Functional studies of the interstrand cross-link repair protein, SNM1A and its beta-CASP domain

Beverlee Buzon
McMaster University, beverleebuzon@gmail.com

FUNCTIONAL STUDIES OF SNM1A AND ITS BETA-CASP DOMAIN

FUNCTIONAL STUDIES OF INTERSTRAND CROSSLINK REPAIR PROTEIN SNM1A AND ITS BETA-CASP DOMAIN

By BEVERLEE BUZON, B.Sc.

A Thesis Submitted to the School of Graduate Studies
in Partial Fulfillment of the Requirements for
the Degree Master of Science

McMaster University MASTER OF SCIENCE (2012) Hamilton, Ontario (Biochemistry and Biomedical Sciences)

TITLE: Functional Studies of Interstrand Cross-linking Repair Protein SNM1A and its Beta-CASP Domain
AUTHOR: Beverlee Buzon, B.Sc. (Ryerson University)
SUPERVISOR: Dr. Murray Junop
NUMBER OF PAGES: 125

ABSTRACT

Interstrand cross-linking (ICL) damage to DNA is cytotoxic as it blocks replication and transcription. This cytotoxicity is exploited in anti-cancer therapies, but increased ICL repair limits the efficacy of these chemotherapies. SNM1A (sensitive to nitrogen mustard 1A), of the beta-CASP family of nucleases, has been shown to participate in the initiation of one of the ICL repair processes. Biochemical studies of SNM1A have been limited due to insolubility and instability of SNM1A in bacteria and insect cell lines and toxicity in human cell lines. Work reported in this thesis describes a novel and efficient method of generating active protein from inclusion body expression of the beta-CASP domain of SNM1A. This refolded beta-CASP domain shows 5' exonuclease activity on single stranded and double stranded DNA *in vitro*. Nevertheless, this domain alone is unable to complement *pso2* null ICL repair defects in *S. cerevisiae* after exposure to ICL agents. These functional studies of the beta-CASP domain of SNM1A will be helpful in directing future research on its role in ICL repair. Additionally, this will aid future structural and inhibitor studies of this essential interstrand cross-link repair protein, SNM1A.

ACKNOWLEDGEMENTS

Firstly, I would like to thank my committee members Dr. Murray Junop, Dr. William Sheffield and Dr. Russell Bishop for their guidance during this difficult project and support during the difficult circumstances during this thesis. I am also indebted to them for their patience and work in their revisions to this thesis to help me achieve something that I did not believe I had the capacity to do. I would also like to acknowledge those who have also help review this thesis, including Jackie Strecker, Michelle Dowling, and Chris Brown.

I would like to thank "Team ICL", with special thanks to Dr. Tracy Tiefenbach for her guidance from the start of this project. I am grateful to the members of the Junop lab, especially who helped keep me sane and cared about me even when I was insane, including Mac Mok, Michelle Dowling, Dr. Tracy Tiefenbach, Dr. Sara Andres, Wilson Lee and Simon Huang.

I want to express my love and gratitude to my parents believing in me even when I stopped believing in myself. Your love has helped me become the daughter that you could be proud of. I also would like to give a special thanks to Ian Tam for standing behind me every step of the way.

Finally, I am thankful to Him. This would not be possible without You.

DEDICATION

For my mother. Thank you showing me the strength and courage to keep moving forward, despite the obstacles and challenges.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

53BP1	p53-binding protein 1
APC	anaphase promoting complex
ATM	ataxia telangiectasia mutated
CPSF73	cleavage and polyadenylation specificity factor 73
CFU	colony-forming units
dsDNA	double stranded DNA
DTT	dithiothreitol
EDTA	ethylenediaminetetraacetate
EME1	essential meiotic endonuclease 1
ERCC1	excision repair cross-complementing rodent repair deficiency, complementation group 1
FAM	carboxyfluorescein
FAN1	Fanconi anemia associated nuclease 1
FPLC	fast protein liquid chromatography
GSH	glutathione (reduced)
GSSG	glutathione (oxidized)
His	histidine
ICL	interstrand crosslink
IMAC	immobilized metal affinity chromatography
IPTG	isopropyl β-D-1-thiogalactopyranoside

LB	Luria Burtani
MBL	metallo beta-lactamase
MRE11	meiotic recombination 11
MUS81	methansulfonate, UV sensitive protein 81
MWCO	molecular weight cut-off
NTA	nitrilotriacetic acid
OD	optical density
PAGE	polyacrylamide gel electrophoresis
PCNA	proliferating cell nuclear antigen
PCR	polymerase chain reaction
PIAS1	protein inhibitor of activated STAT protein 1
PIP	PCNA interacting protein box
PSO2	psoralen sensitive protein 2
SDS	sodium dodecyl sulfate
SNM1A	sensitive to nitrogen mustard 1A
ssDNA	single stranded DNA
SUMO	small ubiquitin-related modifier
TAPS	N-Tris(hydroxymethyl)methyl-3-aminopropanesulfonic acid
TCEP	Tris(2-carboxyethyl)phosphine
XPF	xeroderma pigmentosa, complementation group F
YPD	yeast peptone dextrose

