisplatin crosslink (in bold). **(B)** Nuclease assay of SNM1A on crosslinked and duplex DNA. **(C)** Nuclease assay of Lambda exonuclease on crosslinked and duplex DNA.

The finding that SNM1A has translesion nuclease activity on psoralen and cisplatin ICLs suggested that SNM1A might act directly on these ICLs as an endonuclease. Psoralen intercalation unwinds DNA by approximately 25°, resulting in significant local distortion and base stacking disruption (**Figure 24**A) (Noll et al., 2006). Cisplatin induces a helical bend of 47º and causes cytosine bases to adopt an extra-helical conformation, unwinding the DNA by 110º (**Figure 24**B) (Noll et al., 2006). To address possible endonuclease activity on unstacked bases of cisplatin and psoralen ICLs, the ICL substrates from **Figure 22** and **Figure 23** were modified such that the 5'P was changed to a 5'F. This eliminated all possible exonuclease-mediated translesion activity. In the context of this DNA structure however, no product was observed with either 8MOP or cisplatin ICL substrates (**Figure 24**A and **Figure 24**B). Thus, within the context of fully complemented duplex DNA with a single ICL, helical distortion and base unstacking are not sufficient to stimulate SNM1A endonuclease activity.
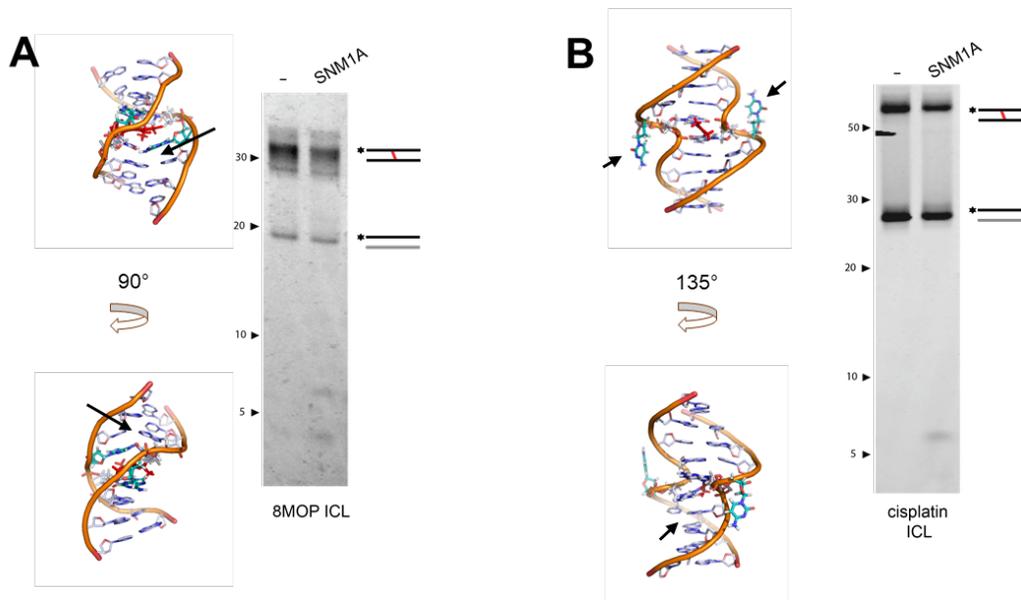
Figure 24: Analysis of direct SNM1A endonuclease activity on cisplatin and psoralen crosslinked DNA

(A) Structure of psoralen crosslink (PDB ID: 204D) and nuclease assay on 5'fluorophore labeled 8MOP ICL substrate. **(B)** Structure of cisplatin crosslink (PDB ID: 1DDP) and nuclease assay of SNM1A on cisplatin ICL substrate. Black arrows indicate unstacked bases.

3.2.3 Comparing SNM1A Nuclease Activities

Our work probing SNM1A translesion endonuclease activity utilized substrates that are unlikely present during ICL repair, particularly in replication-coupled repair. It is thought that DNA 5' to the ICL is unpaired and cleaved, presenting DNA with a 5'P end for exonuclease activity. It is currently thought that the role of SNM1A in ICL repair is to trim DNA around the ICL as an exonuclease, particularly since SNM1A has ability to act as a translesion nuclease. Given that SNM1A also demonstrates endonuclease activity, we wondered if these nuclease activities were consistent with a model in which SNM1A cuts 5' to an ICL to create a nick that would be used for full processing of an ICL adduct using exonuclease activity. Therefore, to observe how SNM1A might act in context of this model, we created a substrate to emulate ICL-dependent replication fork stalling with a 5' flap and a downstream ICL incorporated in the duplex region (**Figure 25**A). Duplex DNA on the unlabeled strand represents the leading strand approaching the crosslink, while the 5' flap on the labelled strand serves as the lagging strand. Although SNM1A had minimal translesion nuclease activity on SJG-136 crosslinked DNA, this ICL-inducing agent was used to generate the crosslink since heat reversing the ICL facilitated analysis of only labelled products. The distinct nuclease functions of SNM1A (endonuclease, exonuclease, and translesion nuclease activity) were compared using the same substrate with different modifications (**Figure 25**B).
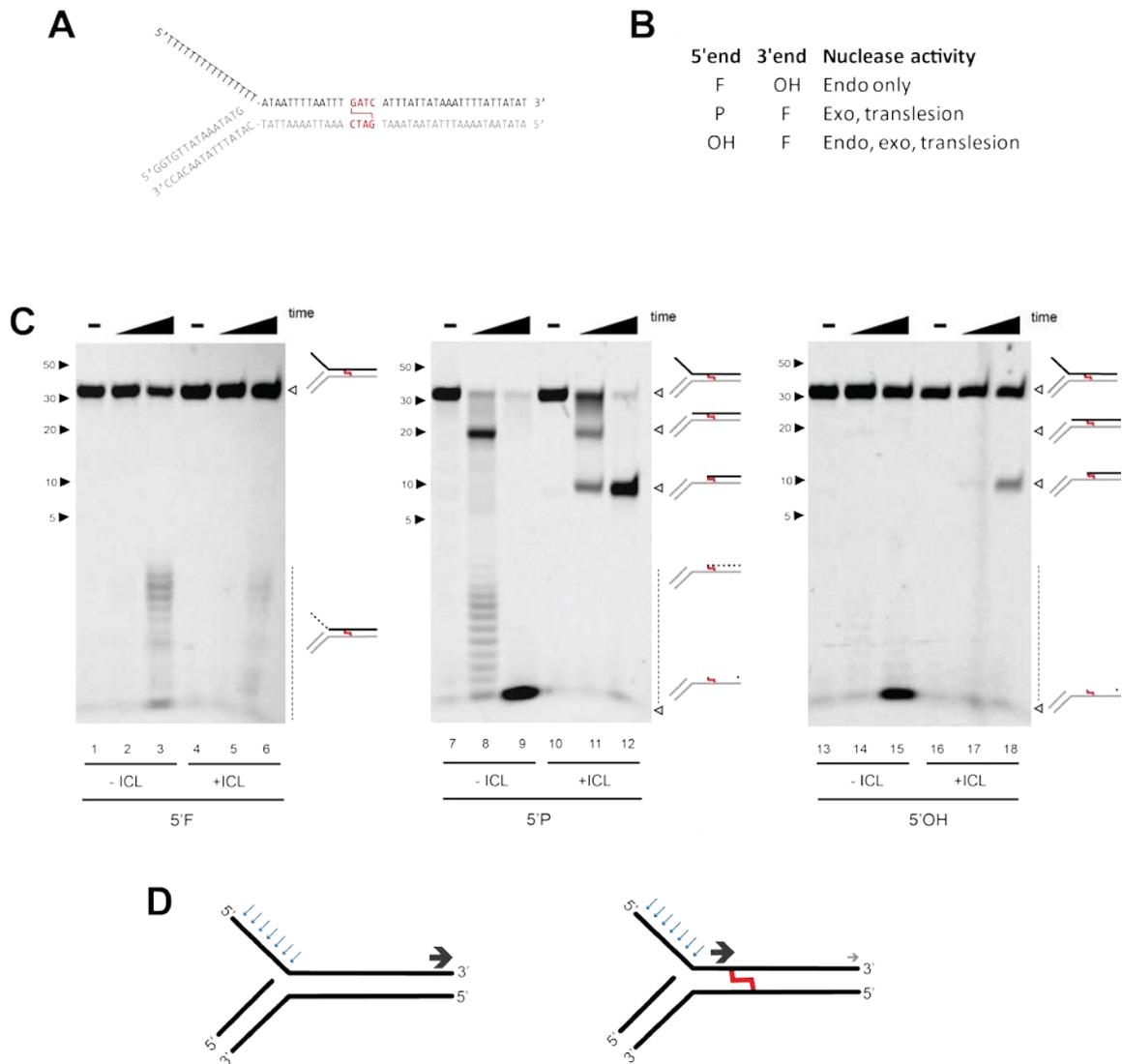
Figure 25: SNM1A nuclease activities on crosslinked replication fork-like substrate.

(A) DNA substrate representing a stalled replication fork (5'flap) containing a single SJG-136 crosslink. **(B)** Substrate end-modifications and corresponding nuclease activities. **(C)** Time course assays of substrate in (A) with 5'end modifications in (B). **(D)** Summary of nuclease events in (C). Arrows perpendicular to DNA indicate initial endonuclease cleavage. Solid arrow parallel to DNA indicates exonuclease product. Grey arrow indicates translesion nuclease product.

As shown in **Figure 25**C, uncrosslinked and crosslinked substrates containing a 5'F generated a similar series of endonuclease products within the unpaired region (**Figure 25**C, lanes 3, 6). Decreased processing of the crosslinked substrates was noted and may reflect decreased affinity of SNM1A for this substrate. When the same substrate contained a 5'P and 3'F, a 40nt product corresponding to digestion of the flap up to the single-to-double strand transition was observed with both crosslinked and uncrosslinked DNA at 10 minutes (lanes 8, 11 and 14). However, unlike the fully digested uncrosslinked substrate, incorporating an ICL in the 5'P substrate resulted in a strong block of exonuclease progression at the crosslink (lanes 11, 12). This block was also seen with the 5'OH substrate (lane 17, 18), indicating that the nick on the flap produced by SNM1A endonuclease function permitted subsequent exonuclease activity to the ICL. Taken together, these results suggest that SNM1A has the required activities to perform strand incision to unhook and fully remove an ICL adduct.

3.3   High-Throughput Screening for SNM1A Inhibitors

The demonstrated nuclease activities of SNM1A may have a greater role in ICL than previously appreciated, particularly if SNM1A is required as a key nuclease for ICL unhooking. SNM1A depletion and, more importantly, catalytically inactive SNM1A in human cells are hypersensitive to ICL damage. This suggests that if SNM1A nuclease activity is inhibited *in vivo*, cells may be sensitized to crosslinking agents. We therefore used a high-throughput screening (HTS) approach to identify small molecules that inhibit SNM1A nuclease activity.

An HTS assay requires a "window" in which the signal to noise is sufficiently limited in order to identify statistically significant changes in signal. The Z' score indicates a screening window that spans at least 12 standard deviations from the mean of the positive and negative controls, reducing identification of false positives. A fluorescence-based assay monitoring SNM1A exonuclease activity had been reported and was further developed within our lab (Sengerová et al., 2012; Huang, 2013). This assay used a ssDNA substrate containing a 5'P with a internal fluorophore (fluorescein) and quencher (black hole quencher 1) pair. SNM1A exonuclease activity results in separation of the fluorophore and quencher and increased fluorescence (**Figure 26**A). We used this assay to detect inhibition of SNM1A exonuclease activity through observation of decreased fluorescence compared to uninhibited SNM1A. Assay development in our lab demonstrated a wide assay window, with a Z' factor as high as 0.88 (Huang, 2013). Unfortunately, the assay that had been developed required a considerable amount of

purified SNM1A protein, limiting the number of compounds that could be screened.

3.3.1 Assay Optimization

Given the challenges of purifying SNM1A, we needed to further optimize the HTS assay parameters to minimize the required amount of SNM1A since simply reducing the amount of SNM1A in the assay negatively affected the screening window. BSA is commonly used in assays to prevent non-specific interaction, particularly to the plastic reaction vessel, without interfering with nuclease reactions. The original HTS assay design omitted BSA due to concerns of non-specific binding of BSA to small molecules in the screen, particularly more hydrophobic compounds, thus reducing potential effects on SNM1A. However, as seen in **Figure 26**B, addition of BSA dramatically increased the activity of SNM1A by 2.7-fold, likely by reducing SNM1A interactions with the plastic wells. Although addition of BSA had the potential of limiting the number of hits identified by HTS, adding BSA improved SNM1A activity enough to justify its use.

SNM1A activity was further improved by changing the pH from the original assay reaction (pH 7.2). It was previously shown that the pH optimum of SNM1A ranged from 8.6-9.1 (Sengerova, 2011). Furthermore, work in our lab suggested that higher pH rendered SNM1A more soluble. As seen in the reaction curve in **Figure 26**C, substrate was fully processed within the first 20 minutes at pH 9.1 whereas the reaction at pH 7.2 had not completed at 40 minutes. Increasing the pH also increased the signal of fluorescein, which has a pK_a of 6.75 (Doughty, 2010). Indeed, in order to adequately compare the reaction rate at pH 7.2 and 9.1, all points generated at pH 9.1 had to be

divided by a factor of 2.8 to account for the increased fluorescence.

Together, modifications to the original reactions conditions had sizable changes to the HTS assay (**Figure 26**D). Under these new conditions, the endpoint of the assay was set at 40 minutes, as this time point ensured unambiguous completion of digestion. All subsequent reactions were based on this standardized SNM1A activity. Addition of BSA increased final product formation nearly 3-fold. Changing the SNM1A reaction buffer to pH 9.1 not only increased the reaction rate to a predicable end-point, but also increased fluorescence signal. Collectively, these two changes decreased the amount of SNM1A required per reaction by 90% from 50nM to 5nM.
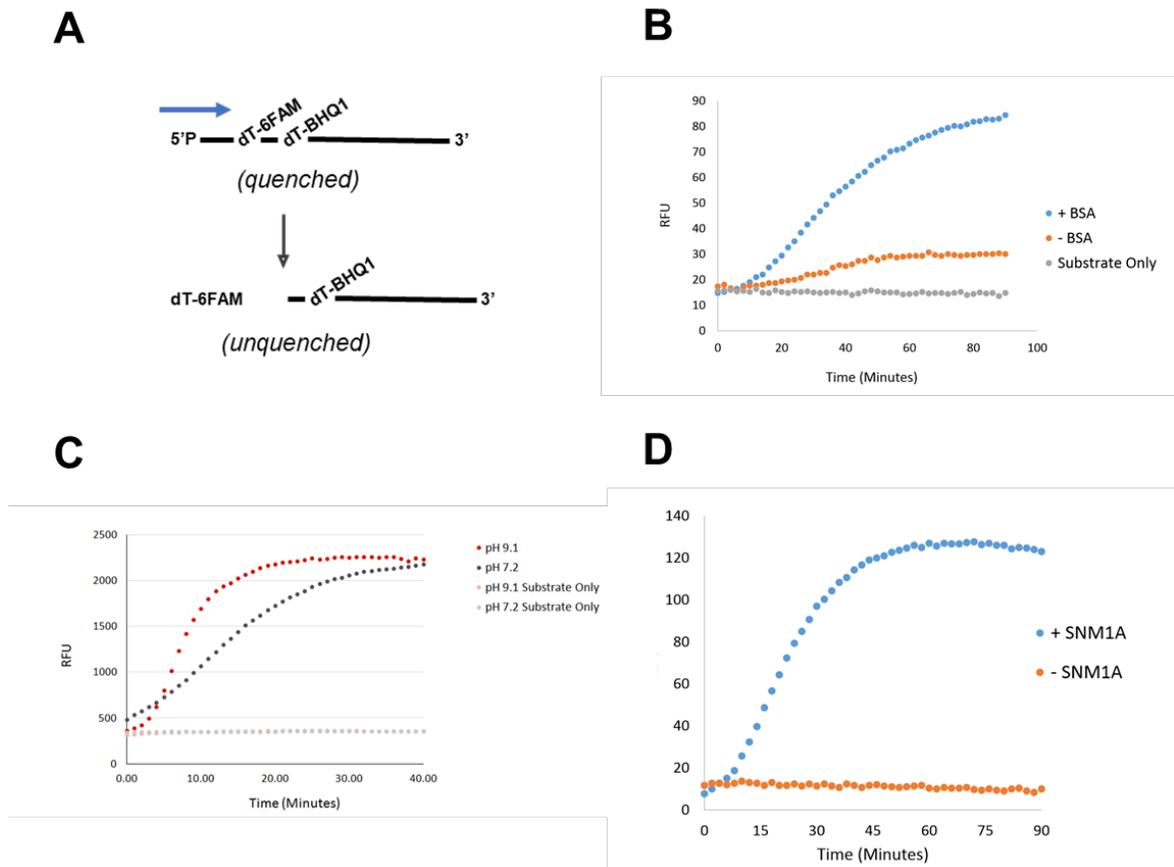
Figure 26: Optimization of SNM1A exonuclease activity for high-throughout screening (HTS)

(**A**) HTS substrate and schematic of nuclease assay. (**B**) Effect BSA on SNM1A assay. (**C**) Effect of pH on SNM1A assay. (**D**) Total effect of optimized parameters on SNM1A nuclease activity.

3.3.2 K_M Determination

Small molecules can act as reversible inhibitors competitively, uncompetitively, and non-competitively. While non-competitive inhibition is independent of substrate concentration, detection of competitive and uncompetitive inhibitors can be confounded by enzyme and substrate concentrations in a reaction. An ideal HTS assay uses the same substrate concentration as the K_M of the enzyme in order to capture all inhibitors, regardless of class. At K_M, both unoccupied and substrate-occupied enzyme are equivalent. Excess substrate outcompetes competitive inhibitors, while insufficient substrate does not generate enough signal to identify a non-competitive inhibitor compared to the control.

Using the optimized reaction conditions, the K_M of SNM1A was found to be approximately 8nM (**Figure 27**A). While this concentration would maximize the number of hits in the screen, the resulting screening window was not sufficiently wide enough. Protein and DNA concentrations were subsequently adjusted for HTS assay conditions while maintaining K_M. To ensure that all parameters, including reaction conditions, protein, and DNA concentrations, as well as robotic liquid handling would yield statistically significant hits, the Z' score was redetermined under fully automated conditions (**Figure 27**B). This resulted in a Z' score of 0.61 (**Figure 27**C). Although this was less robust than the original Z' score with the unoptimized reaction conditions, ten times as many compounds could be screened with the improved HTS assay while maintaining a suitable screening window.
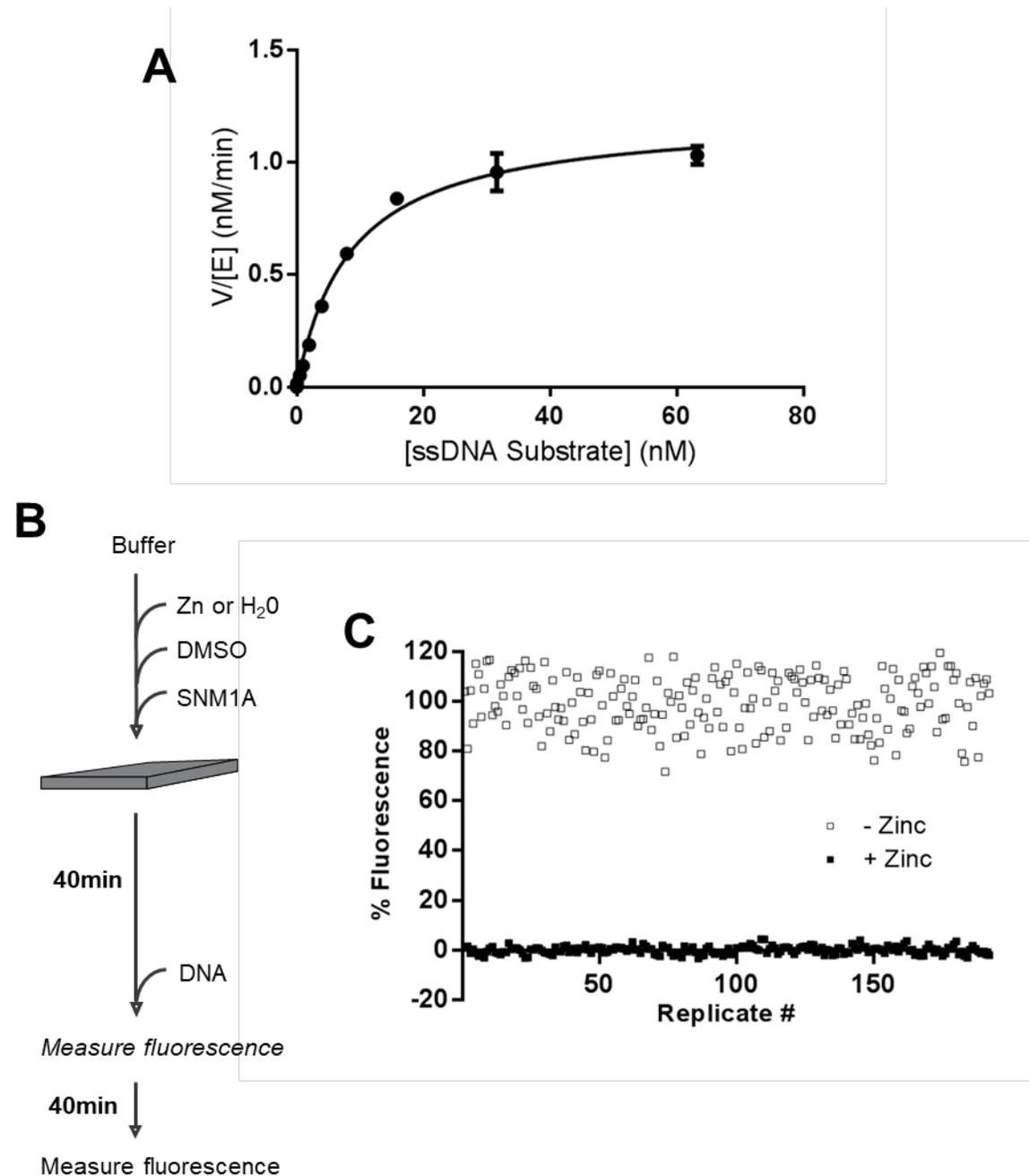
Figure 27: K_M and Z' determination of SNM1A in optimized HTS assay

(**A**) K_M determination. (**B**) Schematic of automated HTS assay. (**C**) Z' determination of HTS assay with uninhibited and inhibited SNM1A.