CHAPTER 1: INTRODUCTION

1.1 ICL damage

An interstrand cross-link (ICL) is the covalent bond joining opposing strands of duplex DNA. This type of DNA damage is toxic, as it inhibits both replication and transcription leading to replication fork collapse, double strand breaks, and eventually, if unrepaired, cell death (Noll et al, 2006). As few as forty ICLs can result in cell death in eukaryotes, making this type of DNA damage one of the most lethal (Noll et al, 2006).

ICL agents are bifunctional compounds that can cause a multitude of lesions, such as DNA-protein adducts, mono-adducts, intrastrand cross-links, and ICLs (Dronkert & Kanaar, 2001). ICLs generally only comprise a small percentage of lesions formed, but the ICL adducts damage *both* strands of DNA (Dronkert & Kanaar, 2001). This results in clastogenic events, such as chromosomal rearrangements, deletions, and other gross chromosomal alterations, consistent with double strand breaks and hyper-recombination (Dronkert & Kanaar, 2001). Although all ICL agents covalently link DNA, not all distort the helical backbone in the same manner. These differences in structural aberrations have possible implications for how and when a particular adduct is recognized and repaired (Dronkert & Kanaar, 2001).

ICLs can be formed endogenously from byproducts of lipid peroxidation, such as malondialdehyde and crotonaldehyde (Noll et al, 2006). Malondialdehyde is a known mutagen and carcinogen. It was originally thought that malondialdehyde formed ICLs under acidic conditions only (Niedernhofer et al, 2003). However, in the absence of primary amines, which compete for the reaction with the aldehyde moiety, these ICLs may persist under physiological conditions (Noll et al, 2006). Crotonaldehyde is an alpha, beta-unsaturated aldehyde which acts as a bifunctional alkylating agent, cross-linking the N2 of guanines in 5'-GC-3' sequences (Liu et al, 2006). These ICLs can spontaneously revert to monoadducts (Noll et al, 2006). The extent and stability of cross-links formed endogenously and their biological consequences are thus unclear.

Natural exogenous ICL sources are from a class of furocoumarins, which encompasses psoralen and its derivatives. These organic compounds are found in foods like parsnips, celery, and citrus fruits. Humans, on average, have a dietary daily intake of 2mg of psoralen (Sayre & Dowdy, 2008). Plant extracts containing psoralens have been used therapeutically since ancient Egyptian times for the treatment of psoriasis and vitiligo (Dronkert & Kanaar, 2001). For treatment of vitiligo, patients are prescribed about 50mg of psoralen per dose (Sayre & Dowdy, 2008). Psoralens covalently bond DNA when irradiated with long wave ultraviolet light, preferentially forming adducts with thymines in 5'-AT-3' and 5'-TA-3' sequences (Noll et al, 2006). Intercalation of psoralen unwinds the helix by approximately 25 º, resulting in significant local distortion (Guainazzi & Scharer, 2010).

1.2 ICL damage in chemotherapy

Given the potent cytoxicity of ICL agents, their use has been exploited in anti-cancer treatments. By far, the most common exposure to ICL agents is through chemotherapeutic drugs (Noll et al, 2006). The use of ICL agents for treatment of malignancies was first suggested during World War I, when it was noted that survivors of sulfur mustard attacks showed reduced numbers of leukocytes (Hirsch, 2006). The first paper on the effects of nitrogen mustards on lymphomas, by Goodman and Wintrobe, was published in 1946. Since then, ICL agents have been a mainstay of cancer therapy; these include cisplatin, nitrogen mustards, and mitomycin C (Hirsch, 2006). New anti-cancer ICL agents, like rationally designed SJG-136, are also in development (Hartley et al, 2004).

Cisplatin is one of the most widely used platinum-based anti-cancer drugs (Galluzzi et al, 2011). Derivatives of this compound, carboplatin and oxaloplatin, demonstrate reduced side effects, including decreased nephrotoxicity and neurotoxicity, but are not as effective as cisplatin (Galluzzi et al, 2011). Cisplatin is, therefore, one of the first line drugs in the treatment of solid mass tumors, and it is especially efficacious in the treatment of ovarian and testicular cancers (Galluzzi et al, 2011). Cisplatin is spontaneously activated within the cytoplasm of the cell, where concentrations of chloride ions are relatively low (Noll et al, 2006). This generates highly reactive mono- or bi-

aquated cisplatin. ICLs comprise about 5 to 8% of cisplatin adducts (Dronkert & Kanaar, 2001). Cisplatin targets 5'-GC-3' sequences to covalently link the N7 of guanines, inducing a helical bend of 47 º. In addition, it causes cytosine bases to adopt an extra-helical conformation, unwinding the DNA by 110 º (Guainazzi, 2011).