3.3.3   Pilot Screen

To ensure proper execution of the automated method in **Figure 28**A, an initial pilot screen was run with a single plate of off-patent drug compounds, assayed in duplicate. In a 384-well plate, 320 compounds were tested alongside 64 uninhibited and zinc-inhibited controls, included for internal Z' factor determination. The pilot Z' score for each of the duplicate plates was 0.53 and 0.65, indicating that the screening window was sufficient to identify inhibitors. The high control was also used to determine the hit rate cut-off, which was set at three standard deviations below the high control mean.

Using this hit rate cut-off based on the high controls, in this case 53.3% (boxed area in **Figure 28**B), the initial pilot screening identified 39 compounds which decreased fluorescence (red dots in **Figure 28**B). This constituted a hit rate of 12.2%, significantly higher than 1%, the expected hit rate for an HTS screen. This difference likely reflects the small sample size of the pilot screen. Interestingly, 30 of the 39 compounds contained heterocyclic nitrogen groups, where 12 of those have a nitrogen-containing polycyclic aromatic core. This core, similar to the nitrogenous base of DNA, may rationalize the high number of possible inhibitors identified, since these compounds may compete for DNA binding residues within SNM1A. Overall, the results of the pilot screen were reasonable and indicated that the HTS assay for SNM1A inhibition was ready for a full screening campaign.
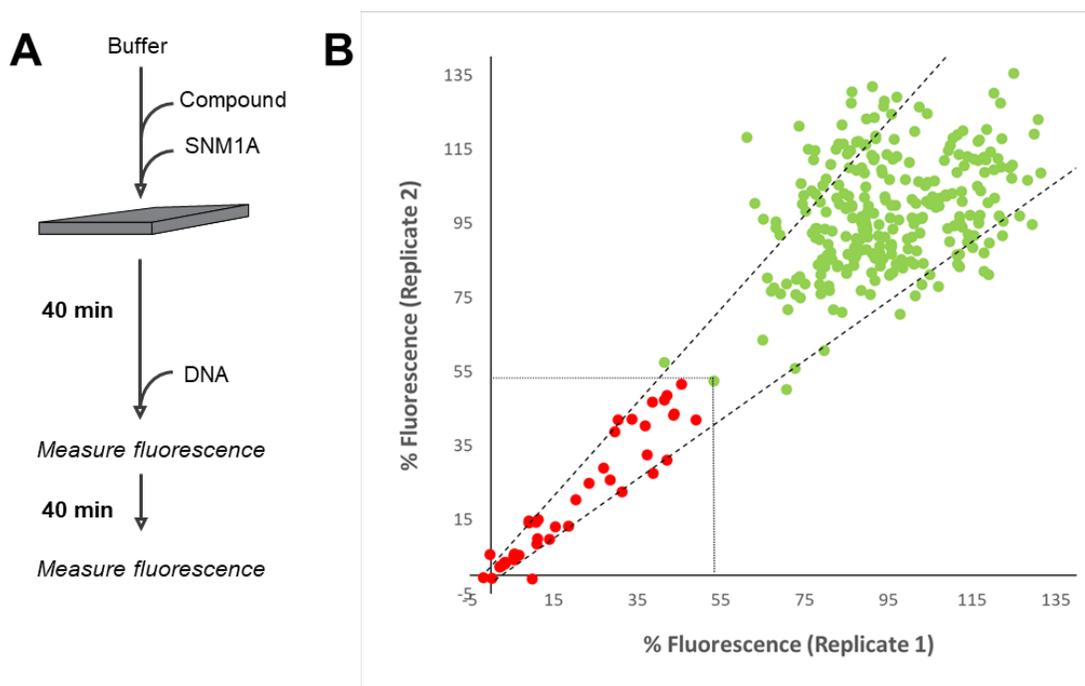
Figure 28: Pilot assay for SNM1A inhibitors

(A) Schematic of pilot assay **(B)** Replica plot of pilot assay of 320 compounds. Diagonal dashed lines indicate standard deviation of high controls. Perpendicular dotted lines indicate cut-off for hit identification. Red dots indicate pilot assay hits and potential SNM1A inhibitors. Green dots indicate compounds that did not meet the hit cut-off level for SNM1A inhibition.

3.3.4 Primary HTS Screen of Bioactive Library

We limited the scope of our screening campaign to a curated library of just under four thousand bioactive compounds, encompassing natural products, off-patent FDA-approved drugs, and druglike synthetic small molecules. Limiting screening to the bioactive library was expected to improve success of further inhibitor characterization using cell-based assays. Given the number of compounds assayed, data normalization was crucial to properly identify hits. Final product fluorescence values for each compound was normalized with respect to high and low controls on each individual plate to adjust minimum and maximum signal. Controls-based normalization permits fluorescence comparison amongst plates, not just within individual plates. **Figure 29**A illustrates variability of each plate and the effect of controls-based normalization to allow plate to plate comparison.

The influence of outliers can be minimized using the interquartile mean (IQM), which represents the average of the second and third quartiles, or all values between 80 to 240 of a sorted data set of 320 compounds. IQM-based normalization flags plates that overrepresent compounds in the bottom and top-most quartiles. This was particularly important because many of the bioactive compounds are intrinsically fluorescent. The effect of IQM-based normalization is evident in **Figure 29**B, where the data points are more consistent after normalization. Using histogram analysis, the effect of IQM-based normalization transformed the horizontal spread of data towards a Gaussian distribution and minimized effect of outliners in the outer quartiles (**Figure 29**C).

Using the assay protocol in **Figure 31**A, 3941 small molecules on 26 separate plates were screened for SNM1A inhibition in duplicate. The high controls-based hit rate cut-off was 54.22% (boxed area in **Figure 31**B). After both controls-based and IQM-based normalization, the number of hits decreased from 114 to 53 compounds (red dots in **Figure 31**B). The normalized number of hits represented a 1.3% hit rate, as expected for an HTS screen.
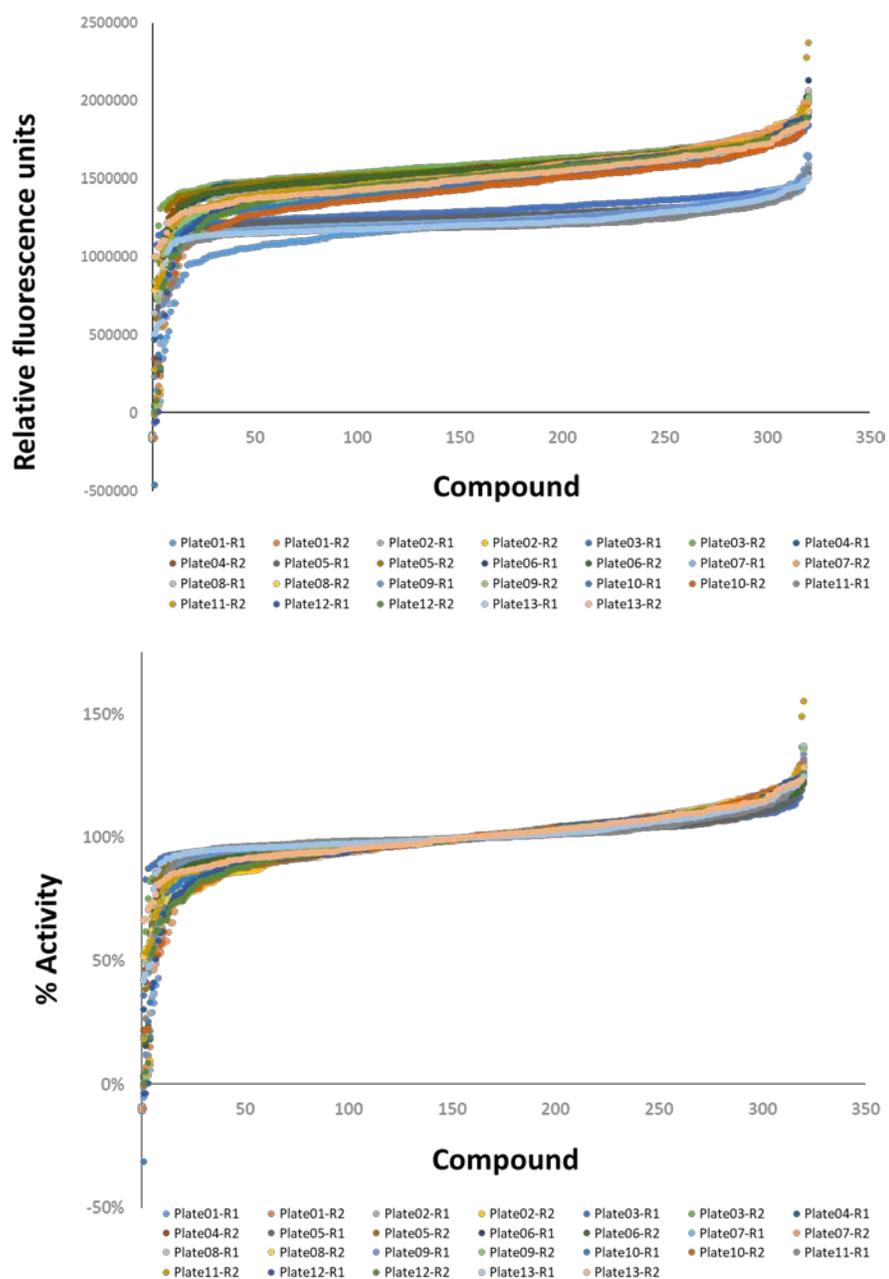
Figure 29: Controls-based normalization

Scatterplot of compounds before (**A**) and after (**B**) controls-based normalization.
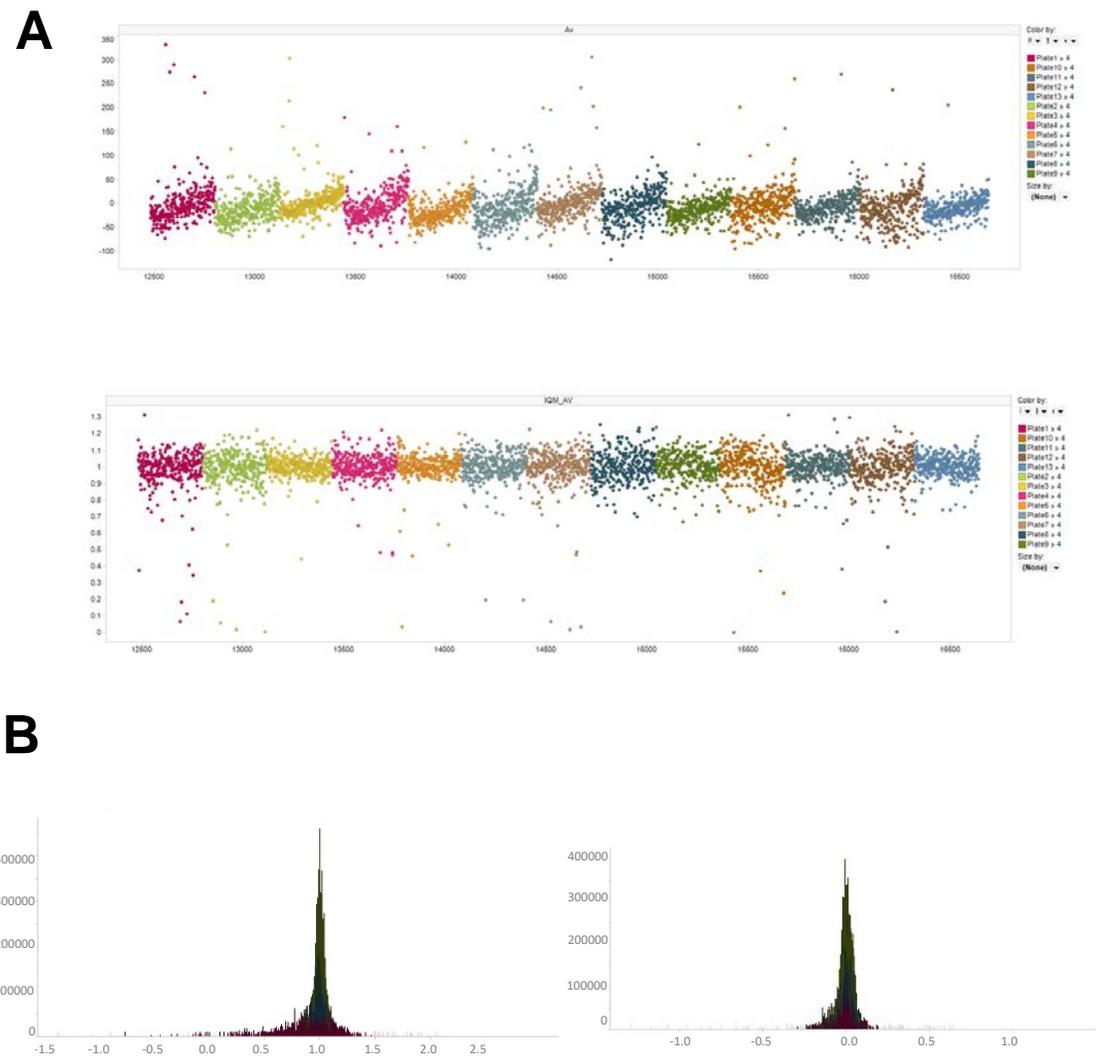
Figure 30: Interquartile mean normalization

(A) Scatter plot of compound plates before (top) and after (bottom) interquartile mean (IQM) normalization. **(B)** Histogram plot of all compounds before and after IQM normalization.
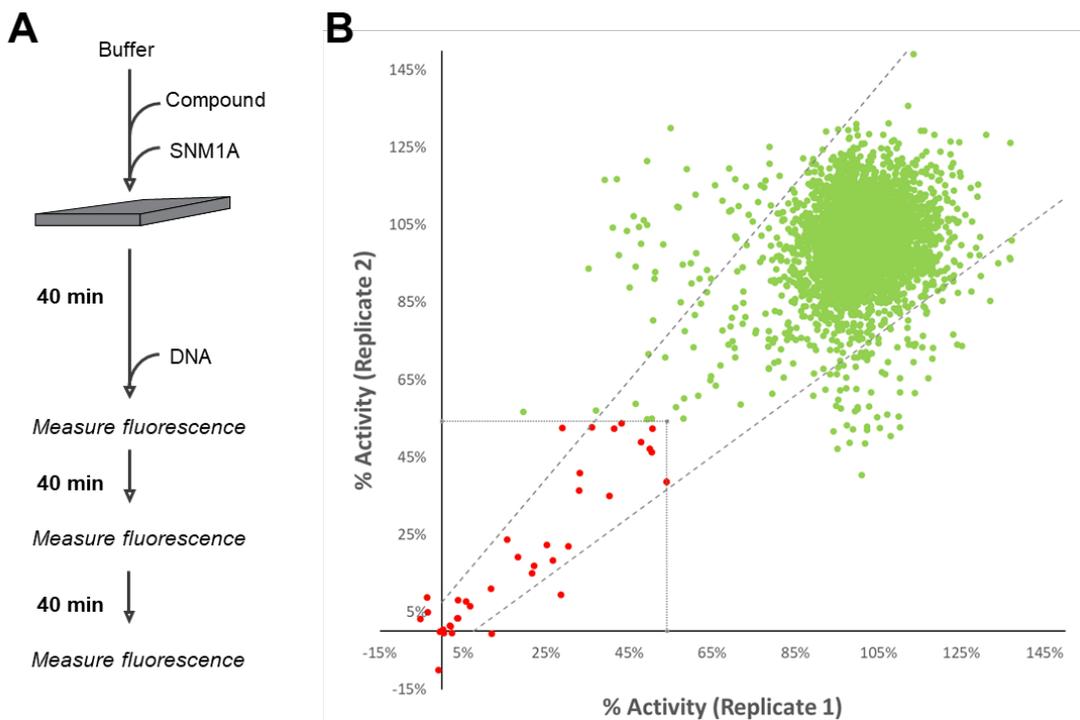
Figure 31: Primary HTS screen for SNM1A inhibitors

(A) Schematic of HTS assay **(B)** Replica plot of HTS campaign of 3941 compounds. Diagonal dashed lines indicate standard deviation of high controls. Perpendicular dotted lines indicate assay cut-off for possible inhibition. Red dots indicate HTS assay hits and possible SNM1A inhibitors. Green dots indicate compounds which do not meet the cut-off and do not inhibit SNM1A.

3.3.5   Validation of HTS Hits by Dose-Response Analysis

To further characterize identified compounds from the primary HTS screen, a secondary dose-response assay was used to validate that the decreased fluorescence was due to the compound of interest. The dose-response assay was not only to validate that the compound of interest reproducibly showed inhibition, but also to eliminate false positives due to plate position. Inhibitors of interest were required to demonstrate decreasing fluorescence with increasing compound concentration. Starting with the same concentration used in the primary HTS assay as the high concentration, compounds were diluted two-fold eight times and assayed. Although 53 compounds met the controls-based cut-off, the top 64 compounds were retested to maximize plate requirements and SNM1A use. Of the 64 compounds tested, 22 displayed reasonable dose-dependent responses (**Figure 32**) and were selected for further confirmation using a secondary gel-based assay.
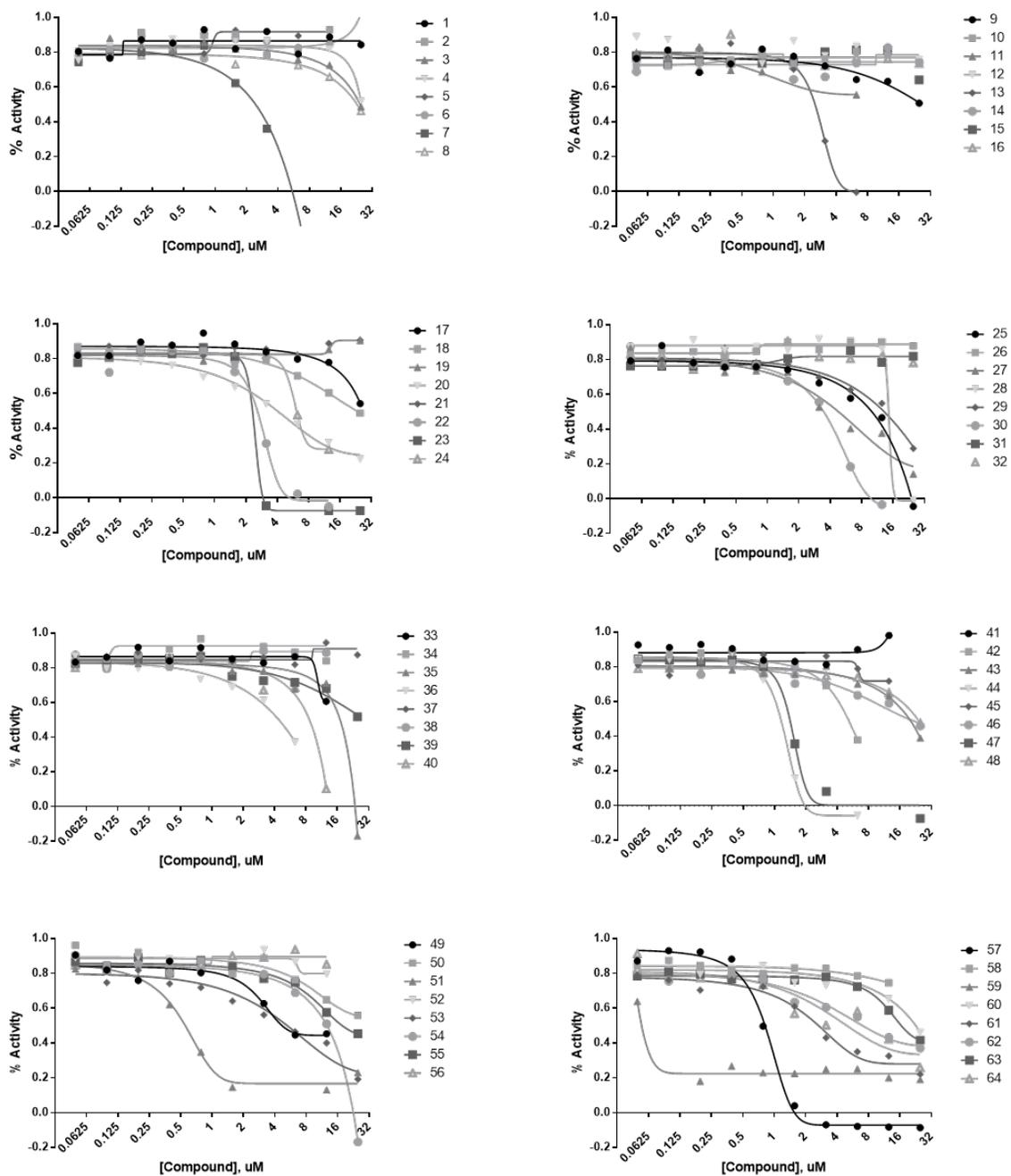
Figure 32: Dose-response analysis of hits identified by primary HTS

Dose-response curves of top 64 compounds identified as hits from HTS screen. Individual compounds are labeled in sets of 8.