Nitrogen mustards are bifunctional alkylating agents that include clinically important anti-cancer drugs like chlorambucil and ifosfamide (Noll et al, 2006). These are used in the treatment of many types of cancers, particularly Hodgkin's and non-Hodgkin's lymphomas. The reactive N,N-bis(2-chloroethyl)amine forms an aziridinium intermediate, which almost exclusively reacts with 5'-GNC-3' in the major groove (Noll et al, 2006). This alkylates the N7 of guanines and induces a slight bend on the helical backbone (Guainazzi & Scharer, 2010). However, this N7 alkylated guanine is somewhat unstable and can undergo further reactions to uncross-link DNA, thus nitrogen mustard ICLs have a half-life of about 2 hours (Dronkert & Kanaar, 2001).

Mitomycin C is an antibiotic, originally isolated from *Streptomyces caespitosus* (Noll et al, 2006). It is used against a variety of different types of cancers. To cross-link DNA, the quinone ring of mitomycin C must be reduced, either enzymatically or chemically, and undergo a series of reactions and rearrangements before it can interact with 5'-GC-3' sequences within the minor groove (Noll et al, 2006). The resulting adduct is cross-linked at N2 of guanines and has minimal local distortion (Guainazzi & Scharer, 2010).

Another ICL agent which results in lesions with minimal DNA perturbation is SJG-136, a pyrrolobenzodiazapine dimer in Phase II of clinical trials for a broad range of cancers (Hartley et al, 2004). It preferentially binds to 5'-purine-GATC-pyrimidine-3', reacting with the N2 of guanines in that minor groove (Hopton & Thompson, 2011). Although this ICL adduct is more stable than ICLs generated with nitrogen mustards, it is heat reversible, resulting in the hydrolysis of the aminal bond (Rahman et al, 2011).

Although the use of ICL agents is widely used in the treatment of malignancies, their efficacy can be limited. Resistance of tumors to ICL agents have largely focused on cisplatin, given its widespread clinical use (Galluzzi et al, 2011). Some cancers are naturally resistant to drugs such as cisplatin (Galuzzi et al, 2011). Development of chemo-resistance to cisplatin is well documented. Despite initial success in cisplatin chemotherapy, tumor resistance to the drug leads to therapeutic failure and off-target toxicity or relapse with acquired cisplatin resistance (Galluzzi et al, 2011).

One mechanism of resistance results from increased repair of ICL lesions (Galluzzi et al, 2011). Excision repair cross-complementing rodent repair deficiency complementation group 1 (ERCC1), a nuclease involved in the repair of ICLs, has been shown to be up-regulated in cisplatin-resistant tumors and may serve as a bio-marker of the efficacy of cisplatin in treatment planning (Galluzzi et al, 2011). Given the clinical significance of the repair of ICLs, an understanding of this DNA damage response to cisplatin, and ICLs in general, is a necessity.

1.3 ICL damage repair in humans

Cells have evolved specialized pathways to combat the onslaught of DNA damaging agents, such as UV radiation, reactive oxygen species, and irradiation. However, significant exposure to ICL agents is generally deliberate, such as the use of ICL agents in anti-cancer therapies. Additionally, most of these chemotherapeutic ICL agents are synthetically made, with only mitomycin C being the exception. Despite this, mechanisms for ICL repair show conservation in budding yeast to humans (Noll et al, 2006).

ICL repair in all eukaryotes makes selective use of proteins from three different DNA repair pathways: the nucleotide excision repair, translesion synthesis, and homologous recombination pathways (Noll et al, 2006). Higher eukaryotes have also evolved a fourth pathway, the Fanconi anemia pathway, for the repair of ICLs (Deans & West, 2011). Depending when the damage is recognized (before or after replication), different subsets of these proteins are recruited.

In replication-independent repair, the ICL may be directly detected by DNA damage machinery. This is thought to occur primarily because of distortion of the helical structure of the DNA or alternatively when transcriptional machinery encounters an ICL adduct. ERCC1 and XPF (xeroderma pigmentosa, complementation group F) belong to

the nucleotide excision repair pathway and have been shown to be essential in the repair of ICLs, regardless of the cell cycle phase (Noll et al, 2006). ERCC1-XPF facilitate the removal of bulky adducts during nucleotide excision repair. Consistent with its activity, it is thought that during ICL repair, ERCC1-XPF make the first incision that begins unhooking the crosslink (Wang et al, 2011). RAD18, an ubiquitin E3-ligase, monoubiquitinates proliferating cell nuclear antigen (PCNA), a processivity factor and DNA clamp (Yang et al, 2010). PCNA recruits error-prone translesion synthesis polymerase zeta, which is able to bypass the ICL adduct to complete the repair of one strand of damaged DNA (Ho & Scharer, 2010). Another round of unhooking and lesion bypass is required to fully repair the DNA.