3.3.6 Secondary Gel-Based Validation and Characterization

Although the primary screen was robust and reproducible, its fluorescence readout was problematic given that many compounds in the bioactive library are intrinsically fluorescent. Therefore, performing an orthogonal assay was critical to confirm SNM1A exonuclease inhibition. Gel-based exonuclease activity assays utilizing a ssDNA substrate containing a 5'P and 3'F were used to validate inhibition of SNM1A. Two compound concentrations were tested; 25uM (as used in the primary screen) and 6.25uM. At this concentration, two compounds (44 and 54) exhibited problematic strong fluorescence in **Figure 33**A and were not further analyzed. Of the 22 compounds identified from the dose-response assay, 9 compounds showed at least moderate (>50%) inhibition of SNM1A (**Figure 33**B). Some differences were noted between % inhibition in the gel-based assay and % activity in the fluorescence-based assay. These differences may reflect the use of different substrates in each assay. Since the fluorophore-quencher pair in the HTS substrate was three bases from the 5'P, only three nuclease events would be required for positive identification. However, the gel-based assay utilized a substrate with a fluorophore 30 nucleotides from the 5' end. Therefore, one would expect that some compounds identified using the HTS screen would fail to meet to gel-based assay parameters for nuclease inhibition.
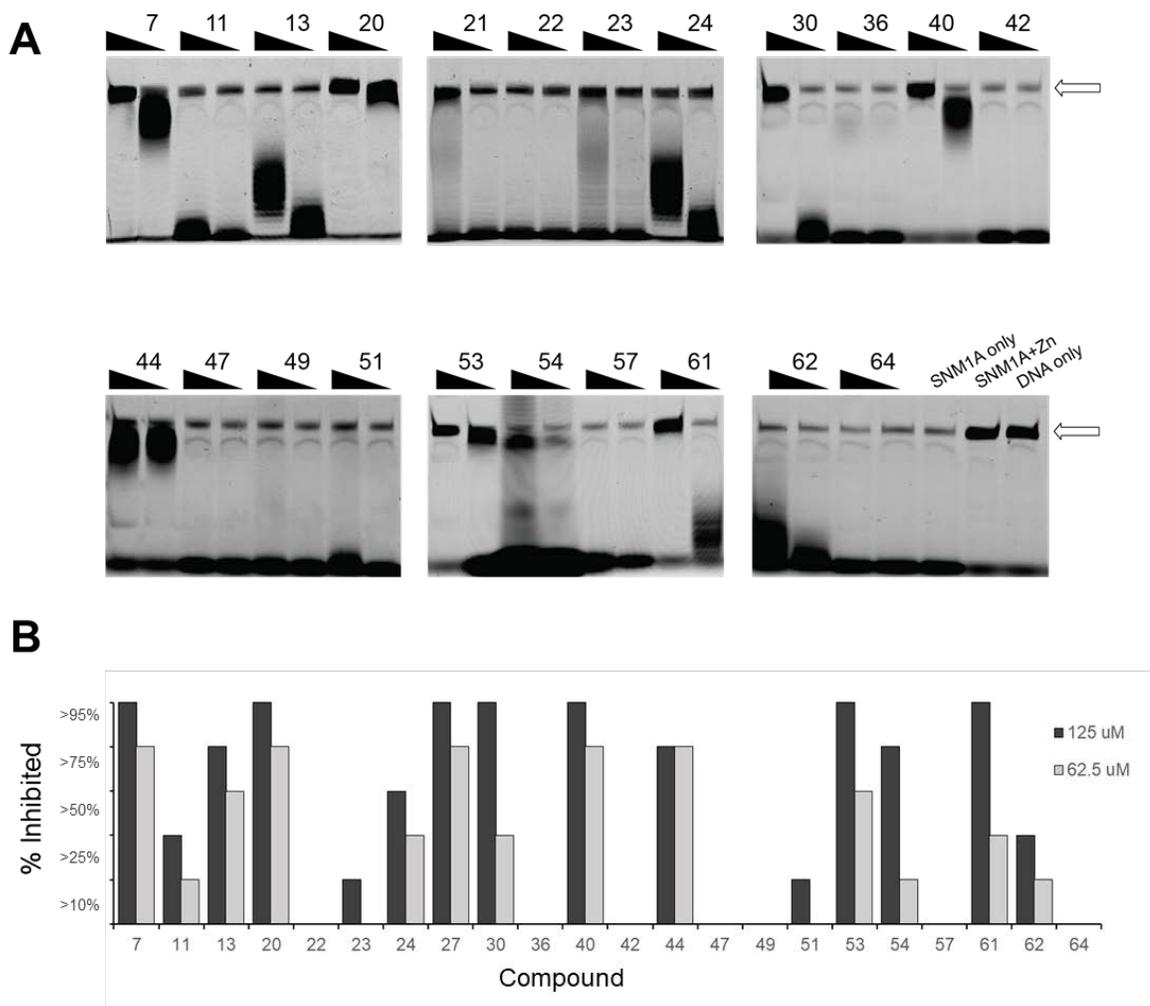
Figure 33: Gel-based secondary screen for of SNM1A inhibitors.

(A) Dose response inhibition assay. (B) Quantification of SNM1A nuclease activity with HTS hits. Values are normalized with respect to nuclease activity of uninhibited SNM1A.

3.3.7 IC$_{50}$ Determination

Using K$_M$ concentrations determined in **Figure 18**A, the concentration of each compound required to inhibit half the nuclease activity (IC$_{50}$) of SNM1A was determined. This measure of inhibitor potency also allowed comparison between exonuclease and endonuclease inhibition. As shown in **Figure 34**A, theaflavin digallate, ATA, DIDS were the most potent inhibitors of exonuclease activity, acting in the sub-nanomolar range. Furthermore, theaflavin and ATA appear to preferentially inhibit exonuclease activity since more than a 10-fold increase in inhibitor concentration was required to inhibit endonuclease activity to similar levels (**Figure 34**B). Conversely, only 6-OH DOPA and amentoflavone seemed to exert more endonuclease over exonuclease inhibition, although this difference is much less pronounced. These differences are significant, as they may help to address the endonuclease and exonuclease requirements for SNM1A activity in ICL repair.

A

B

Inhibitor	Exonuclease IC50 (uM)	Endonuclease IC50 (uM)	Fold difference of endo/exo
Theaflavin digallate	0.08	1.9	25.1
Aurintricarboxylic acid [ATA]	0.12	1.6	13.3
4,4-diisothiocyanostilbene-2,2-sulfonic acid sodium salt [DIDS]	0.36	2.1	6.0
Reactive Blue 2	1.68	22.1	13.2
Methyl-7-deshydroxypyrogallin-4-carboxylate [DHPG]	2.84	NA.	-
GW5074	12.46	30.7	2.5
6-hydroxy DOPA	35.25	25.2	0.7
Amentoflavone	38.01	27.1	0.7
Epigallocatechin gallate	68.22	195.4	2.9

Figure 34: IC$_{50}$ determination of small molecule SNM1A inhibitors

(**A**) Log scale graph of IC$_{50}$ concentrations of SNM1A inhibitors. (**B**) Comparison of SNM1A exonuclease and endonuclease inhibition.

To fully characterize the inhibitors of SNM1A *in vitro*, we investigated the mode of action of the compounds by generating inhibited kinetic curves to compare with the uninhibited kinetic curve. Changes in inhibitor concentration without significant change in V_{max} are indicative of reversible competitive inhibition, where both inhibitor and substrate vie for the active site of an enzyme. Theaflavin and DIDS competitively inhibited SNM1A, as evident by the increased K_M without significant changes to V_{max} in the presence of inhibitor (**Figure 35**A). Changes in inhibitor concentration while maintaining a similar K_M suggests non-competitive inhibition, in which the inhibitor binds to the enzyme-substrate complex. Non-competitive inhibition was observed for amentoflavone and Reactive Blue 2 since SNM1A activity had a similar K_M, but differed in V_{max} in the presence of inhibitor (**Figure 35**B). All other compounds had showed mixed inhibition since neither K_M or V_{max} stayed constant after changing DNA concentrations.
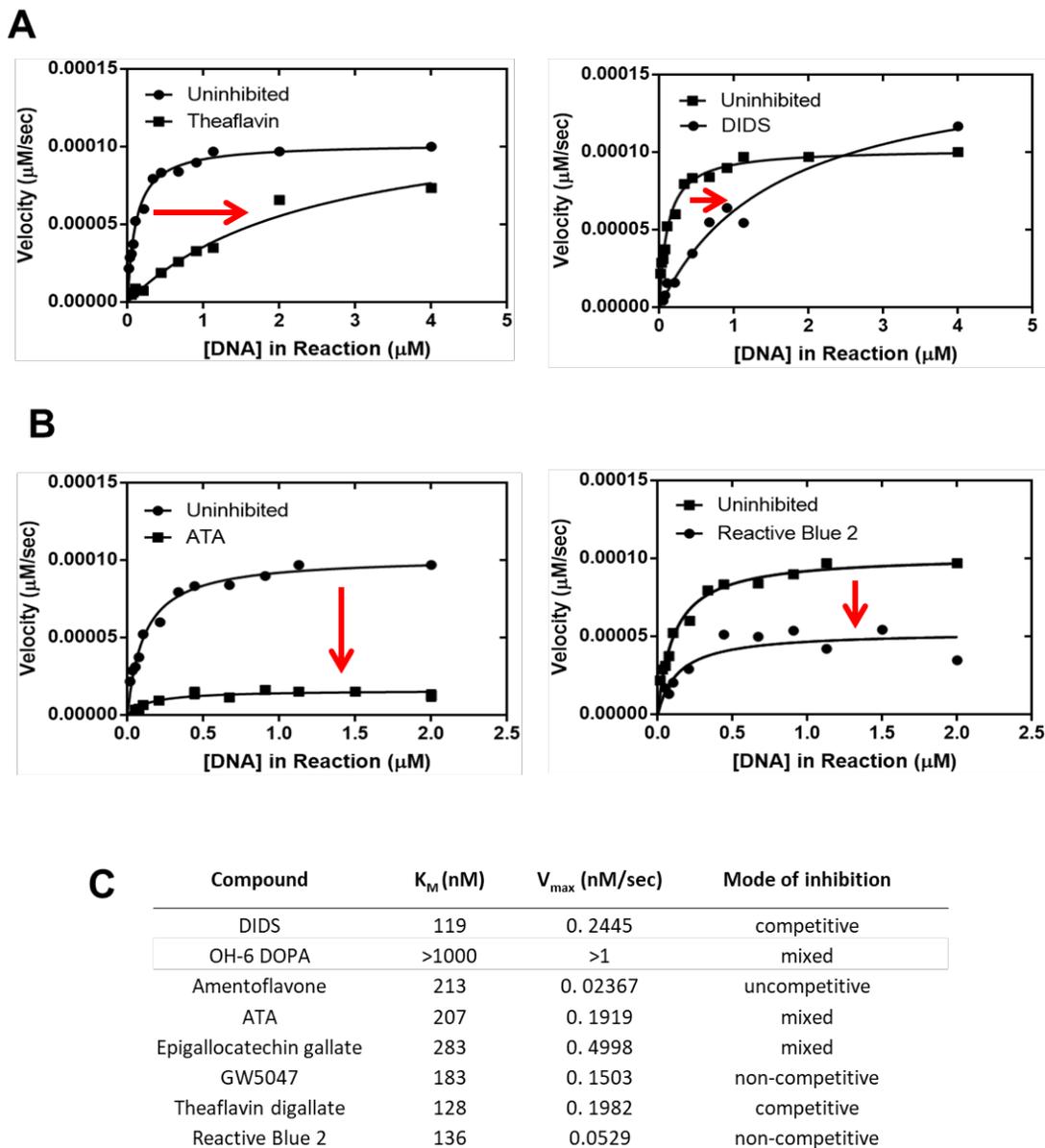
Figure 35: SNM1A inhibitor mode of action analysis

(**A**) Competitive inhibitors of SNM1A. (**B**) Non-competitive inhibitors of SNM1A. (**C**) Summary of mode of action inhibition analysis.

3.3.8 Nuclease Specificity

A good inhibitor should not only be potent, but also specific. To investigate the specificity of identified SNM1A inhibitors, compounds were analyzed against common exonucleases used in molecular biology. We tested Lambda exonuclease and RecJ exonuclease (both of which possess 5'-3' exonuclease activity, employed to create 3' overhangs in homologous recombination) as well as Pso2, for possible inhibition by SNM1A inhibitors (Han et al., 2006; Mitsis and Kwagh, 1999). As seen in **Figure 36**A, all SNM1A inhibitors were able to also inhibit Pso2 exonuclease activity at 25uM. This is a reasonable result given the functional homology between SNM1A and Pso2. Lambda exonuclease was moderately inhibited by ATA and considerably inhibited by theaflavin digallate. Little to no inhibition was observed for RecJ with inhibitors of SNM1A.

Inhibition of standard molecular biology endonucleases were also tested for inhibition. DNA single-strand specific endonucleases (Mung bean nuclease from *Vigna radiata* and S1 nuclease from *Aspergillus oryzae*) as well as structure-specific endonuclease (T7 endonuclease), were assayed for endonuclease activity using a stem-loop substrate in the presence and absence of SNM1A inhibitors (Kabotyanski et al., 1995; Mitsunobu et al., 2014). Minimal inhibition was observed for Mung bean and S1 nuclease, but 7 of the 8 inhibitors showed significant inhibition of T7 endonuclease (**Figure 36**B).

Overall, all compounds that inhibit SNM1A (as well as Pso2), at least one other exonuclease or endonuclease to some extent. Inhibition of commonly used nucleases

suggest a non-specific effect of the SNM1A inhibitor. However, variable inhibition of some nucleases and not others may reflect differences in protein and inhibitor concentrations. As molecular biology tools, these nucleases are measured in activity units, not molarity. Given this, specificity of the nine SNM1A inhibitors are an approximation, and further work is required to determine their non-specific effects, particularly in context of nucleases present in eukaryotic cells.
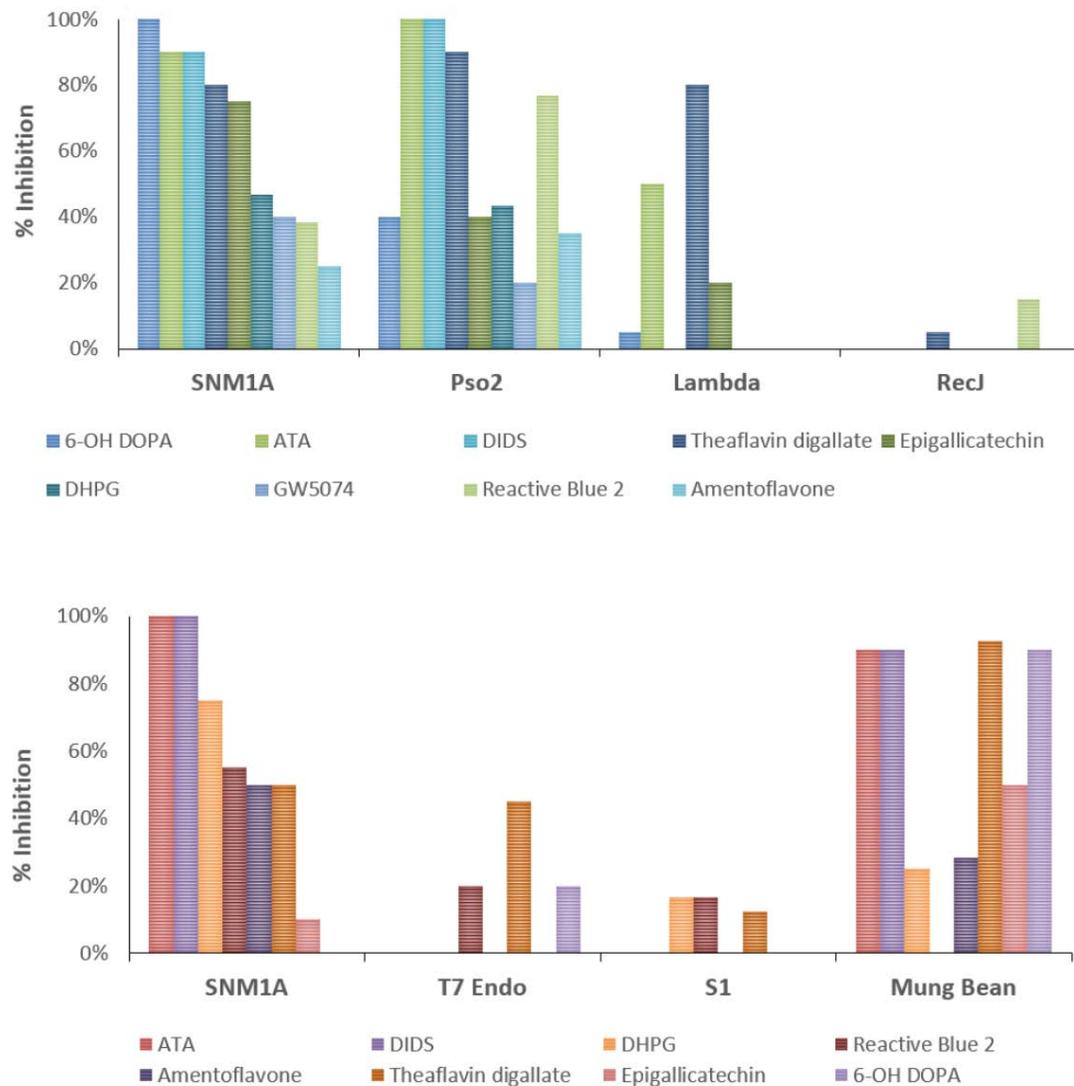
Figure 36: Analysis of SNM1A inhibitor specificity

(**A**) Comparison of exonuclease inhibition by SNM1A inhibitors. (**B**) Comparison of endonuclease inhibition by SNM1A inhibitors.

3.4   Cell-Based Cisplatin Sensitization Assay with SNM1A Inhibitors

In vitro, SNM1A is inhibited by at least 9 compounds. The question remained, do these compounds inhibit ICL repair *in vivo*? To answer this question, a cell-based viability assay was used to determine the effect of each of the SNM1A inhibitors after ICL damage induction. One difficulty with approach task is utilizing an ICL-inducing agent that results in ICL damage only. Cisplatin was chosen for these assays since it is the most clinically relevant ICL-inducing agent. Since cisplatin is a first-line chemotherapeutic in the treatment of many malignancies (such as cervical cancer), we made use of the cervical cancer-derived HeLa cell line for this assay. Importantly, a severe ICL hypersensitivity phenotype has been observed for SNM1A in HeLa cells (Wang et al., 2011).

Several assay parameters were optimized prior to testing possible cisplatin sensitization. Cell survival and proliferation were assayed with alamarBlue, in which blue resazurin is irreversibly reduced to red/pink fluorescent resorufin to measure the redox potential and aerobic capacity of cells (Rampersad, 2012). While this assay has a colorimetric output, under the conditions tested, a reproducible signal was difficult to obtain, and thus a fluorometric signal was measured instead. Other parameters requiring optimization for consistent measurement of cell survival included determination of the size and number of cells per well. In the end, 96-well plates were sufficient for reliably seeding and minimizing variability in duplicate wells. Finally, cisplatin kill curves were

generated in quadruplicate to determine the sub-lethal concentration of cisplatin required for ICL damage induction (**Figure 37**A).

Possible cisplatin sensitization was tested in HeLa cells by incubating cells with SNM1A inhibitor and inducing ICL damage prior to determining cell survival (**Figure 37**B). Cells were exposed to compound only (25uM) or with cisplatin IC_{10} and IC_{90} concentrations. The use of these two cisplatin concentrations provided a window to observe how an SNM1A inhibitor might potentiate cisplatin cytotoxicity. As shown in **Figure 37**C, exposure to GW5074 and possibly theaflavin digallate, alone, reduced cell survival (in grey bars). All other inhibitors did not decrease cell viability compared to the DMSO control. As expected, in the presence of cisplatin at its IC_{90}, cell survival with inhibitor (in navy bars) was essentially the same as cisplatin alone.

When a sub-lethal concentration of cisplatin was used (i.e. IC_{10} of cisplatin), nearly all inhibitors appeared to potentiate the effect of cisplatin (in blue bars). Due to variability observed with DIDS and DHPG, results for these two compounds were inconclusive. ATA, GW5074, and epigallocatechin modestly enhanced cisplatin cytotoxicity at least additively. Strikingly, four compounds (6-OH DOPA, amentoflavone, Reactive Blue 2, and theaflavin digallate) in combination with cisplatin (at IC_{10}) were as lethal as cisplatin alone at its IC_{90}. The observed synergy is significant, as it indicates that, while the inhibitor itself is non-toxic to the cell, it significantly potentiates the cytotoxicity of cisplatin.
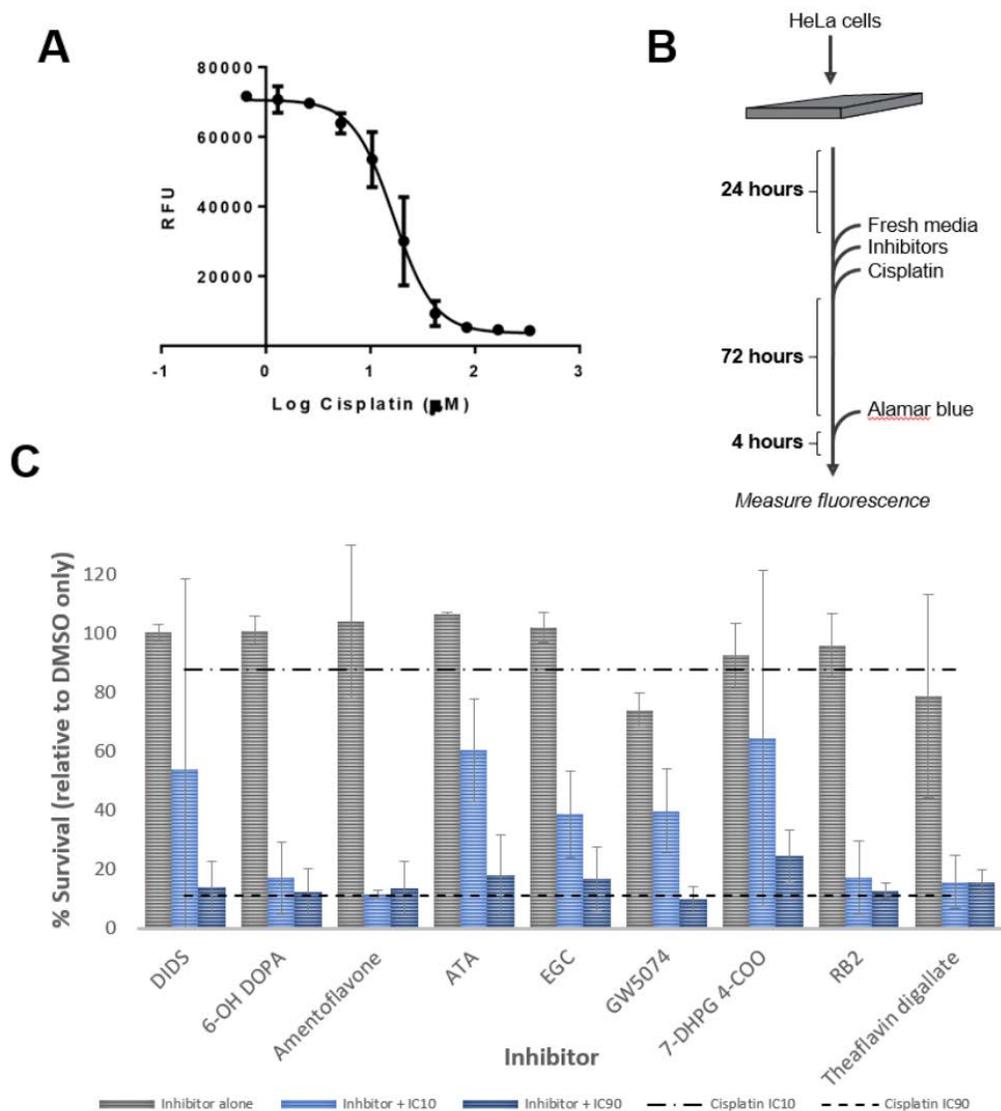
Figure 37: Cisplatin sensitization of HeLa by SNM1A inhibitors

(A) Cisplatin IC_{50} determination. **(B)** Schematic of cell-based cisplatin sensitization assay. **(C)** Graph of cisplatin sensitization by SNM1A inhibitors.

3.5   Summary of SNM1A Compound Inhibition

A promising inhibitor should be effective, specific, potent, and non-toxic. *In vitro* and cell-based analysis of SNM1A inhibitors suggest some compounds may be more promising than others, as summarized in **Table 1**.

Both **amentoflavone** and **6-OH DOPA** had similar inhibitor profiles, demonstrating low affinity for SNM1A, no toxicity alone, and synergism with cisplatin. However, amentoflavone had least cross-reactivity with other nucleases, while 6-OH DOPA inhibited three of five nucleases to some degree. **Reactive Blue 2** had moderate potency with an IC_{50} in the low micromolar range and moderate specificity, but displayed high synergism with cisplatin. **Theaflavin** digallate had the highest affinity for SNM1A *in vitro*, with an IC_{50} in the low- to mid-nanomolar range, however, it also inhibited all tested nucleases, suggesting that inhibition may be a non-specific effect. It is not unexpected then that, while theaflavin digallate potentiated the cytotoxicity of cisplatin to the greatest extent, it also showed toxicity alone.

Table 1: Summary of SNM1A inhibitors

Compound	Exo IC$_{50}$ (uM)	Endo IC50 (uM)	Inhibitor mode of action	Nuclease cross-inhibition	Compound alone (% survival)	Compound + cisplatin (% survival)
Theaflavin digallate	0.07715	1.937	Competitive	Significant	79	16
ATA	0.1185	1.575	Mixed	Moderate	107	61
DIDS	0.3556	2.116	Competitive	Moderate	100	54
Reactive Blue 2	1.68	3	Non competitive	Moderate	96	17
DHPG	2.839	22.12	Not determined	Mild	93	65
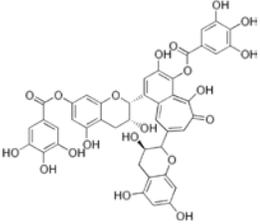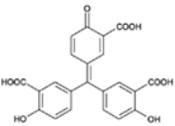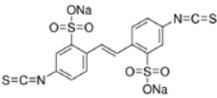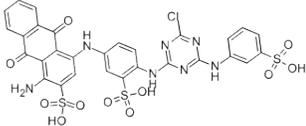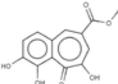
Compound	Exo IC$_{50}$ (uM)	Endo IC50 (uM)	Inhibitor mode of action	Nuclease cross-inhibition	Compound alone (% survival)	Compound + cisplatin (% survival)
6-OH DOPA	35.25	27.05	Mixed	Moderate	101	17
Amentoflavone	38.01	30.73	Uncompetitive	Minor	105	12
Epigallocatechin	68.22	195.4	Mixed	Moderate	102	39
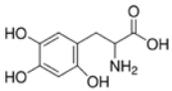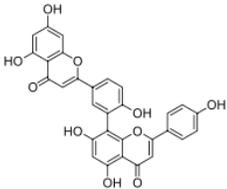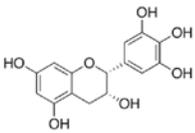
Although **epigallocatechin** was only cytotoxic in combination with cisplatin, the moderate effect was also accompanied by modest non-specific inhibition and the lowest affinity for SNM1A. **ATA** has been shown to be a non-specific nuclease inhibitor. Surprisingly, while it potently inhibited SNM1A, only two of five nucleases tested were inhibited by ATA. ATA was non-toxic, but only mildly potentiated cisplatin cytotoxicity. **GW5074** had an additive effect with cisplatin, but had low affinity and mild specificity, in addition to being the most toxic, making it a poor inhibitor of SNM1A. Both **DHPG** and **DIDS** had moderate to high affinity for SNM1, and mild to moderate non-specific inhibition, respectively. Given the variable effect with cisplatin, these last compounds should be tested again before concluding the value of these two compounds in SNM1A inhibition.

CHAPTER FOUR: DISCUSSION

Understanding the mechanisms of ICL repair process is of particular importance, given its implications in anti-cancer therapy. ICL repair requires concerted incisions to initiate careful removal of DNA damaged on both strands. SNM1A is important for proper repair of ICLs in human cells, but the molecular mechanisms underlying its contributions remain unclear. Despite the well-established role of Pso2 in yeast ICL repair, the importance of SNM1A in human ICL repair has been slowly emerging. Prior complementation studies demonstrating that SNM1A is able to rescue ICL repair defects in *pso2*-deficient cells suggest that SNM1A and Pso2 share common activities required for processing ICL damage (Hazrati et al., 2008). Indeed, both share 5' phosphate-dependent exonuclease activity and have been implicated in ICL repair (Hejna et al., 2007; Li et al., 2005). Despite these similarities, whether Pso2 and SNM1A also share endonuclease activity remained unclear.

Prior work likely missed endonuclease activity of SNM1A because this activity is much weaker than exonuclease activity *in vitro*. Indeed, exonuclease activity of SNM1A was found to be three orders of magnitude greater than endonuclease activity under the conditions tested. Since exonuclease activity of SNM1A is more robust than endonuclease activity, substrates labelled at the 3' end understandably failed to capture endonuclease activity. Endonuclease products are only clearly observed with a 5' labeled substrate since they are not subject to 5'-3' exonuclease activity.

Truncation of SNM1A to its catalytic domain may have also contributed to the

ability to observe SNM1A endonuclease activity. The SNM1A expression construct used in our studies was limited to the catalytic MBL/β-CASP domain (residues 698-1040), which differed from a previously characterized construct (SNM1A$^{N\Delta608}$) by 90 amino acids (Sengerová et al., 2011). In addition, the finding that freeze-thawing SNM1A$^{N\Delta608}$ diminished exonuclease activity and restrained endonuclease activity suggests that additional residues N-terminal to the MBL/β-CASP domain may hinder the ability to detect intrinsic endonuclease activity of SNM1A. Along the same lines, variations in the extent of SNM1A exonuclease processing have been reported that correlate to differences in length of the N-terminus. While full-length SNM1A is capable of processing ssDNA as an exonuclease, its ability to act on dsDNA is far reduced (Hazrati et al., 2008; Hejna et al., 2007). Full truncation to the catalytic core of SNM1A permits robust exonuclease activity on both ssDNA and dsDNA. Decreasing activity is observed with increasing length of sequence N-terminal to the MBL/β-CASP domain. Even the NusA solubility fusion N-terminal to the SNM1A catalytic core in our contruct prevented observation of endonuclease activity. Only when NusA was removed by TEV cleavage was endonuclease activity observed. These observations support the idea that intrinsic nuclease activities of SNM1A may be masked or regulated by the N-terminal regions of SNM1A.

4.1 Regulation of SNM1A Endonuclease Activity

Negative regulation of nuclease activity of the SNM1A MBL/β-CASP domain is expected *in vivo* since unauthorized nuclease function would be deleterious to the cell.

SNM1A regulation may be akin to auto-inhibition of the catalytic core of SNM1C (Niewolik et al., 2017). The β-CASP domain of SNM1C is sufficient for V(D)J recombination and endonuclease activity *in vitro*, suggesting that the non-conserved C-terminus of SNM1C negatively inhibits the nuclease activity of SNM1C (Poinsignon et al., 2004). Mutations in SNM1C identified from patients with SCID were used to elucidate key interactions between the beta-CASP domain and non-conserved regulatory regions (Niewolik et al., 2017). *In vitro,* V(D)J recombination with full-length SNM1C was increased by mutations disrupting self-association and regulation of the nuclease domain (Niewolik et al., 2017). Likewise, patients with SCID have mutations in similar residues that strengthen self-interaction of SNM1C, suggesting that the mechanism underlying the SCID phenotype is based on the inability of SNM1C to disassociate from itself. While it is unclear if full-length SNM1A is a homodimer in solution, it is possible that the N-terminus autoinhibits the endonuclease and/or exonuclease activity of SNM1A.

Although SNM1A endonuclease activity is relatively weak *in vitro*, it is quite likely that SNM1A interactions with other proteins may modulate SNM1A nuclease activity. Multiple proteins, many of which act in response to DNA damage, have been shown to interact with SNM1A. Some of these factors (as discussed below) may be required to localize SNM1A to sites of damage. It is also possible that SNM1A interactions with these proteins may modulate the endonuclease activity of SNM1A.

4.1.1   SNM1A Post-Translational Modification

Post-translational modification of enzymes is a key means of regulation given its

reversible nature. Indeed, full-length SNM1C requires DNA-PKcs autophosphorylation for endonuclease activity (Goodarzi et al., 2006). SNM1A has been shown to be phosphorylated *in vitro* by ATM (Akhter and Legerski, 2008). ATM belongs to the same PIKK kinase family as DNA-PKcs. It has been shown that ATM is required for SNM1A foci formation in response to IR and UVA damage (Akhter and Legerski, 2008a). ATM-dependent foci formation suggests that ATM phosphorylation is required for SNM1A mobilization within the cell, but it does not exclude the possibility that this phosphorylation modulates SNM1A nuclease activity. SNM1A in mice has been identified to be phosphorylated at serines 574 and 578 via phospho-proteome mass spectrometry (Huttlin et al, 2010). At this point, it is unclear if this phosphorylated residue is ATM-mediated or is mediated by another kinase.

Regulation of SNM1A nuclease activity may also result from sumoylation. The β-CASP domain of SNM1A interacts directly with sumoylation E3 ligase PIAS1 (Ishiai et al., 2004). Foci formation and cell survival were both found to be impaired in response to DNA damage when the SNM1A and PIAS1 interaction was disrupted. Interestingly, binding of SNM1A to PIAS1 is markedly reduced when SNM1A residues thought to be required for metal binding (H994, D838) are mutated (Ishiai et al., 2004). It is not surprising that catalytic mutants would decrease cell survival, but how these catalytic residues promote association between SNM1A and PIAS1 is unclear. While protein unfolding might account for these changes, at least one of these mutations (H994A) had no reported effect on SNM1A activity, suggesting no significant change to the secondary structure that would disrupt the SNM1A-PIAS1 binding (Sengerová et al., 2012). It is

possible that these residues induce a slight conformational change that promotes the interaction between PIAS1 and SNM1A. Since PIAS1 mediates sumoylation, it is also possible that interaction of SNM1A with PIAS1 results in the modification and modulation of SNM1A. Further work on the sumoylation state of SNM1A and its effect on nuclease activity are required.

4.1.2 SNM1A Protein-Protein Interactions

Although SNM1A does not require heterodimerization as other nucleases do (e.g. XPF-ERCC1, MUS81-EME1/2), a transient protein-protein interaction may prolong SNM1A at the site of damage. SNM1A directly interacts with CSB, a dsDNA-binding ATPase of the SWI/SNF family that is required to remodel DNA in NER (Iyama et al., 2015). CSB depletion results in hypersensitivity to cisplatin in stationary cells and MMC in proliferating cells, demonstrating a role for CSB in ICL repair (Enoiu et al., 2012; Licht et al., 2003). Direct binding of CSB was shown to stimulate the nuclease activity of SNM1A$^{N\Delta608}$. Increased catalysis by SNM1A corresponded to an increased affinity of SNM1A for DNA mediated by CSB (Iyama et al., 2015). This CSB-dependent effect may increase the ability for SNM1A to catalyze a weak endonuclease reaction. CSB has been shown to facilitate ICL unhooking, but it is unclear how SNM1A and CSB collectively affect this process (Iyama et al., 2015). CSB possesses ATPase activity that has been shown to remodel DNA and contains a conserved helicase domain. Unlike other SWI/SNF family members, helicase activity has not been observed for CSB *in vitro* (Selby and Sancar, 1997). Nevertheless, SNM1A-CSB complex formation clearly

modulates SNM1A activity on DNA, perhaps by inducing structural changes to DNA that facilitate SNM1A cleavage.

Another key partner of SNM1A is PCNA, which is better known for its central function in TLS. Monoubiquitination of PCNA not only recruits TLS polymerases, but also other repair factors in response to damage (Maga and Hubscher, 2003). Disruption of SNM1A-PCNA interaction through the conserved UBZ domain and PIP box of SNM1A impairs foci formation and cell survival in response to ICL damage (Yang et al., 2010). PCNA is a reasonable candidate for factors that may stimulate SNM1A endonuclease activity because it localizes SNM1A to cleavage sites. A similar mechanism is employed in bacteria, whereby the β-clamp (the PCNA homolog) regulates endonuclease activity of MutL in DNA mismatch repair. MutL endonuclease activity is required to cut mismatched DNA, but lacks nuclease activity in the absence of the β-clamp (Pillon et al., 2011). Although the small interface of MutL and the β-clamp promotes a transient interaction, it is sufficient to induce a conformational change required for MutL endonuclease activity (Pillon et al., 2011). Interaction of SNM1A and PCNA provides a clear mechanism for modulating SNM1A nuclease activity, particularly since monoubiquitination is a key modification that occurs in response to DNA damage.

Finally, while it is not known if SNM1A directly binds RPA, SNM1A is notably the only single strand-specific endonuclease involved in ICL repair. RPA binds to and stabilizes ssDNA to prevents random incision of unpaired DNA. Since SNM1A cleavage would occur on RPA-bound ssDNA, there is likely some interplay between these proteins. Interestingly, RPA preferentially binds to pyrimidine-rich patches, analogous to the

preferential activity of SNM1A for the same tracts of pyrimidine ssDNA. RPA determines the polarity and coordinates incisions by XPF-ERCC1 and XPG during NER (De Laat et al., 1998). RPA may similarly coordinate the incision by SNM1A in ICL repair. Hyperphosphorylation of RPA in response to DNA damage is mediated by 53BP1, an interacting partner of SNM1A (Yoo et al., 2005). Signalling amongst SNM1A, 53BP1 and RPA may therefore facilitate displacement of RPA and permit access of SNM1A to ssDNA for strand cleavage.

4.2   Translesion Nuclease Activity SNM1A

It had been generally assumed that the function of SNM1A in ICL repair is to trim ICL adducts using exonuclease activity. Reducing the number of nucleotides that flank a crosslink destabilizes the duplex and permits the ICL adduct to be flipped out. This in turn increases TLS by up to 80% (Ho et al., 2011). Observation of SNM1A digestion past the covalent linkage of an SJG-136 adduct supported an attractive model in which an ICL could be effectively unhooked by translesion nuclease activity of SNM1A. As compelling as the results of SNM1A translesion nuclease activity were on an SJG-136 crosslink, several questions remained. Why would SNM1A stall at the +3 position of an SJG-136 crosslink? Could translesion nuclease activity be due to hairpin-like endonuclease cleavage? Could substrate design result in an artifact suggestive of translesion nuclease activity? The fact that Lambda exonuclease can digest past the first covalent linkage of SJG-136 suggests that the adduct is not stable and supports the claim that SNM1A lacks translesion nuclease activity on this type of crosslink.

Recent work demonstrated that after RPA-mediated XPF-ERCC1 incision on the 5'
side of an ICL, SNM1A appeared to have translesion nuclease activity past a triazole ICL
(Abdullah et al., 2017). In this case, digestion was limited to a few nucleotides past the
ICL, which is inconsistent with the robust exonuclease activity that would be anticipated
past a crosslink. In this study, click chemistry was used to form the ICL (Abdullah et al.,
2017). As a result, the crosslinker (a hydrophobic amino hexyl chain) may interact with
the DNA or protein to inhibit processing. Without probing the substrate used in their
study with another exonuclease, it is not possible to assess if SNM1A has translesion
nuclease activity on this type of ICL.

Still, the possibility that SNM1A translesion nuclease activity might exist required
further investigation. Given that nucleotides with destabilized bases are more prone to
SNM1A cleavage, it was reasonable to test a hypothesis in which endonuclease activity of
SNM1A is sufficient to incise a distorting ICL with unpaired bases (i.e. from psoralen or
cisplatin). While SNM1A endonuclease activity did not appear to directly mediate ICL
unhooking, distorted ICLs were clearly bypassed by SNM1A translesion activity. Lambda
exonuclease stalling at cisplatin and psoralen adducts demonstrated the stability of the
substrate and specificity of SNM1A translesion nuclease activity. It is possible that
trimming to the ICL is required for access into the SNM1A active site, reconciling why
direct endonuclease incision was not observed. Alternatively, the destabilized duplex may
be required for approach by SNM1A, akin to TLS polymerases which more readily
approach highly distorting lesions at the cost of more easily bypassing the lesion (Roy
and Schärer, 2016). How SNM1A acts on lesions that vary widely in structure may be

more difficult to address since less distorting ICLs, such as MMC and malondialdehyde, are either unreactive or unstable, respectively, *in vitro*. As methods to synthesize ICLs of different conformations develop, the question of the types of ICL structures that can be acted upon by SNM1A can be answered more fully.

4.3 Possible Roles of SNM1A in ICL Repair

A longstanding question in the field of ICL repair is which endonuclease(s) are required to generate nicks flanking an ICL adduct. Work presented here demonstrates that the endonuclease activity of SNM1A has the ability to participate in ICL unhooking. Although several endonucleases could potentially fulfill the role of the other endonuclease in ICL unhooking, only SNM1A has been shown to work in the same pathway as XPF-ERCC1 (Wang et al., 2011).

4.3.1 Replication-Dependent Repair

In agreement with the observed epistasis between SNM1A and XPF-ERCC1, it has been suggested that XPF-ERCC1 may create an initial nick that serves as a substrate for SNM1A 5' exonuclease action in replication-dependent ICL repair. The preferred substrate of XPF-ERCC1 is a 3' splayed arm structure (De Laat et al., 1998b). If XPF-ERCC1 were to act on the 3' end of a simple fork, the incision would be 3' to the ICL, creating a substrate in which exonuclease trimming by SNM1A would move away from the ICL. Very recent work, however, demonstrated that RPA can position XPF-ERCC1 at a 3' splayed arm or 3' flap such that XPF-ERCC1 cleavage is positioned 6 nucleotides into the duplex region on the 5' side of an ICL(Abdullah et al., 2017). Such cleavage

would therefore generate substrate for SNM1A exonuclease trimming, however, our lab and others have shown that the 5' nicked intermediate is a poor substrate for SNM1A exonuclease activity, likely since access to the 5' phosphate is sterically hindered (Sengerová et al., 2012).

During replication-dependent ICL repair, regions of unpaired DNA occur on the lagging strand, 5' to the ICL, providing an ideal substrate for SNM1A endonuclease activity. In such case, SNM1A endonuclease activity on the lagging strand would generate the necessary 5' phosphate for SNM1A exonuclease trimming up to and past an ICL. Monoubiquitination is a key modification which coordinates factors in ICL repair. Unlike all other endonucleases implicated in ICL unhooking, SNM1A is not known to interact with monoubiquitinated FANCI-FANCD2 (mediated by the FA complex) or SLX4, which are main factors of replication-dependent ICL repair. However, the UBZ domain of SNM1A does interact with monoubiquitinated PCNA (Yang et al., 2010). Modification of PCNA is mediated by RAD18-RAD6, which is activated in response to accumulated RPA on a lagging strand (Huttner and Ulrich, 2008). It could therefore be possible that the RPA-coated ssDNA on the lagging strand triggers RAD18-RAD6 in response to a replication fork stalling by an ICL. PCNA, associated with leading strand synthesis machinery, is monoubiquitinated by RAD18-RAD6, providing a signal for recruitment of SNM1A though monoubiquitinated PCNA. The converged replication fork signals for FA complex assembly and recruitment of XPF-ERCC1 though SLX4 association with monoubiquitinated FANCD2-FANCI. With both XPF-ERCC1 and SNM1A present at the crosslink, unhooking could then be accomplished by incision of the leading and lagging

strand by XPF-ERCC1 and SNM1A, respectively. The unhooked ICL would be processed by SNM1A exonuclease and translesion nuclease activity in preparation for downstream repair by TLS and HR.
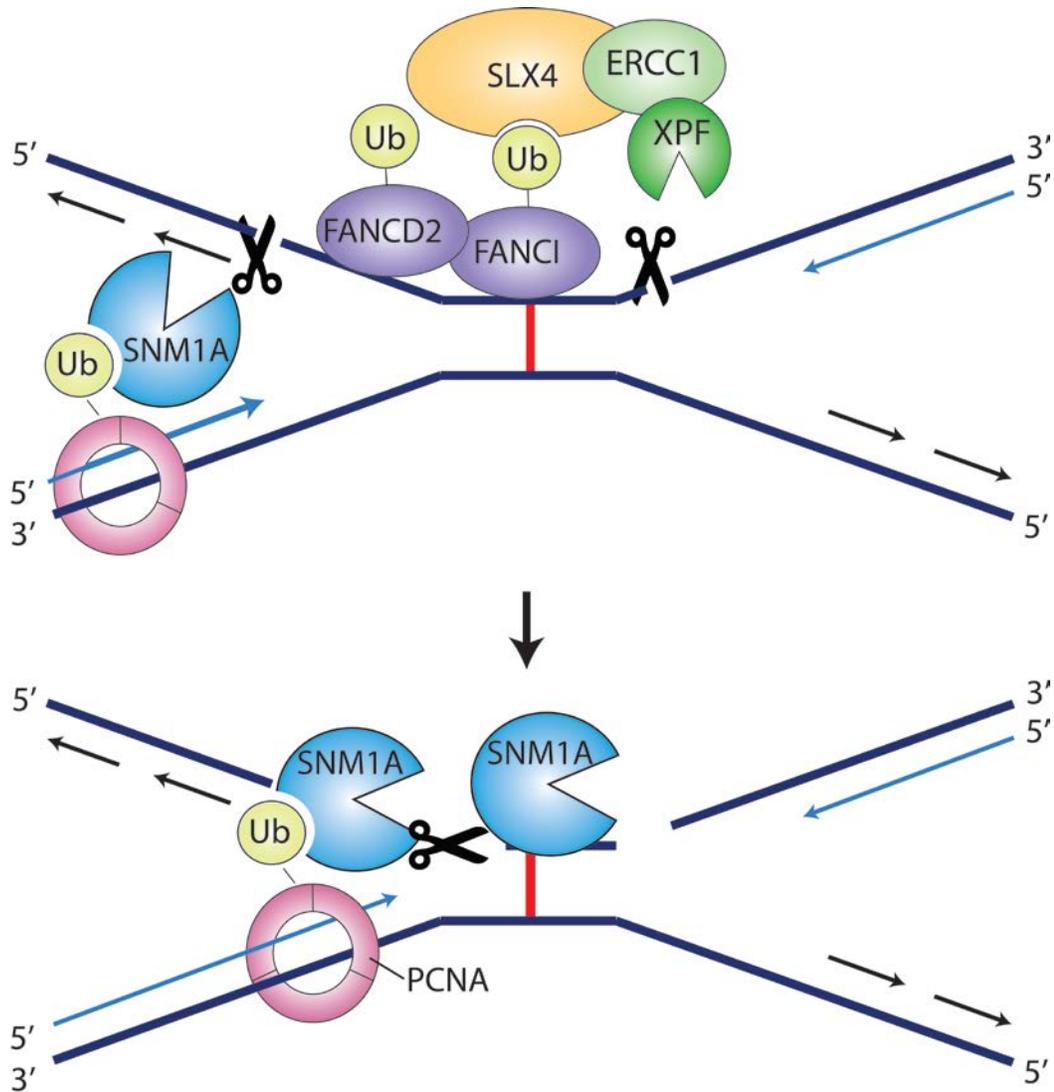
Figure 38: Possible functions of SNM1A in replication-dependent ICL repair

FA-mediated monoubiquitination of FANCD2-FANCI recruits 3'flap endonuclease ERCC1-XPF to cut 3' of the ICL lesion. RAD18-mediated monoubiquitinated PCNA recruits SNM1A to the converged replication fork. SNM1A endonuclease activity nicks unpaired DNA in the lagging strand or XPF-ERCC1 cleaves duplex 5' to the ICL, creating a 5'phosphate for SNM1A exonuclease activity trimming. SNM1A translesion nuclease activity could then bypass the ICL adduct, flipping out a single crosslinked base. A lesion tolerant polymerase may then replace SNM1A for TLS, repair being completed by HR.

4.3.2 Replication-Independent Repair

While replication-dependent repair is thought to be the main mechanism during which crosslinked DNA is restored, RIR is important for non-cycling or terminally differentiated cells. Understanding of transcription-coupled ICL repair is limited, but depletion studies suggest important roles for XPA, CSB and several TLS polymerases. A role for SNM1A in RIR is supported by interaction and epistasis studies of SNM1A and CSB (Iyama et al., 2015). CSB has been implicated in ICL unhooking, although this protein alone has no nuclease function (Iyama et al., 2015). CSB may stimulate SNM1A, accounting for the requirement of CSB in unhooking.

In support of a role of SNM1A in transcription-coupled repair, SNM1A directly interacts with PCNA, which coordinates TLS polymerases during repair in TC-NER. The enhanced hypersensitivity of SNM1A in response to ICL-inducing agents may also explain the epistasis of SNM1A specifically with XPF-ERCC1 as opposed to other structure-specific endonucleases that act on flaps and forks formed during replication. Given the evidence of SNM1A for RIR repair, there are two scenarios in which SNM1A may act in this pathway. First, SNM1A exonuclease activity may act downstream of XPF-ERCC1 after initial cleavage on the 5' side of a transcription bubble. Second, DNA remodelling by CSB at the damage site by may expose ssDNA and allow access of SNM1A to the ICL on the 3' side of the ICL. In this case, ICL trimming by SNM1A would be unnecessary. Regardless, SNM1A may be localized to the repair site by association with PCNA, CSB, and potentially RPA.
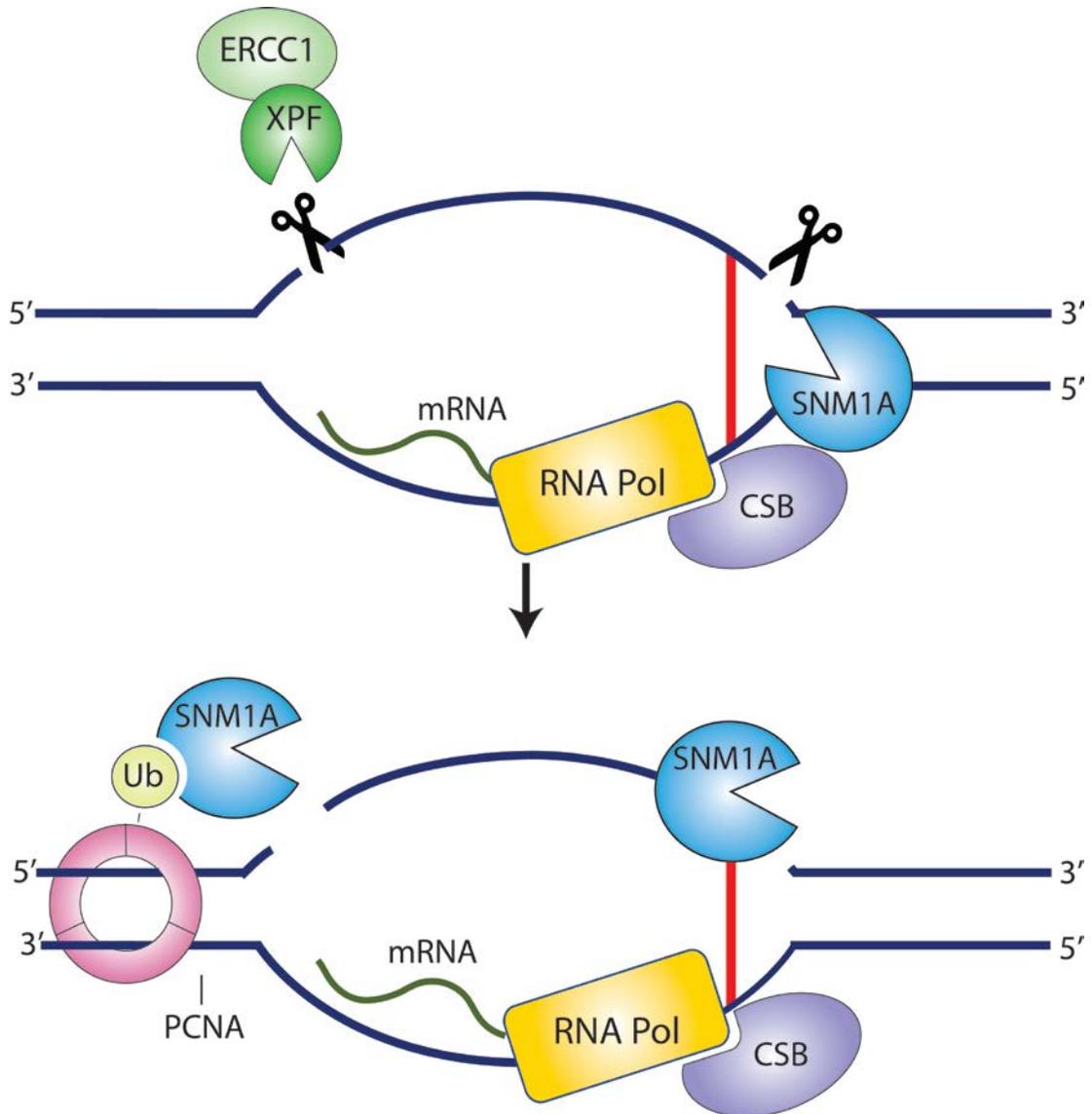
Figure 39: Possible functions of SNM1A in transcription-coupled ICL repair

RNA Pol II stalling recruits CSB for DNA remodeling to open DNA. SNM1A recruitment via CSB may enable endonuclease incision of ssDNA. XPF-ERCC1 incision 5' of ICL lesion could generate a 5' phosphate for exonuclease trimming. PCNA might recruit SNM1A for trimming and lesion bypass. A lesion tolerant polymerase may displace SNM1A from PCNA to perform TLS, and repair completion by a second round of NER.

4.4   Inhibition of ICL Repair

Inability to properly repair DNA can result in tumorigenesis, however anti-cancer therapies heavily rely on damaging DNA in malignant cells to promote apoptosis. As such, anti-cancer efforts have focused on understanding the mechanisms of DNA repair to exploit its inhibition in malignant cells. The benefit of targeted inhibition of DNA repair is twofold. First, persisting DNA damage by decreased repair may extend the cytotoxic effects of DNA damaging agents. However, initial tumorigenesis is often indicative of defects in DNA repair. Cell survival in the presence of genomic instability therefore relies on other functional repair pathways to evade apoptosis (Nickoloff et al., 2017). Thus, targeting these alternative repair pathways, on which a cancerous cell becomes dependent, removes the survival mechanisms of that malignant cell (Curtin, 2012). This is best exemplified by the successful development of PARP inhibitors as a chemotherapeutic drug.

PARP1 is an early factor required for single-strand DNA break (SSB) and DSB repair. PARP-dependent DSB repair occurs by error-prone alternative end-joining (Alt-EJ) and becomes important for cell survival once HR repair is no longer functional (Wang et al., 2006). Mutations in both copies of either HR factor BRCA1 or BRCA2 are embryonic lethal, but a single mutated copy can be tolerated (Hakem et al., 1998). Mutation of the second copy of BRCA1/2 impairs HR and genomic instability results in tumorigenesis. Of heritable ovarian and breast cancers, 90% harbour BRCA mutations that quickly become chemoresistant (Paluch-Shimon et al., 2016). PARP inhibition not

only prevents the repair of SSBs which converted into lethal DSBs during replication, but also blocks Alt-EJ required for replication fork restart, thereby eradicating tumor cells (Dziadkowiec et al., 2016). PARP inhibitors thus exemplify the ability to target and inhibit DNA repair to combat cancers and demonstrate a rational approach towards chemotherapeutic drug development.

4.4.1 Inhibitors of XPF-ERCC1 to date

Similar to strategies used for PARP inhibition, a rational approach towards inhibition of XPF-ERCC1 by small molecules has also been attempted. Not only do clinical studies demonstrate that XPF-ERCC1 is a predictor of tumor response to chemotherapy in many cancers, but mechanistic studies also support the essential role of XPF-ERCC1 in ICL unhooking (McNeil and Melton, 2012). Thus, several HTS campaigns directed towards inhibition of XPF-ERCC1 have been conducted. *In silico* screens targeting *de novo* complex formation of XPF-ERCC1, as well as *in vitro* screens focusing on disruption of the XPF and ERCC1 interface, identified small molecules that inhibited heterodimerization of XPF-ERCC1 (McNeil et al., 2015). These inhibitors proved ineffective in cells dues to the stability of XPF-ERCC1 interaction.

Further screening of small molecules targeting XPF endonuclease activity identified two promising compounds (Arora et al., 2014). One of the XPF inhibitors targeted the conserved metal binding site and was therefore non-specific, while the second compound demonstrated high toxicity. Ultimately, although these compounds

were effective in sensitizing cells to DNA damage, the effects were not ICL-repair specific.

4.4.2 Current Status of SNM1A Inhibitors

Directed inhibition of XPF-ERCC1 in ICL repair remains problematic, since XPF-ERCC1 is required for both ICL and NER repair. Given the epistasis of XPF-ERCC1 and SNM1A in response to ICL-inducing agents, inhibition of SNM1A may enable more specific inactivation of ICL repair compared to XPF-ERCC1. Recognizing this, efforts to find SNM1A inhibitors have gained significant attention. The metallo-β-lactamase fold encompassed within the MBL/β-CASP domain hydrolyzes an array of substrates, including β-lactams. Screening a panel of β-lactam antibiotics for SNM1A inhibition resulted in four hits, each containing a cephalosporin core (Lee et al., 2016). While researchers could demonstrate SNM1A inhibitory effects of cephalosporins in the micromolar range *in vitro*, *in vivo* effects were unsuccessful since these β-lactamase inhibitors function at the extracellular level, with low cell membrane permeability (Lee et al., 2016). Given this, we chose to carry out HTS using only bioactive compounds for inhibition of SNM1A to facilitate downstream testing in cells. Using this strategy, we were able to successfully identify inhibitors that not only inhibit SNM1A *in vitro*, but also exhibit synergism with cisplatin in cell-based systems.

4.5 Limitations of SNM1A Inhibitors

Although identification of bioactive SNM1A inhibitors is significant, there are caveats to SNM1A inhibition involving specificity. Inhibition must target the correct

nuclease and correct pathway to prevent off-target effects.

4.5.1 Off-Target Inhibition of β-CASP Nucleases

Given the essential and specific role of SNM1A in ICL repair, SNM1A inhibition represents a preferable target compared to XPF-ERCC1. However, given conservation of the catalytic core, inhibition of other β-CASP nucleases may limit the efficacy of SNM1A inhibitors. Mammalian cells have three DNA β-CASP processing paralogs and one RNA β-CASP homolog. Inhibition of these β-CASP nucleases could disrupt many important cell functions, including telomere maintenance, NHEJ, V(D)J recombination, and RNA processing (Yan et al., 2010). The active site of β-CASP proteins is well-conserved, with five motifs coordinating and participating in catalysis. Given the conservation of the metal binding site of β-CASP nucleases, active site inhibitors must be carefully designed to specifically target SNM1A. In this regard, structure determination of SNM1A bound to inhibitor would be helpful to better understand specificity and mechanism of action. Parallel structure-function analysis of SNM1A mutants and DNA-bound structure determination will be important to help elucidate how the inhibitors disrupt activity and guide further inhibitor modification to improve target specificity. Interestingly, HTS campaigns directed at inhibition of SNM1C hairpin-opening activity have been performed. Of the nine SNM1A inhibitors identified in our studies, none have been identified as inhibitors of SNM1C. Testing compounds that inhibit SNM1C for SNM1A inhibition may help understand co-inhibition of β-CASP nucleases in general.

4.5.2 Off-Target Effects of SNM1A Inhibition

While inhibition of SNM1A is expected to shut down ICL repair more specifically than other factors involved in ICL repair, SNM1A inhibition may disrupt less defined cellular roles of SNM1A. It is unclear how SNM1A is involved DNA damage response to IR since SNM1A is required for proper foci formation, but hypersensitivity to IR is absent in SNM1A depleted cells (Akhter and Legerski, 2008b). It is possible that a crosslink-like substrate is formed during ionization of DNA and redundancy with SNM1C masks IR hypersensitivity. Further studies in foci formation with catalytically inactive SNM1A and/or co-depletion of wild-type SNM1A and SNM1C may help elucidate the role of SNM1A in response to IR.

Redundancy with SNM1B and SNM1C have been observed for induction of apoptosis induced by etoposide. DNA fragmentation was shown to be diminished in *SNM1A*$^{-/-}$ cells, indicating that catalytically active SNM1A is important for apoptotic DNA fragmentation (Hosono et al., 2011). However, SNM1A inhibition is unlikely to have a significant effect on this process since no additive effect was observed with depletion of SNM1C and SNM1B, which are also involved in etoposide-induced apoptosis.

SNM1A has also been implicated in cell cycle regulation in G1 and mitosis. While arrest of the G1 checkpoint is likely required to ensure efficient DNA repair before entry into replication, the role of SNM1A in the early mitotic checkpoint is unknown (Akhter and Legerski, 2008b; Akhter et al., 2004a). Depletion of SNM1A results in aneuploidy and micronuclei, indicative of chromosomal instability, not DNA breaks (Akhter et al., 2004b). DNA repair proteins have been shown to be required for proper chromosome

segregation, but it is unclear if the chromosome undergoes active repair (Hustedt and Durocher, 2016). Determining if and how the nuclease activity of SNM1A contributes to chromosomal stability during mitosis is an important consideration for efforts to pursue an SNM1A inhibitor.

4.5.3 Off-Target Effects from Small Molecule Inhibitors

Given that a library of bioactive compounds was screened for SNM1A inhibitors, it is reasonable that SNM1A inhibitors likely have other biological effects. The targets of theaflavin digallate, amentoflavone, 6-OH DOPA, and Reactive Blue 2 are noteworthy since these four compounds demonstrated high synergism with cisplatin.

Flavonoids encompass a family of ubiquitous polyphenolic secondary metabolites of plants with antioxidant activity. Theaflavin digallate, as well as SNM1A inhibitor epigallocatechin, are flavonoids derived from tea that demonstrate chemoprotective effects in *in vivo* models when delivered before and during exposure to carcinogens (Kanwar et al., 2012). *In vitro*, theaflavin digallate can bind to DNA structures (including dsDNA, quadraplex DNA, and DNA histones) (Mikutis et al., 2013). Non-specific DNA binding may explain the observed inhibition of all molecular biology nucleases tested. This compound has also been reported to inhibit a variety of kinases (including JNK, Erk, and ABL kinase) as well as NF-κB in macrophages (Sur and Panda, 2017). Toxicity was noted for cells exposed to theaflavin digallate alone. This toxicity may be a result of inhibition of kinases and potentially proteins interacting with DNA. Further work would be required to increase the specificity theaflavin digallate for SNM1A.

6-OH DOPA (not to be confused L-DOPA used in the treatment of Parkinson's disease) is a potent excitotoxin. 6-OH DOPA decreases noradrenaline concentrations, resulting in the degeneration of neurons in the noradrenaline nerve terminals (Sachs and Jonsson, 1972). High doses administered to rodents result in Parkinson's symptoms, giving researchers the ability to probe mechanisms of this disease (Olney et al., 1990). Potential neurotoxicity of 6-OH DOPA may be an important off-target effect that may limit drug efficacy when used for cisplatin sensitization, thus requiring chemical modifications to increase specificity for SNM1A.

Reactive Blue 2 is a synthetic anthraquinone and dye within the same family as Brilliant Blue G, trypan blue, and Cibacron Blue 3GA (Ferreira et al., 2016). Reactive Blue 2 is a selective agonist of some PY2 purinergic G-protein coupled receptors (Reilly et al., 1987). For unknown reasons, this compound has been discontinued for industrial and research purposes.

Interestingly, both Reactive Blue 2 and 6-OH DOPA were identified as APE1 nuclease inhibitors (Simeonov et al., 2009). APE1 is required to nick the DNA backbone 5' of an abasic site during BER (Freudenthal et al., 2015). While both SNM1A and APE1 are both metal-dependent nucleases, the basis of cross-inhibition of SNM1A and APE1 is unclear as both proteins have distinct active sites and mechanisms of incision. APE1 produces a 5'deoxyribose 5'phosphate and 3'OH using a single Mg^{2+} to coordinate the phosphate backbone (Tsutakawa et al., 2013). Interestingly, the active site of Nfo (the APE homolog in bacteria) contains at least two Zn^{2+} ions that may be similar to those found in SNM1A (Tsutakawa et al., 2013). It is possible that the sulfate groups of

Reactive Blue 2 mimic the phosphate backbone of of DNA or possibly chelate the divalent metal ion(s) within the active site. However, Reactive Blue 2 only moderately inhibited other nucleases, suggesting a more specific mechanism. Further structural work comparing inhibitor binding to SNM1A and APE1 is required to help understand the basis of cross-inhibition of these proteins.

As previously described in the introduction, persistent abasic sites represent reactive lesions, leading to formation of ICLs (Greenberg, 2014). Thus, increasing the number of ICLs and reducing the extent of repair may explain the significant synergy observed for SNM1A and APE1 inhibitors. APE1 upregulation has been implicated in cisplatin resistance and may be a prediction marker for cancer prognosis with cisplatin therapy (Al-Attar et al., 2010). As such, the link between BER and ICL repair needs to be further understood and remains an important area of investigation.

4.5.4 Functional Redundancy of SNM1A

Efficacy of SNM1A inhibition may be limited by functional redundancy of other nucleases. Given the central role of unhooking in ICL repair, it is not surprising that redundancy might exist among endonucleases to ensure that this critical step proceeds. For example, both XPF and MUS81 share the same active site and function as 3' flap endonucleases (Ciccia et al., 2008). If XPF-ERCC1 fails to generate the first nick to unhook an ICL, MUS81-EME1 may be able to act on the same substrate in attempts to restart the collapsed replication fork (Wang et al., 2011). Likewise, nucleases that act on a

5' flap or bypass an ICL lesion could substitute for SNM1A endonuclease cleavage or SNM1A translesion nuclease activity, respectively.

The best candidate is FAN1 that is a functionally redundant protein to SNM1A. FAN1 has been shown to act as a 5'flap endonuclease (Yoshikiyo et al., 2010). Although other 5'flap endonucleases have been implicated in ICL repair (SLX1-SLX4), SNM1A and FAN1 share other unique nuclease functions absent in other nucleases. Both SNM1A and FAN1 possess 5' phosphate dependent 5'-3'exonuclease activity (Li et al., 2005; Yoshikiyo et al., 2010). This strict requirement may reflect a poorly understood requirement for repair of ICLs. Furthermore, the exonuclease activity of FAN1, which cuts every fourth nucleotide, can bypass an ICL (Pizzolato et al., 2015; Wang et al., 2014). Thus, SNM1A and FAN share similar translesion nuclease activity. As an endonuclease, FAN1 cleaves a 5'flap within the duplex region (Zhao et al., 2014). The structure specificity of SNM1A appears to be more relaxed with ability to cut single-stranded DNA alone. Although ICL unhooking by FAN1 has been shown to occur *in vitro*, an unhooking defect was not observed in Xenopus extracts depleted of FAN1 (Klein Douwel et al., 2014). Unhooking by SNM1A was not tested in these studies since it was not known to have endonuclease activity. Testing the effect of SNM1A depletion on unhooking, as well as possible redundancy with FAN1 may reveal unhooking defects not previously observed. Indeed, the mild hypersensitivity of $SNM1A^{-/-}$ MEF cells in response to ICL-inducing agents is exacerbated by FAN1 knockdown, suggesting that SNM1A hypersensitivity may be partially masked by FAN1 function (Thongthip et al., 2016). ICL hypersensitivity in cells depleted for both SNM1A and FAN1 approached hypersensitivity

of SLX4 depletion (Thongthip). This finding suggests that two sub-pathways for ICL repair may exist: one using SNM1A and one using FAN1. Disruption of both pathways would effectively halt all ICL repair. In *S. pombe*, *fan1Δpso2Δ* cell show significant decreases in survival in response to ICL exposure, but only mild hypersensitivity when knocked-out individually (Fontebasso et al., 2013). Notably, Fan1 is absent in *S. cerevisiae*, suggesting that the unique role of Pso2 in budding yeast may be due to a singular protein able to unhook, trim, and bypass ICLs. Finally, both SNM1A and FAN1 are implicated in mitotic checkpoint regulation. SNM1A and APC are associated to regulate the prophase-metaphase transition, whereas FAN1 is a substrate of APC mediating mitotic exit (Akhter et al., 2004b; Lai et al., 2012). Thus, although the mechanisms underlying nuclease activities of SNM1A and FAN1 differ, functional redundancy between these two proteins may help to ensure efficient 5' cleavage and translesion nuclease activity is able to occur during ICL repair.

CHAPTER FIVE: CONCLUSIONS AND FUTURE DIRECTIONS

The repair of ICLs is a complex process, requiring highly coordinated nucleases at multiple stages during repair. SNM1A is a key nuclease in human ICL repair, but its precise function in this repair process is unclear. The aims of this thesis were twofold: to investigate the nuclease activities of SNM1A and to inhibit its activity using small molecules.

5.1   Summary of SNM1A Nuclease Activities

We demonstrate that, like other members of the β-CASP family of nucleases, SNM1A has structure-specific endonuclease activity. This endonuclease activity is single-strand specific, with a preference for pyrimidine nucleotides. Co-purification of SNM1A exonuclease and endonuclease activity, as well as ability of SNM1A to complement *Δpso2* hairpin-opening defects in yeast, confirm that SNM1A possesses endonuclease activity.

We also show that SNM1A can bypass highly distorting ICLs, including cisplatin and psoralen, but direct incision of the ICL does not appear to be the basis of translesion nuclease activity. We were unable to validate SNM1A translesion nuclease activity on SJG-136 ICLs and further in-depth analysis suggested that ICL crosslink stability may be responsible for the discrepancy between our findings and previous reports. Nevertheless, the findings that SNM1A not only has 5'-3' exonuclease activity, but also single-strand

specific endonuclease activity and translesion nuclease activity on distorting ICLs broaden the potential role(s) of SNM1A in ICL repair.

The three distinct nuclease activities of SNM1A reflect a non-specific highly flexible nuclease. These activities suggest that SNM1A may be sufficient to unhook an ICL adduct and generate a 5' phosphate for further trimming up to and past an ICL. Intriguingly, of the nucleases implicated in human ICL repair, only for SNM1A has ssDNA been shown to be sufficient for endonuclease cleavage. This contrasts with all other structure-specific endonucleases involved in ICL that cleave the double-to-single strand junction of DNA flaps within the duplex region. Uncontrolled single-strand cleavage during replication could detrimentally generate a DSB, therefore it is unsurprising that nuclease activity of SNM1A must be strictly regulated. Furthermore, comparatively weak endonuclease activity of SNM1A may actually prevent inadvertent incision in the absence factors required to licence DNA cleavage. Further work exploring remarkable plasticity of the SNM1A activity site will shed light on the molecular mechanisms of SNM1A.

5.2   Summary of SNM1A Inhibition

Using a fluorescence-based exonuclease assay optimized for SNM1A activity, we screened ~4,000 compounds for inhibition of SNM1A. Of the 53 compounds initially identified as possible hits, 22 showed dose-response inhibition of SNM1A. An orthogonal gel-based assay further confirmed 9 small molecules as inhibitors of SNM1A, with an IC_{50} in the mid-nanomolar to mid-micromolar range. We determined the mode of action

and analyzed each inhibitor for non-specific inhibition of other nucleases. Finally, while inherent toxicity of the drug alone was observed for two of the SNM1A inhibitors, we successfully identified compounds that act synergistically with cisplatin. Further analysis and optimization of these SNM1A inhibitors provide excellent opportunities for specifically inhibiting repair of ICLs.

5.3 Future Directions

An obvious question arising from our work characterizing SNM1A activity is which of the three nuclease activities of SNM1A are required for ICL repair? In order to address this question, separation of function mutants will be required to investigate the contributions of each activity within the cell. The ability to generate such mutants, however, is complicated by the fact that all nuclease activities are dependent on the same active site. Further work towards understanding the mechanics of each nuclease activity of SNM1A will be needed to create division of function mutants in order to determine the role(s) of SNM1A in ICL repair. Structural determination of SNM1A in complex with DNA will be required to help guide structure-function analysis to differentiate the distinct requirements for substrate binding. Alternatively, guided mutagenesis based on homology modeling and sequence conservation may help identify important elements of SNM1A nuclease activity.

Structure determination via x-ray crystallography should be pursued as follow-up studies of SNM1A inhibitors. Mode of action was determined for only half of the

SNM1A inhibitors, which distinctly affected either K_M or V_{max}. However, other SNM1A inhibitors displayed mixed inhibition, influencing both kinetic parameters. Inhibitor-bound structures of SNM1A are needed to help elucidate the precise mechanism by which inhibitor act, particularly in cases with differential inhibition of exonuclease and endonuclease activity.

Both separation of function and inhibitor studies must be accompanied by cell-based work to establish the role of SNM1A in human ICL repair. Although SNM1A knockouts have been generated in mice and DT40 chicken cells, these cells may not recapitulate the role of SNM1A in human cells. It is thus paramount that human cells lacking SNM1A be studied. Most work in SNM1A depletion studies within human cells have used siRNA to knockdown SNM1A levels, but a stable knockout cell line is necessary to minimize variability for proper comparative studies. We have attempted to fully knockout SNM1A in HeLa cells using Crispr-Cas9, but have encountered difficulty obtaining isolates. Difficulty with SNM1A depletion has also been encountered by other groups, where aneuploidy and growth defects have been observed. Nonetheless, a cell line without SNM1A (perhaps conditionally knocked out) is paramount to ascertain of the effect of cisplatin synergy is a direct result of SNM1A inhibition in the cell.

The ability to eliminate wild-type SNM1A will also facilitate division of function analysis by complementation with SNM1A mutants. While current studies support a role of ICL unhooking by SNM1A, unhooking by SNM1A has yet to be definitively demonstrated. Work towards analysis of crosslinked genomic DNA from yeast using 2D

gel electrophoresis had been undertaken, but did not yield reproducible results. Modified comet tail assays, in which the length of the tail of broken DNA is reduced in the presence of a crosslink, may better provide the capacity to monitor and measure unhooking. Using modified comet tail assays, CSB has been shown to promote unhooking. Since CSB helps recruit SNM1A to sites of damage, determination of if and how SNM1A unhooks DNA is required.

5.4   Conclusions and Relevance

Studies on SNM1A date back more than twenty five years ago, with emphasis investigating the role of SNM1A in DNA repair (Ahkter et al., 2005; Dronkert et al., 2000b; Ishiai et al., 2004). Because of the mild ICL hypersensitivity in mice and non-epistatic relationship with the FA pathway, the importance of SNM1A in ICL repair has been largely overlooked (Dronkert et al., 2000b; Hemphill et al., 2008b). However, the FA pathway represents a mechanism to repair ICLs in only one of four stages of the cell cycle. Although ICLs are most disruptive to DNA replication, it does not mean that ICL repair is not critical during other phases of the cell cycle. RIR is particularly crucial in non-cycling or terminally differentiated cells (e.g. stem-cells, neurons, cardiac myocytes) (Williams et al., 2012). Significant accumulation of damage within the stem cell compartment has substantial consequences on downstream differentiated cells, as evident in blood disorders of FA syndrome (Trujillo et al., 2011; Rosenberg et al., 2003). Damage to mainly quiescent stem cells may result in genomic instability, tumorigenesis, and persistent tumors (Singh et al., 2003). Neuropathology and neurogeneration are well

established with mutations in NER and are emerging for ICL repair (Chen et al., 2016; Madabhushi et al., 2014; Wang et al., 2014). Taken together, it would appear that in the absence of replication, the capacity to repair ICLs is important and must not be overlooked.

Mild hypersensitivity of SNM1A may differ depending on the complement of expressed proteins within a given cell. Indeed, ICL repair capacity differs amongst cell types, as evident by tissue-dependent variations in XPF-ERCC1 expression levels and innate chemoresistance (Kirschner and Melton, 2010). Furthermore, differences in ICL repair may be organism-specific. Absence of genes, such FAN1 in budding yeast or MUS81 in DT40 cells, or presence of some FANC proteins in only higher eukaryotes reflect altered genetic interactions, changing epistatic relationships of ICL factors (Fontebasso et al., 2013; Kikuchi et al., 2013; McHugh et al., 2012). Unlike dedicated pathways for repair of DNA breaks (e.g. HR, NHEJ, SSB) or erroneous DNA bases (e.g. NER, TLS, MMR), ICL repair makes use of multiple pathways. This suggests a selective pressure to repair ICLs that has not evolved from a single pathway for crosslink repair. Damage induced by ICLs is often derived from external sources and not within the cell. Exposure to natural forms of crosslinking agents vary, depending diet and/or exposure to microorganisms. It has only been century since the creation of the first synthetically derived ICL-inducing compounds. In all, exposure to ICL-inducing agents not only varies, but continues to evolve.

Variation in responses to ICL exposure should not hinder the work towards inhibition of ICL repair as a means of tumor sensitization. In fact, these differences could

be exploited to target specific cancers. Similar to BRCA1/2 haplo-insufficiency as an indication for treatment using PARP inhibitors, genetic markers of ICL repair may eventually guide the use of ICL compounds for individual chemotherapy regimens. Increased ERCC1 expression within a tumor suggest that malignant cells proficient in DNA repair may better respond to ICL-based chemotherapy when coupled with SNM1A inhibition, given the relationship between SNM1A and XPF-ERCC1

Some have argued that inhibitors of SNM1A are unlikely to be developed as clinically useful drugs. Although SNM1A depletion results in significant hypersensitivity *ex vivo* in human cells, research to date suggests that SNM1A hypersensitivity is mild in *in vivo* mouse models. However, there is currently not enough known to understand, let alone conclude, about the importance of SNM1A in human ICL repair. Specific inhibition of SNM1A in the cell would greatly facilitate the dissection of ICL repair by eliminating the contribution of SNM1A without genetic manipulation.

Compounds that inhibit ICL repair act in opposition to mechanisms which drive chemoresistance, thus potentiating the cytotoxicity of ICL-inducing agents. It may be possible to achieve the therapeutic effect of cisplatin without cisplatin-associated toxicities with co-administration with an ICL repair inhibitor. Cisplatin sensitization would not only reduce the side-effects of chemotherapy, but could also fully eliminate malignant cells of the primary tumor. Until we comprehend the intricacies of crosslink repair, however, can we attempt to reverse the process for our benefit in hopes to ultimately combat chemoresistance and cancer.

APPENDIX 1: Oligonucleotide sequences

Substrate	Oligo	Sequence	Bases	Modification
Gap	gap.T	G*AGGGCGAGCCCGATTTTTTCCGCTTGACCC-AAGTAAGATTTTTTGCAGATACTTAAC*A*C	60	5'6FAM
	gap.B1	A*GCGGAAAAAATCGGGCTCGCCC*T*C	25	--
	gap.B2	G*TGTTAAGTATCTGCAAAAAATC *T*T	25	--
5' overhang	5'overhang.T	T*TTTTTTTTTTTTTTTACTGAGTCCTACAGAAGG*A*T	35	5'6FAM
	overhang.B	G*AT CCT TCT GTA GGA CTC A*G*T	21	--
5' pseudo Y	5'overhang.T	T*TTTTTTTTTTTTTTTACTGAGTCCTACAGAAGG*A*T	35	5'6FAM
	5'O/H.polyT.B	G*ATCCTTCTGTAGGACTCAGTTTTTTTTTTTTTTTT*T*T	36	
5' flap	5'overhang.T	T*TTTTTTTTTTTTTTTACTGAGTCCTACAGAAGG*A*T	35	5'6FAM
	5'O/H.B	G*ATCCTTCTGTAGGACTCAGTCATATTTATAACA*C*C	36	--
	SJG 5'flap.B3	G*GTGTTATAAATA*T*G	15	--
3' overhang	3'overhang.T	A*CTGAGTCCTACAGAAGGATCTTTTTTTTTTTTT*T*T	36	5'6FAM
	overhang.B	G*ATCCTTCTGTAGGACTCA*G*T	21	--
3' pseudoY	3'overhang.T	A*CTGAGTCCTACAGAAGGATCTTTTTTTTTTTTT*T*T	36	5'6FAM
	3'O/H.polyT.B	T*TTTTTTTTTTTTTTTGATCCTTCTGTAGGACTCA*G*T	36	--
3' flap	3'overhang.T	A*CTGAGTCCTACAGAAGGATCTTTTTTTTTTTTT*T*T	36	5'6FAM
	3'O/H.B	C*ATATTTATAACACCGATCCTTCTGTAGGACTCA*G*T	36	--
	SJG 5'flap.B3	G*GTGTTATAAATA*T*G	15	--
Symmetrical bubble	bubble.T	A*GGCTGTGTTAAGTATCTGGTTTTTTTTTTT-GCTCGCCCTCAGGTCGAC*A*A	50	5'6FAM
	bubble.B	T*TGTCGACCTGAGGGCGAGCCCGATGAATTCCAGATAC TTAACACAGC*C*T	50	--
Heterologous loop	bubble.T	A*GGCTGTGTTAAGTATCTGG-TTTTTTTTTTGCTCGCCCTCAGGTCGAC*A*A	50	5'6FAM
	loop.B	T*TGTCGACCTGAGGGCGAGCCCAGA TA CTTAACACAGC*C*T	40	--
Stem loop	stemloop.T	GCTGACTGAGTCCTACAGAAGGATCTTTTTTTTTTGATCCTT CTGTAGGACTCAGTCAG	60	5'6FAM
Paired hairpin	hairpin20	G*ATTACTACGGTAGAGCTACGTAGCTCTACCGTAGTA*A*T	35	5'6FAM
Y-structure	5'Y.T	G*TACGTAACCTGACTGCTATCGACTGGACTTGATGCCG*T*C	40	5'6FAM
	Y.B1	G*ACGGCATCAAGTCCAGTCGCAGATGGCGTATAGCAGT*T*A	40	--
	Y.B2	T*AACTGCTATAC-GCCATCTGATAGCAGTCAGGTTACGT*A*C	41	--
5'P	5'P	TTTTTTTTTTTTTTTTTTTT	20	3'6FAM, 5'Phos
5'OH	5'OH	TTTTTTTTTTTTTTTTTTTT	20	3'6FAM
5'F (PolyT)	PolyT	T*TTTTTTTTTTTTTTTTTT*T*T	23	5'6FAM
PolyA	PolyA	A*AAAAAAAAAAAAAAAAA*A*A	23	5'6FAM
PolyAC	PolyAC	A*AACCCAAACCCAAACCC*A*A	20	5'6FAM
SJG ICL 5' P	SJG(AT).5O/H.5'P.3F.T	TTTTTTTTTTTTTTTTTATAATTTTAATTTGATC-ATTTATTATAAATTTTATTAT*A*T	56	3'6FAM, 5'Phos
SJG ICL 5'Y OH	SJG(AT).5O/H.3F.T	T*T*TTTTTTTTTTTTTTTATAATTTTAATTTGATCATTTATTATAAATTTTATTAT*A*T	56	3'6FAM
SJG ICL 5'Y F	SJG(AT).5O/H.5F.T	T*T*TTTTTTTTTTTTTTTATAATTTTAATTTGATCATTTATTATAAATTTTATTAT*A*T	56	5'6FAM
	SJG(AT:5'O/H).B	A*TATAATAAATTTATAATAAATGATCAAATTAAAATTATCATATTTAT AACA*C*C	56	--
	SJG 5'flap.B3	G*GTGTTATAAATA*T*G	15	--
Short ICL	pSJG-T-Cy3	ATA ATT TGATCA TCT ATTA*T	20	3'CY3, 5'Phos,
	SJG-B-Cy3	A*T*AATAGATGATCAAATTAT	20	--
Long ICL	SJG(AT).T	ATA ATT TG ATC ATT TAT TAT AAA TTT TAT TAT AT	33	3'CY3, 5'Phos,
	SJG(AT).B	A*TATAATAAATTTATAATAAATGATCAAATT*A*T	33	--
GC1	SJG long.T R2	TGTCCTTTTAGATCTGTGTTGTATTTGGT*G	30	3'6FAM, 5'Phos
	SJG long.B R2	C*ACCAAATACAACACAGATCTAAAAGGA*C*A	30	--
GC2	SJG long.T R1	TGTGGTTTTAGAAATGTGTTGTATTTGGT*G	30	3'6FAM, 5'Phos
	SJG long.B R1	C*ACCAAATACAACACAGATCTAAAA*C*C	27	--
Cisplatin ICL	CP30.T	T*C*TCCTTCTGCTCTCTTCCTTCCTCTTT*C*C	30	3'6FAM+5'Phos or 5'6FAM
	CP30.B	G*GAAAGAGGAAGGAAGAGAGCAGAAGGA*G*A	30	--
Psoralen ICL	PSO(GC).T	C*G*GCCATCGCGCGGCGG*G*G	20	3'6FAM+5'Phos or 5'6FAM
	PSO(GC).B	C*CCCGCCGCGCGATGGC*C*G	20	--

APPENDIX 2: Primary HTS hits

Compound	IQM average
HARMALOL	0.14
FLUORESCEIN	0.14
MERBROMIN	0.16
#64	0.17
Tyrphostin 51	0.17
HARMALOL HYDROCHLORIDE	0.17
4,4'-DIISOTHIOCYANOSTILBENE-2,2'-SUFONIC ACID SODIUM SALT	0.17
OXOKAHWEOL, 16-	0.18
HAEMATOPORPHYRIN	0.18
HARMALOL HYDROCHLORIDE	0.18
Reactive Blue 2	0.19
HEMATEIN	0.19
EPIGALLOCATECHIN	0.2
TETRAHYDROPAPAVERINE	0.2
beta-SITOSTEROL	0.2
ACRIFLAVINIUM HYDROCHLORIDE	0.2
PATULIN	0.23
WEDELOLACTONE	0.27
DIHYDROTANSHINONE	0.27
GAMBOGIC ACID	0.27
THEAFLAVIN DIGALLATE	0.27
7-DESHYDROXYPYROGALLIN-4-CARBOXYLIC ACID	0.29
CHLORANIL	0.3
AC-93253 iodide	0.31
AMENTOFLAVONE	0.35
ACTINOMYCIN D	0.37
HOMIDIUM BROMIDE	0.37
METHYL 7-DESHYDROXYPYROGALLIN-4-CARBOXYLATE	0.37
GOSSYPOL	0.39
Calmidazolium chloride	0.41
Hispidin	0.42
Ruthenium red	0.42
3-OXOURSAN (28-13)OLIDE	0.42
PELLITORINE	0.43
6-Hydroxy-DL-DOPA	0.43

PHYTOL	0.43
ELLAGIC ACID	0.45
Me-3,4-dephostatin	0.45
RIBOFLAVIN	0.45
GERALDOL	0.47
DOXORUBICIN	0.48
KOPARIN	0.49
ALOE-EMODINE	0.5
NF 023	0.5
BU99006	0.51
GW5074	0.51
PPNDS tetrasodium	0.52
ERYTHROSINE SODIUM	0.52
SULFURETINE	0.53
DAUNORUBICIN	0.53
Chlorhexidine	0.53
Tyrphostin AG 879	0.55
DOXORUBICIN	0.55
MARITIMEIN	0.56
Idarubicin	0.56
Suramin sodium salt	0.56
ALEXIDINE HYDROCHLORIDE	0.56
BENZETHONIUM CHLORIDE	0.56
SHIKONIN	0.58
BRAZILIN	0.58
3,4'-DIHYDROXYFLAVONE	0.58
2,3-DICHLORO-5,8-DIHYDROXYNAPTHOQUINONE	0.58
STREPTONIGRIN	0.59

REFERENCES

Abdullah, U.B., McGouran, J.F., Brolih, S., Ptchelkine, D., El-Sagheer, A.H., Brown, T., and McHugh, P.J. (2017). RPA activates the XPF-ERCC1 endonuclease to initiate processing of DNA interstrand crosslinks TL - 36. EMBO J. *36 VN-r*, 2047–2060.

Abe, T., Ishiai, M., Hosono, Y., Yoshimura, A., Tada, S., Adachi, N., Koyama, H., Takata, M., Takeda, S., Enomoto, T., et al. (2008). KU70/80, DNA-PKcs, and Artemis are essential for the rapid induction of apoptosis after massive DSB formation. Cell. Signal. *20*, 1978–1985.

Admiraal, S.J., and O'Brien, P.J. (2015). Base excision repair enzymes protect abasic sites in duplex DNA from interstrand cross-links. TL - 54. Biochemistry *54 VN-r*, 1849–1857.

Ahkter, S., Richie, C.T., Zhang, N., Behringer, R.R., Zhu, C., and Legerski, R.J. (2005). Snm1-Deficient Mice Exhibit Accelerated Tumorigenesis and Susceptibility to Infection TL - 25. Mol. Cell. Biol. *25 VN-r*, 1007110078.

Ahmad, A., Robinson, A.R., Duensing, A., van Drunen, E., Beverloo, H.B., Weisberg, D.B., Hasty, P., Hoeijmakers, J.H.J., and Niedernhofer, L.J. (2008). ERCC1-XPF Endonuclease Facilitates DNA Double-Strand Break Repair. Mol. Cell. Biol. *28*, 5082–5092.

Aida, T., Takebayashi, Y., Shimizu, T., Okamura, C., Higasimoto, M., Kanzaki, A., Nakayama, K., Terada, K., Sugiyama, T., Miyazaki, K., et al. (2005). Expression of copper-transporting P-type adenosine triphosphatase (ATP7B) as a prognostic

factor in human endometrial carcinoma. TL - 97. Gynecol. Oncol. *97 VN-r*, 41–45.

Akhter, S., and Legerski, R.J. (2008). SNM1A acts downstream of ATM to promote the G1 cell cycle checkpoint. TL - 377. Biochem. Biophys. Res. Commun. *377 VN-*, 236–241.

Akhter, S., Richie, C.T., Deng, J.M., Brey, E., Zhang, X., Patrick, C., Behringer, R.R., and Legerski, R.J. (2004). Deficiency in SNM1 abolishes an early mitotic checkpoint induced by spindle stress. Mol. Cell. Biol. *24*, 10448–10455.

Allerston, C.K., Lee, S.Y., Newman, J.A., Schofield, C.J., McHugh, P.J., and Gileadi, O. (2015). The structures of the SNM1A and SNM1B/Apollo nuclease domains reveal a potential basis for their distinct DNA processing activities TL - 43. Nucleic Acids Res. *43 VN-r*, 11047–11060.

Bae, J.-B., Mukhopadhyay, S.S., Liu, L., Zhang, N., Tan, J., Akhter, S., Liu, X., Shen, X., Li, L., and Legerski, R.J. (2008). Snm1B/Apollo mediates replication fork collapse and S Phase checkpoint activation in response to DNA interstrand cross-links. Oncogene *27*, 5045–5056.

Benitez, A., Yuan, F., Nakajima, S., Wei, L., Qian, L., Myers, R., Hu, J.J., Lan, L., and Zhang, Y. (2014). Damage-dependent regulation of MUS81-EME1 by Fanconi anemia complementation group A protein. Nucleic Acids Res. *42*, 1671–1683.

Bennett, C.B., Lewis, A.L., Baldwin, K.K., and Resnick, M.A. (1993). Lethality induced by a single site-specific double-strand break in a dispensable yeast plasmid TL - 90. Proc. Natl. Acad. Sci. *90 VN-r*, 5613–5617.

Blood, A.-A.D. (1988). A test for Fanconi's anemia. Blood.

Bogliolo, M., Schuster, B., Stoepker, C., Derkunt, B., Su, Y., Raams, A., Trujillo, J.P., Minguillón, J., Ramírez, M.J., Pujol, R., et al. (2013). Mutations in ERCC4, Encoding the DNA-Repair Endonuclease XPF, Cause Fanconi Anemia. Am. J. Hum. Genet. *92*, 800–806.

Brendel, M., Bonatto, D., Strauss, M., Revers, L.F., Pungartnik, C., Saffi, J., and Henriques, J.A.P. (2003). Role of PSO genes in repair of DNA damage of Saccharomyces cerevisiae. Mutat. Res. - Rev. Mutat. Res. *544*, 179–193.

Brooks, P.J., and Theruvathu, J.A. (2005). DNA adducts from acetaldehyde: implications for alcohol-related carcinogenesis. TL - 35. Alcohol *35 VN-r*, 187–193.

Buckley, R.H., Schiff, R.I., Schiff, S.E., Markert, M.L., Williams, L.W., Harville, T.O., Roberts, J.L., and Puck, J.M. (1997). Human severe combined immunodeficiency: Genetic, phenotypic, and functional diversity in one hundred eight infants TL - 130. J. Pediatr. *130 VN-*, 378–387.

Budzowska, M., Graham, T.G., Sobeck, A., Waga, S., and Walter, J.C. (2015). Regulation of the Rev1-pol complex during bypass of a DNA interstrand cross-link. EMBO J. *34*, 1971–1985.

Callebaut, I., Moshous, D., Mornon, J.-P., and de Villartay, J.-P. (2002). Metallo-beta-lactamase fold within nucleic acids processing enzymes: the beta-CASP family. Nucleic Acids Res. *30*, 3592–3601.

Cannistra, S.A. (2004). Cancer of the ovary. TL - 351. N. Engl. J. Med. *351 VN-*, 2519–2529.

Carfi, A., Pares, S., Duée, E., Galleni, M., Duez, C., Frère, J.M., and Dideberg, O. (1995). The 3-D structure of a zinc metallo-beta-lactamase from Bacillus cereus reveals a new type of protein fold. TL - 14. EMBO J. *14 VN-r*, 4914–4921.

Castor, D., Nair, N., Déclais, A.-C., Lachaud, C., Toth, R., Macartney, T.J., Lilley, D.M.J., Arthur, J.S.C., and Rouse, J. (2013). Cooperative control of holliday junction resolution and DNA repair by the SLX1 and MUS81-EME1 nucleases. Mol. Cell *52*, 221–233.

Cattell, E., Sengerová, B., and McHugh, P.J. (2010). The SNM1/Pso2 family of ICL repair nucleases: from yeast to man. TL - 51. Environ. Mol. Mutagen. *51 VN-r*, 635–645.

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

Chen, J., Calhoun, V.D., Perrone-Bizzozero, N.I., Pearlson, G.D., Sui, J., Du, Y., and Liu, J. (2016). A pilot study on commonality and specificity of copy number variants in schizophrenia and bipolar disorder. Transl. Psychiatry *6*, e824–e824.

Ciccia, A., McDonald, N., and West, S.C. (2008). Structural and Functional Relationships of the XPF/MUS81 Family of Proteins. Annu. Rev. Biochem. *77*, 259–287.

Clauson, C., Scharer, O.D., and Niedernhofer, L. (2013). Advances in understanding the complex mechanisms of DNA interstrand cross-link repair. Cold Spring Harb Perspect Biol *5*, a012732.

Cybulski, K.E., and Howlett, N.G. (2011). FANCP/SLX4: a Swiss army knife of DNA

interstrand crosslink repair. TL - 10. Cell Cycle *10 VN-r*, 1757–1763.

Davies, A.A., Huttner, D., Daigaku, Y., Chen, S., and Ulrich, H.D. (2008). Activation of
  Ubiquitin-Dependent DNA Damage Bypass Is Mediated by Replication Protein A
  TL - 29. Mol. Cell *29 VN-r*, 625–636.

Deans, A.J., and West, S.C. (2011). DNA interstrand crosslink repair and cancer. Nat.
  Rev. Cancer *11*, 467–480.

Demuth, I., Digweed, M., and Concannon, P. (2004). Human SNM1B is required for
  normal cellular response to both DNA interstrand crosslink-inducing agents and
  ionizing radiation. Oncogene *23*, 8611–8618.

Demuth, I., Bradshaw, P.S., Lindner, A., Anders, M., Heinrich, S., Kallenbach, J.,
  Schmelz, K., Digweed, M., Meyn, S.M., and Concannon, P. (2008). Endogenous
  hSNM1B/Apollo interacts with TRF2 and stimulates ATM in response to ionizing
  radiation TL - 7. DNA Repair (Amst). *7 VN-re*, 1192–1201.

Dominski, Z. (2007). Nucleases of the metallo-beta-lactamase family and their role in
  DNA and RNA metabolism. Crit. Rev. Biochem. Mol. Biol. *42*, 67–93.

Doughty, M.J. (2010). PH dependent spectral properties of sodium fluorescein ophthalmic
  solutions revisited. Ophthalmic Physiol. Opt. *30*, 167–174.

Dronkert, M., and Kanaar, R. (2001). Repair of DNA interstrand cross-links TL - 486.
  Mutat. Res. Repair *486 VN-*, 217–247.

Dronkert, M.L., de Wit, J., Boeve, M., Vasconcelos, M.L., van Steeg, H., Tan, T.L.,
  Hoeijmakers, J.H., and Kanaar, R. (2000a). Disruption of mouse SNM1 causes
  increased sensitivity to the DNA interstrand cross-linking agent mitomycin C.

Mol. Cell. Biol. *20*, 4553–4561.

Eckstein, N. (2011). Platinum resistance in breast and ovarian cancer cell lines. J. Exp. Clin. Cancer Res. *30*, 91.

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

Fanconi, G. (1967). Familial constitutional panmyelocytopathy, Fanconi's anemia (F.A.). I. Clinical aspects. TL - 4. Semin. Hematol. *4 VN-re*, 233–240.

Fernandez, F.J., Lopez-estepa, M., and Vega, M.C. (2011). Nucleases of Metallo-B-Lactamase and Protein Phosphatase Families in DNA Repair. Intech.

Folmer, V., Soares, J.C.C., Gabriel, D., and Rocha, J.B.B. (2003). A high fat diet inhibits delta-aminolevulinate dehydratase and increases lipid peroxidation in mice (Mus musculus). TL - 133. J. Nutr. *133 VN-*, 2165–2170.

Fontebasso, Y., Etheridge, T.J., Oliver, a. W., Murray, J.M., and Carr, a. M. (2013). The conserved Fanconi anemia nuclease Fan1 and the SUMO E3 ligase Pli1 act in two novel Pso2-independent pathways of DNA interstrand crosslink repair in yeast. DNA Repair (Amst). *12*, 1011–1023.

Fousteri, M., and Mullenders, L.H.F. (2008). Transcription-coupled nucleotide excision repair in mammalian cells: molecular mechanisms and biological effects TL - 18. Cell Res. *18 VN-r*.

Fricke, W.M., and Brill, S.J. (2003). Slx1 - Slx4 is a second structure-specific endonuclease functionally redundant with Sgs1 - Top3. Genes Dev. *17*, 1768–1778.

Fuss, J.O., and Tainer, J.A. (2011). XPB and XPD helicases in TFIIH orchestrate DNA duplex opening and damage verification to coordinate repair with transcription and cell cycle via CAK kinase TL - 10. DNA Repair (Amst). *10 VN-r*, 697–713.

Galluzzi, L., Senovilla, L., Vitale, I., Michels, J., Martins, I., Kepp, O., Castedo, M., and Kroemer, G. (2012). Molecular mechanisms of cisplatin resistance. Oncogene *31*, 1869–1883.

Gatti, L., and II, Z.-F. (2005). Overview of tumor cell chemoresistance mechanisms. Chemosensitivity Vol. II.

Gilljam, K.M., Feyzi, E., Aas, P.A., Sousa, M., Müller, R., Vågbø, C.B., Catterall, T.C., Liabakk, N.B., Slupphaug, G., Drabløs, F., et al. (2009). Identification of a novel, widespread, and functionally important PCNA-binding motif TL - 186. J. Cell Biol. *186 VN-*, 645–654.

Goodarzi, A. a, Yu, Y., Riballo, E., Douglas, P., Walker, S. a, Ye, R., Härer, C., Marchetti, C., Morrice, N., Jeggo, P. a, et al. (2006). DNA-PK autophosphorylation facilitates Artemis endonuclease activity. EMBO J. *25*, 3880–3889.

Goodman, L.S., Wintrobe, M.M., Dameshek, W., Goodman, M.J., Gilman, A., and Mclennan, M.T. (1946). Nitrogen mustard therapy: Use of Methyl-Bis(Beta-Chloroethyl)amine Hydrochloride and Tris(Beta-Chloroethyl)amine Hydrochloride for Hodgkin's Disease, Lymphosarcoma, Leukemia and Certain Allied and Miscellaneous Disorders. J. Am. Med. Assoc. *132*, 126–132.

Gotlieb, W.H., Bruchim, I., Ben-Baruch, G., Davidson, B., Zeltser, A., Andersen, A., and Olsen, H. (2007). Doxorubicin levels in the serum and ascites of patients with

ovarian cancer. TL - 33. Eur. J. Surg. Oncol. *33 VN-r*, 213–215.

Greenberg, M.M. (2013). Abasic and oxidized abasic site reactivity in DNA: enzyme

inhibition, cross-linking, and nucleosome catalyzed reactions TL - 47. Acc. Chem.

Res. *47 VN-r*, 646–655.

Gregory, R.C., Taniguchi, T., and D'Andrea, A.D. (2003). Regulation of the Fanconi

anemia pathway by monoubiquitination. Semin. Cancer Biol. *13*, 77–82.

Guervilly, J.-H., Takedachi, A., Naim, V., Scaglione, S., Chawhan, C., Lovera, Y.,

Despras, E., Kuraoka, I., Kannouche, P., Rosselli, F., et al. (2015). The SLX4

Complex Is a SUMO E3 Ligase that Impacts on Replication Stress Outcome and

Genome Stability TL - 57. Mol. Cell *57 VN-r*, 123–137.

Gupta, A.K., and Anderson, T.F. (1987). Psoralen photochemotherapy. TL - 17. J. Am.

Acad. Dermatol. *17 VN-r*, 703–734.

Han, E.S., Cooper, D.L., Persky, N.S., Sutera, V.A., Whitaker, R.D., Montello, M.L., and

Lovett, S.T. (2006). RecJ exonuclease: substrates, products and interaction with

SSB TL - 34. Nucleic Acids Res. *34 VN-r*, 1084–1091.

Hanada, K., Budzowska, M., Modesti, M., Maas, A., Wyman, C., Essers, J., and Kanaar,

R. (2006). The structure-specific endonuclease Mus81-Eme1 promotes conversion

of interstrand DNA crosslinks into double-strands breaks. TL - 25. EMBO J. *25

VN-r*, 4921–4932.

Harrap, K.R. (1985). Preclinical studies identifying carboplatin as a viable cisplatin

alternative. TL - 12 Suppl A. Cancer Treat. Rev. *12 Suppl A*, 21–33.

Hartley, J.A., and Hochhauser, D. (2012). Small molecule drugs - optimizing DNA

damaging agent-based therapeutics. TL - 12. Curr. Opin. Pharmacol. *12 VN-r*, 398–402.

Hartley, J.A., Spanswick, V.J., Brooks, N., Clingen, P.H., McHugh, P.J., Hochhauser, D., Pedley, B.R., Kelland, L.R., Alley, M.C., Schultz, R., et al. (2004a). SJG-136 (NSC 694501), a Novel Rationally Designed DNA Minor Groove Interstrand Cross-Linking Agent with Potent and Broad Spectrum Antitumor Activity: Part 1: Cellular Pharmacology, In vitro and Initial In vivo Antitumor Activity TL - 64. Cancer Res. *64 VN-r*, 6693–6699.

Haynes, B., Saadat, N., Myung, B., and Shekhar, M.P.V. (2015). Crosstalk between translesion synthesis, Fanconi anemia network, and homologous recombination repair pathways in interstrand DNA crosslink repair and development of chemoresistance. Mutat. Res. Mutat. Res. *763*, 258–266.

Hazrati, A., Ramis-Castelltort, M., Sarkar, S., Barber, L.J., Schofield, C.J., Hartley, J.A., and McHugh, P.J. (2008a). Human SNM1A suppresses the DNA repair defects of yeast pso2 mutants. TL - 7. DNA Repair (Amst). *7 VN-re*, 230–238.

Hejna, J., Philip, S., Ott, J., Faulkner, C., and Moses, R. (2007). The hSNM1 protein is a DNA 5'-exonuclease. Nucleic Acids Res. *35*, 6115–6123.

Hemphill, A.W., Bruun, D., Thrun, L., Akkari, Y., Torimaru, Y., Hejna, K., Jakobs, P.M., Hejna, J., Jones, S., Olson, S.B., et al. (2008b). Mammalian SNM1 is required for genome stability. Mol. Genet. Metab. *94*, 38–45.

Henriques, J.A., and Moustacchi, E. (1980). Isolation and characterization of pso mutants sensitive to photo-addition of psoralen derivatives in Saccharomyces cerevisiae.

TL - 95. Genetics *95 VN-r*, 273–288.

Holloman, W.K. (2011). Unraveling the mechanism of BRCA2 in homologous

    recombination TL - 18. Nat. Struct. Mol. Biol. *18 VN-r*, 748.

Holzer, A.K., and Howell, S.B. (2006). The internalization and degradation of human

    copper transporter 1 following cisplatin exposure. TL - 66. Cancer Res. *66 VN-r*,

    10944–10952.

Hosono, Y., Abe, T., Ishiai, M., Takata, M., Enomoto, T., and Seki, M. (2011). The role of

    SNM1 family nucleases in etoposide-induced apoptosis. Biochem. Biophys. Res.

    Commun. *410*, 568–573.

Huen, M.S.Y., Sy, S.M.H., and Chen, J. (2010). BRCA1 and its toolbox for the

    maintenance of genome integrity. TL - 11. Nat. Rev. Mol. Cell Biol. *11 VN-r*,

    138–148.

Ishiai, M., Kimura, M., Namikoshi, K., Yamazoe, M., Yamamoto, K., Arakawa, H.,

    Agematsu, K., Matsushita, N., Takeda, S., Buerstedde, J.-M., et al. (2004). DNA

    cross-link repair protein SNM1A interacts with PIAS1 in nuclear focus formation.

    Mol. Cell. Biol. *24*, 10733–10741.

Jackson, S.P., and Durocher, D. (2013). Regulation of DNA Damage Responses by

    Ubiquitin and SUMO. Mol. Cell *49*, 795–807.

Jasin, M., and Rothstein, R. (2013). Repair of Strand Breaks by Homologous

    Recombination TL - 5. Cold Spring Harb. Perspect. Biol. *5 VN-re*.

Jin, H., and Cho, Y. (2017). Structural and functional relationships of FAN1 TL - 56.

    DNA Repair (Amst). *56 VN-r*, 135–143.

Kabotyanski, E.B., Zhu, C., Kallick, D. a, and Roth, D.B. (1995). Hairpin opening by

single-strand-specific nucleases. Nucleic Acids Res. *23*, 3872–3881.

Kasahara, K., Fujiwara, Y., Nishio, K., Ohmori, T., Sugimoto, Y., Komiya, K., Matsuda,

T., and Saijo, N. (1991). Metallothionein content correlates with the sensitivity of

human small cell lung cancer cell lines to cisplatin. TL - 51. Cancer Res. *51 VN-r*,

3237–3242.

Kashiyama, K., Shimada, M., Sasaki, K., Fawcett, H., Wing, J.F., Lewin, S.O., Carr, L.,

Li, T.-S., Yoshiura, K.-I., Utani, A., et al. (2013). Malfunction of Nuclease

ERCC1-XPF Results in Diverse Clinical Manifestations and Causes Cockayne

Syndrome, Xeroderma Pigmentosum, and Fanconi Anemia. Am. J. Hum. Genet.

92, 807–819.

Kelland, L. (2007). The resurgence of platinum-based cancer chemotherapy. TL - 7. Nat.

Rev. Cancer *7 VN-re*, 573–584.

Kikuchi, K., Narita, T., Pham, V.T., Iijima, J., Hirota, K., Keka, I.S., Mohiuddin, Okawa,

K., Hori, T., Fukagawa, T., et al. (2013). Structure-Specific Endonucleases Xpf

and Mus81 Play Overlapping but Essential Roles in DNA Repair by Homologous

Recombination. Cancer Res. *73*, 4362–4371.

Kirschner, K., and Melton, D.W. (2010). Multiple roles of the ERCC1-XPF endonuclease

in DNA repair and resistance to anticancer drugs. Anticancer Res. *30*, 3223–3232.

Klein Douwel, D., Boonen, R., Long, D.T., Szypowska, A.A., Räschle, M., Walter, J.C.,

and Knipscheer, P. (2014). XPF-ERCC1 Acts in Unhooking DNA Interstrand

Crosslinks in Cooperation with FANCD2 and FANCP/SLX4 TL - 54. Mol. Cell

54 VN-r, 460–471.

Köberle, B., Tomicic, M.T., Usanova, S., and Kaina, B. (2010). Cisplatin resistance: preclinical findings and clinical implications. TL - 1806. Biochim. Biophys. Acta *1806 VN-*, 172–182.

Kohno, T., Sakiyama, T., Kunitoh, H., Goto, K., Nishiwaki, Y., Saito, D., Hirose, H., Eguchi, T., Yanagitani, N., Saito, R., et al. (2006). Association of polymorphisms in the MTH1 gene with small cell lung carcinoma risk. Carcinogenesis *27*, 2448– 2454.

Komar, A.A., and Hatzoglou, M. (2014). Cellular IRES-mediated translation TL - 10. Cell Cycle *10 VN-r*, 229–240.

Korita, P. V, Wakai, T., Shirai, Y., Matsuda, Y., Sakata, J., Takamura, M., Yano, M., Sanpei, A., Aoyagi, Y., Hatakeyama, K., et al. (2010). Multidrug resistance-associated protein 2 determines the efficacy of cisplatin in patients with hepatocellular carcinoma. TL - 23. Oncol. Rep. *23 VN-r*, 965–972.

Kosinski, J., Feder, M., and Bujnicki, J.M. (2005). The PD-(D/E)XK superfamily revisited: identification of new members among proteins involved in DNA metabolism and functional predictions for domains of (hitherto) unknown function. BMC Bioinformatics *6*, 172.

Kutler, D.I., Patel, K.R., Auerbach, A.D., Kennedy, J., Lach, F.P., Sanborn, E., Cohen, M.A., Kuhel, W.I., and Smogorzewska, A. (2016). Natural history and management of Fanconi anemia patients with head and neck cancer: A 10-year follow-up TL - 126. Laryngoscope *126 VN-*, 870–879.

De Laat, W.L., Appeldoorn, E., Jaspers, N.G.J., and Hoeijmakers, J.H.J. (1998). DNA structural elements required for ERCC1-XPF endonuclease activity. J. Biol. Chem. *273*, 7835–7842.

Lawley, P.D., and Phillips, D.H. (1996). DNA adducts from chemotherapeutic agents TL - 355. Mutat. Res. Mol. Mech. Mutagen. *355 VN-*, 13–40.

Lewis, A.D., Hayes, J.D., and Wolf, C.R. (1988). Glutathione and glutathione-dependent enzymes in ovarian adenocarcinoma cell lines derived from a patient before and after the onset of drug resistance: intrinsic differences and cell cycle effects. TL - 9. Carcinogenesis *9 VN-re*, 1283–1287.

Li, S., Chang, H.H., Niewolik, D., Hedrick, M.P., Pinkerton, A.B., Hassig, C.A., Schwarz, K., and Lieber, M.R. (2014). Evidence That the DNA Endonuclease ARTEMIS also Has Intrinsic 5′-Exonuclease Activity TL - 289. J. Biol. Chem. *289 VN-*, 7825–7834.

Li, X., Hejna, J., and Moses, R.E. (2005). The yeast Snm1 protein is a DNA 5-exonuclease. DNA Repair (Amst). *4*, 163–170.

Liu, T., Ghosal, G., Yuan, J., Chen, J., and Huang, J. (2010). FAN1 acts with FANCI-FANCD2 to promote DNA interstrand cross-link repair. Science *329*, 693–696.

Ma, Y., Schwarz, K., and Lieber, M.R. (2005). The Artemis:DNA-PKcs endonuclease cleaves DNA loops, flaps, and gaps TL - 4. DNA Repair (Amst). *4 VN-re*, 845–851.

MacKay, C., Déclais, A.C., Lundin, C., Agostinho, A., Deans, A.J., MacArtney, T.J., Hofmann, K., Gartner, A., West, S.C., Helleday, T., et al. (2010). Identification of

KIAA1018/FAN1, a DNA Repair Nuclease Recruited to DNA Damage by Monoubiquitinated FANCD2. Cell *142*, 65–76.

Madabhushi, R., Pan, L., and Tsai, L.-H. (2014). DNA damage and its links to neurodegeneration. Neuron *83*, 266–282.

Maga, G., and Hubscher, U. (2003). Proliferating cell nuclear antigen (PCNA): a dancer with many partners. J. Cell Sci. *116*, 3051–3060.

Mandel, C.R., Kaneko, S., Zhang, H., Gebauer, D., Vethantham, V., Manley, J.L., and Tong, L. (2006). Polyadenylation factor CPSF-73 is the pre-mRNA 3'-end-processing endonuclease. Nature *444*, 953–956.

Mason, J.M., and Sekiguchi, J.M. (2011). Snm1B/Apollo functions in the fanconi anemia pathway in response to DNA interstrand crosslinks. Hum. Mol. Genet. *20*, 2549–2559.

McHugh, P.J., Ward, T.A., and Chovanec, M. (2012). A prototypical Fanconi anemia pathway in lower eukaryotes? Cell Cycle *11*, 3739–3744.

McNeil, E.M., and Melton, D.W. (2012). DNA repair endonuclease ERCC1-XPF as a novel therapeutic target to overcome chemoresistance in cancer therapy. Nucleic Acids Res. *40*, 9990–10004.

Mitsis, P.G., and Kwagh, J.G. (1999). Characterization of the interaction of lambda exonuclease with the ends of DNA TL - 27. Nucleic Acids Res. *27 VN-r*, 3057–3063.

Mitsunobu, H., Zhu, B., Lee, S., Tabor, S., and Richardson, C.C. (2014). Flap endonuclease of bacteriophage T7. 1–6.

Niedernhofer, L.J. (2007). The Fanconi anemia signalosome anchor. TL - 25. Mol. Cell *25 VN-r*, 487–490.

Niedernhofer, L.J., Daniels, J.S., Rouzer, C. a., Greene, R.E., and Marnett, L.J. (2003). Malondialdehyde, a product of lipid peroxidation, is mutagenic in human cells. J. Biol. Chem. *278*, 31426–31433.

Niewolik, D., Pannicke, U., Lu, H., Ma, Y., Wang, L.C.V., Kulesza, P., Zandi, E., Lieber, M.R., and Schwarz, K. (2006). DNA-PKcs dependence of artemis endonucleolytic activity, differences between hairpins and 5′ or 3′ overhangs. J. Biol. Chem. *281*, 33900–33909.

Niewolik, D., Peter, I., Butscher, C., and Schwarz, K. (2017). Autoinhibition of the nuclease ARTEMIS is mediated by a physical interaction between its catalytic and C-terminal domains. J. Biol. Chem. *292*, 3351–3365.

Nishino, T., and Morikawa, K. (2002). Structure and function of nucleases in DNA repair: shape, grip and blade of the DNA scissors. Oncogene *21*, 9022–9032.

Noll, D.M., Mason, T., and Miller, P.S. (2006). Formation and Repair of Interstrand Cross-Links in DNA TL - 106. Chem. Rev. *106 VN-*, 277–301.

Oestergaard, V.H., Langevin, F., Kuiken, H.J., Pace, P., Niedzwiedz, W., Simpson, L.J.,

Ohzeki, M., Takata, M., Sale, J.E., and Patel, K.J. (2007). Deubiquitination of FANCD2 Is Required for DNA Crosslink Repair. Mol. Cell *28*, 798–809.

Ogi, T., Limsirichaikul, S., Overmeer, R.M., Volker, M., Takenaka, K., Cloney, R., Nakazawa, Y., Niimi, A., Miki, Y., Jaspers, N.G., et al. (2010). Three DNA Polymerases, Recruited by Different Mechanisms, Carry Out NER Repair

Synthesis in Human Cells TL - 37. Mol. Cell *37 VN-r*, 714–727.

Olaussen, K.A., Dunant, A., Fouret, P., Brambilla, E., André, F., Haddad, V., Taranchon, E., Filipits, M., Pirker, R., Popper, H.H., et al. (2006). DNA repair by ERCC1 in non-small-cell lung cancer and cisplatin-based adjuvant chemotherapy. TL - 355. N. Engl. J. Med. *355 VN-*, 983–991.

Pannicke, U., Ma, Y., Hopfner, K.P., Niewolik, D., Lieber, M.R., and Schwarz, K. (2004b). Functional and biochemical dissection of the structure-specific nuclease ARTEMIS. EMBO J. *23*, 1987–1997.

Park, S., and Lippard, S.J. (2011). Redox State-Dependent Interaction of HMGB1 and Cisplatin-Modified DNA. Biochemistry *50*, 2567–2574.

Patrick, S.M., Tillison, K., and Horn, J.M. (2008). Recognition of cisplatin-DNA interstrand cross-links by replication protein A. Biochemistry *47*, 10188–10196.

Pawelczak, K.S., and Turchi, J.J. (2010b). Purification and characterization of exonuclease-free Artemis: Implications for DNA-PK-dependent processing of DNA termini in NHEJ-catalyzed DSB repair. DNA Repair (Amst). *9*, 670–677.

Paz, M.M., Zhang, X., Lu, J., and Holmgren, A. (2012). A new mechanism of action for the anticancer drug mitomycin C: mechanism-based inhibition of thioredoxin reductase. TL - 25. Chem. Res. Toxicol. *25 VN-r*, 1502–1511.

Pechura, C.M., and Rall, D.P. (1993). History and analysis of mustard agent and lewisite research programs in the United States. Hist. Anal. Mustard Agent Lewisite Res. Programs United States.

Poinsignon, C., Moshous, D., Callebaut, I., de Chasseval, R., Villey, I., and de Villartay,

J.-P. (2004b). The Metallo-β-Lactamase/β-CASP Domain of Artemis Constitutes the Catalytic Core for V(D)J Recombination TL - 199. J. Exp. Med. *199 VN-*.

Rahman, K.M., Thompson, A.S., James, C.H., Narayanaswamy, M., and Thurston, D.E. (2009). The pyrrolobenzodiazepine dimer SJG-136 forms sequence-dependent intrastrand DNA cross-links and monoalkylated adducts in addition to interstrand cross-links. J. Am. Chem. Soc. *131*, 13756–13766.

Rampersad, S.N. (2012). Multiple applications of Alamar Blue as an indicator of metabolic function and cellular health in cell viability bioassays. Sensors (Basel). *12*, 12347–12360.

Rechkunova, N.I., Krasikova, Y.S., and Lavrik, O.I. (2011). Nucleotide excision repair: DNA damage recognition and preincision complex assembly TL - 76. Biochem. *76 VN-r*, 24–35.

Rekoske, J.E., and Barteau, M.A. (2011). Kinetics, Selectivity, and Deactivation in the Aldol Condensation of Acetaldehyde on Anatase Titanium Dioxide TL - 50. Ind. Eng. Chem. Res. *50 VN-r*, 41–51.

Richie, C.T., Peterson, C., Lu, T., Hittelman, W.N., Carpenter, P.B., and Legerski, R.J. (2002). hSnm1 colocalizes and physically associates with 53BP1 before and after DNA damage. Mol. Cell. Biol. *22*, 8635–8647.

Rodgers, K., and McVey, M. (2016). Error-Prone Repair of DNA Double-Strand Breaks TL - 231. J. Cell. Physiol. *231 VN-*, 15–24.

Rosenberg, P.S., Greene, M.H., and Alter, B.P. (2003). Cancer incidence in persons with Fanconi anemia. TL - 101. Blood *101 VN-*, 822–826.

Roy, U., and Schärer, O.D. (2016). Involvement of translesion synthesis DNA

polymerases in DNA interstrand crosslink repair TL - 44. DNA Repair (Amst). *44

VN-r*, 33–41.

Safaei, R., Holzer, A.K., Katano, K., Samimi, G., and Howell, S.B. (2004). The role of

copper transporters in the development of resistance to Pt drugs. TL - 98. J. Inorg.

Biochem. *98 VN-r*, 1607–1613.

Sarbajna, S., and West, S.C. (2014). Holliday junction processing enzymes as guardians

of genome stability. Trends Biochem. Sci. *39*, 409–419.

Sayre, R.M., and Dowdy, J.C. (2008). The increase in melanoma: Are dietary

furocoumarins responsible? TL - 70. Med. Hypotheses *70 VN-r*, 855–859.

Sengerová, B., Wang, A.T., and McHugh, P.J. (2011). Orchestrating the nucleases

involved in DNA interstrand cross-link (ICL) repair. Cell Cycle *10*, 3999–4008.

Sengerová, B., Allerston, C.K., Abu, M., Lee, S.Y., Hartley, J., Kiakos, K., Schofield,

C.J., Hartley, J. a., Gileadi, O., and McHugh, P.J. (2012). Characterization of the

human SNM1A and SNM1B/Apollo DNA repair exonucleases. J. Biol. Chem.

287, 26254–26267.

Shao, Z., Davis, A.J., Fattah, K.R., So, S., Sun, J., Lee, K.-J., Harrison, L., Yang, J., and

Chen, D.J. (2012). Persistently bound Ku at DNA ends attenuates DNA end

resection and homologous recombination TL - 11. DNA Repair (Amst). *11 VN-r*,

310–316.

Singh, S.K., Clarke, I.D., Terasaki, M., Bonn, V.E., Hawkins, C., Squire, J., and Dirks,

P.B. (2003). Identification of a Cancer Stem Cell in Human Brain Tumors.

CANCER Res. *63*, 5821–5828.

Smeaton, M.B., Hlavin, E.M., Mason, T.M., Noronha, A.M., Wilds, C.J., and Miller, P.S. (2008). Distortion-dependent unhooking of interstrand cross-links in mammalian cell extracts. Biochemistry *47*, 9920–9930.

Stone, M.P., Cho, Y.-J.J., Huang, H., Kim, H.-Y.Y., Kozekov, I.D., Kozekova, A., Wang, H., Minko, I.G., Lloyd, R.S., Harris, T.M., et al. (2008). Interstrand DNA cross-links induced by alpha,beta-unsaturated aldehydes derived from lipid peroxidation and environmental sources. TL - 41. Acc. Chem. Res. *41 VN-r*, 793–804.

Sugasawa, K., Ng, J., Masutani, C., Iwai, S., van der Spek, P.J., Eker, A., Hanaoka, F., Bootsma, D., and Hoeijmakers, J. (1998). Xeroderma Pigmentosum Group C Protein Complex Is the Initiator of Global Genome Nucleotide Excision Repair TL - 2. Mol. Cell *2 VN-re*, 223–232.

Symington, L.S. (2014). End Resection at Double-Strand Breaks: Mechanism and Regulation TL - 6. Cold Spring Harb. Perspect. Biol. *6 VN-re*.

Teply, B.A., and Hauke, R.J. (2016). Chemotherapy options in castration-resistant prostate cancer TL - 32. Indian J. Urol. *32 VN-r*, 262–270.

Tiefenbach, T., and Junop, M. (2012a). Pso2 (SNM1) is a DNA structure-specific endonuclease. TL - 40. Nucleic Acids Res. *40 VN-r*, 2131–2139.

Tomoda, Y., Katsura, M., Okajima, M., Hosoya, N., Kohno, N., and Miyagawa, K. (2009). Functional evidence for Eme1 as a marker of cisplatin resistance TL - 124. Int. J. Cancer *124 VN-*, 2997–3001.

Touzot, F., Callebaut, I., Soulier, J., Gaillard, L., Azerrad, C., Durandy, A., Fischer, A., de

Villartay, J.-P., and Revy, P. (2010). Function of Apollo (SNM1B) at telomere highlighted by a splice variant identified in a patient with Hoyeraal–Hreidarsson syndrome TL - 107. Proc. Natl. Acad. Sci. *107 VN-*, 10097–10102.

Trujillo, J.P., Mina, L.B., Pujol, R., Bogliolo, M., Andrieux, J., Holder, M., Schuster, B., Schindler, D., and Surrallés, J. (2012). On the role of FAN1 in Fanconi anemia. Blood *120*, 86–89.

Trujillo, K.M., Yuan, S.-S.F., Lee, E.P., and Sung, P. (1998). Nuclease Activities in a Complex of Human Recombination and DNA Repair Factors Rad50, Mre11, and p95 TL - 273. J. Biol. Chem. *273 VN-*, 21447–21450.

Usanova, S., Piée-Staffa, a., Sied, U., Thomale, J., Schneider, a., Kaina, B., and Köberle, B. (2011). Cisplatin sensitivity of testis tumour cells is due to deficiency in interstrand-crosslink repair and low ercc1-xpf expression. J. Urol. *186*, 457.

Vaisman, A., and Woodgate, R. (2017). Translesion DNA polymerases in eukaryotes: what makes them tick? TL - 52. Crit. Rev. Biochem. Mol. Biol. *52 VN-r*, 1–30.

Verweij, J., and Pinedo, H.M. (1990). Mitomycin C: mechanism of action, usefulness and limitations. TL - 1. Anticancer. Drugs *1 VN-re*, 5.

Vogel, E.W., Nivard, M.J.M., Ballering, L.A.B., Bartsch, H., Barbin, A., Nair, J., Comendador, M.A., Sierra, L.M., Aguirrezabalaga, I., Tosal, L., et al. (1996). DNA damage, and repair in mutagenesis and carcinogenesis: implications of structure-activity relationships for cross-species extrapolation TL - 353. Mutat. Res. Mol. Mech. Mutagen. *353 VN-*, 177–218.

Wang, A.T., Sengerová, B., Cattell, E., Inagawa, T., Hartley, J.M., Kiakos, K., Burgess-

Brown, N.A., Swift, L.P., Enzlin, J.H., Schofield, C.J., et al. (2011a). Human SNM1A and XPF-ERCC1 collaborate to initiate DNA interstrand cross-link repair. TL - 25. Genes Dev. *25 VN-r*, 1859–1870.

Wang, C.-H., Li, J., Teo, C., and Lee, T. (1999). The delivery of BCNU to brain tumors TL - 61. J. Control. Release *61 VN-r*, 21–41.

Wang, X., Wang, S., Zhou, L., Yu, L., and Zhang, L. (2016). A network-pathway based module identification for predicting the prognosis of ovarian cancer patients TL - 9. J. Ovarian Res. *9 VN-re*, 73.

Wang, Y., Chakravarty, P., Ranes, M., Kelly, G., Brooks, P.J., Neilan, E., Stewart, A., Schiavo, G., and Svejstrup, J.Q. (2014). Dysregulation of gene expression as a cause of Cockayne syndrome neurological disease. Proc. Natl. Acad. Sci. *111*, 14454–14459.

West, S.C., Blanco, M.G., Chan, Y., Matos, J., Sarbajna, S., and Wyatt, H. (2015). Resolution of Recombination Intermediates: Mechanisms and Regulation TL - 80. Cold Spring Harb. Symp. Quant. Biol. *80 VN-r*, 103–109.

Williams, H.L., Gottesman, M.E., and Gautier, J. (2012). Replication-Independent Repair of DNA Interstrand Crosslinks. Mol. Cell *47*, 140–147.

Williams, H.L., Gottesman, M.E., and Gautier, J. (2013). The differences between ICL repair during and outside of S phase. Trends Biochem. Sci. *38*, 386–393.

Wood, R.D. (2010). Mammalian nucleotide excision repair proteins and interstrand crosslink repair. Environ. Mol. Mutagen. *51*, 520–526.

Wyatt, H., Laister, R.C., Martin, S.R., Arrowsmith, C.H., and West, S.C. (2017). The

SMX DNA Repair Tri-nuclease TL - 65. Mol. Cell *65 VN-r*, 848–1869832192.

Wyatt, H.D.M., Sarbajna, S., Matos, J., and West, S.C. (2013). Coordinated actions of SLX1-SLX4 and MUS81-EME1 for holliday junction resolution in human cells. Mol. Cell *52*, 234–247.

Yang, K., Moldovan, G.L., and D'Andrea, A.D. (2010). RAD18-dependent recruitment of SNM1A to DNA repair complexes by a ubiquitin-binding zinc finger. J. Biol. Chem. *285*, 19085–19091.

Yoshikiyo, K., Kratz, K., Hirota, K., Nishihara, K., Takata, M., Kurumizaka, H., Horimoto, S., Takeda, S., and Jiricny, J. (2010). KIAA1018/FAN1 nuclease protects cells against genomic instability induced by interstrand cross-linking agents. Proc. Natl. Acad. Sci. *107*, 21553–21557.

Yu, J., Marshall, K., Yamaguchi, M., Haber, J.E., and Weil, C.F. (2004). Microhomology-Dependent End Joining and Repair of Transposon-Induced DNA Hairpins by Host Factors in Saccharomyces cerevisiae TL - 24. Mol. Cell. Biol. *24 VN-r*, 13511364.

Yun, M.H., and Hiom, K. (2009). CtIP-BRCA1 modulates the choice of DNA double-strand-break repair pathway throughout the cell cycle TL - 459. Nature *459 VN-*.

Zhang, J., and Walter, J.C. (2014b). Mechanism and regulation of incisions during DNA interstrand cross-link repair. DNA Repair (Amst). *19*, 135–142.

Zhang, J., Dewar, J.M., Budzowska, M., Motnenko, A., Cohn, M. a, and Walter, J.C. (2015). DNA interstrand cross-link repair requires replication-fork convergence. Nat. Struct. Mol. Biol.

Zhang, X., Richie, C., and Legerski, R.J. (2002). Translation of hSNM1 is mediated by an internal ribosome entry site that upregulates expression during mitosis. TL - 1. DNA Repair (Amst). *1 VN-re*, 379–390.

Zhao, J., Jain, A., Iyer, R.R., Modrich, P.L., and Vasquez, K.M. (2009). Mismatch repair and nucleotide excision repair proteins cooperate in the recognition of DNA interstrand crosslinks. Nucleic Acids Res. *37*, 4420–4429.

Zhao, Q., Xue, X., Longerich, S., Sung, P., and Xiong, Y. (2014). Structural insights into 5′ flap DNA unwinding and incision by the human FAN1 dimer. Nat. Commun. *5*, 5726.

Zhao, Y., Lu, M., Zhang, H., Hu, J., Zhou, C., Xu, Q., Ul Hussain Shah, A.M., Xu, H., Wang, L., and Hua, Y. (2015). Structural insights into catalysis and dimerization enhanced exonuclease activity of RNase J. Nucleic Acids Res. *43*, 5550–5559.

Zhou, W., Otto, E.A., Cluckey, A., Airik, R., Hurd, T.W., Chaki, M., Diaz, K., Lach, F.P., Bennett, G.R., Gee, H.Y., et al. (2012). FAN1 mutations cause karyomegalic interstitial nephritis, linking chronic kidney failure to defective DNA damage repair. Nat. Genet. *44*, 910–915.