ICL adducts which have escaped detection during interphase trigger the replication-dependent repair pathway Subsequent replication fork collapse signals the recruitment of DNA damage repair proteins. Again, ERCC1-XPF is thought to unhook the ICL adduct (Wang et al, 2011). The incision made to the separating duplex generates a double strand break in DNA. The double strand break and presence of a sister chromatid allows for factors from the homologous recombination pathway to repair the DNA in an error-free manner after the adduct is unhooked (Noll et al, 2006).

Since the damage occurs on both strands of the DNA, concerted strand breaks are required for complete adduct removal, necessitating specialized nucleases for ICL repair. Only ERCC1-XPF, a structure-specific endonuclease, has been directly shown to

participate in the repair of ICLs *in vitro* (Galluzzi et al, 2011). Genetic knockdowns of other genes and subsequent ICL sensitivities have implicated several other nucleases in ICL damage repair; however the contribution of their known nuclease activities in ICL repair can only be inferred.

Several structure-specific endonucleases have been implicated in the repair of ICLs. These nucleases may serve to unhook the ICL adduct from the DNA. Nucleotide excision repair endonuclease ERCC1-XPF can initiate unhooking by generating a nick 5' from the ICL (Wang et al, 2011). This is consistent with the known ability of ERCC1-XPF to recognize and cleave distorted DNA structures, preferentially cleaving the 3' end of splayed arm, bubble and stem-loop DNA structures (Sengerova et al, 2011). Fanconi anemia associated nuclease 1 (FAN1) of the ICL repair Fanconi anemia pathway demonstrates endonuclease activity on 5' flap structures and replication forks (Sengerova et al, 2011). Homologous recombination nuclease, mitotic recombination 11(MRE11), shows hairpin-opening activity. Methansulfonate , UV-sensitive protein 81 (MUS81) and essential mitotic endonuclease 1 (EME1) contributes to replication restart, cleaving collapsed replication fork and 3' flaps (Hanada, 2007). SLX1-SLX4 has been shown to possess activity on stem loops, 5' flaps, and replication forks (Sengerova, 2011). Additionally, SLX4 has the ability to interact with both ERCC1-XPF and MUS81-EME1 (Sengerova, 2011). Finally, it is worth noting that SNM1A's yeast homologue, Pso2 (psoralen-sensitive protein 2), specifically acts on DNA hairpins (Tiefenbach & Junop, 2012).

Translesion synthesis for lesion bypass is more effective when unhooked DNA is resected (Ho, 2011). Therefore DNA processing by an exonuclease may be important to resect the ICL lesion for bypass by translesion synthesis polymerases. FAN1 and Pso2, in addition to structure specific endonuclease activity, also possess 5' exonuclease activity whereas Mre11 also has 3' exonuclease activity (Sengerova et al, 2011, Tiefenbach & Junop, 2012). Apollo has been implicated in the repair of a small subset of ICL agents, demonstrating 5' exonuclease activity *in vitro* (Wang et al, 2011). Finally, SNM1A may be involved in DNA resection in the repair of ICLs.

Although ICL agents are commonly used in cancer therapies, disruption of ICL processes can also result in cancer (Deans &West, 2011). Indeed, polymorphisms of ICL repair protein, SNM1A, are associated with an increased risk of small cell lung carcinoma (Kohno et al, 2006). The repair of ICL adducts is vitally important because repair defects lead to chromosomal instabilities. This is most evident with the rare genetic disorder, Fanconi anemia syndrome (Deans & West, 2011). As previously mentioned, higher eukaryotes have evolved a fourth pathway specifically for the repair of ICLs. Mutations in genes known to act within the Fanconi anemia pathway lead to congenital abnormalities, progressive aplastic anemia and bone marrow failure, as well as high incidence of hematological and squamous cell cancers (Deans & West, 2011). Compared to the general population, the risk of developing leukemia is increased by more than 700-fold (Deans & West, 2011). Patients with Fanconi anemia show increased chromosomal

aberrations, which can be specifically induced by ICL agents (Deans & West, 2011). Although Fanconi anemia syndrome is a complex disorder with many clinical manifestations, it illustrates the importance of effective ICL repair and potential consequences when ICL repair pathways are defective.

1.4 Function of SNM1A in ICL repair

SNM1A was first implicated in ICL repair when it was shown that heterologous expression of human SNM1A in budding yeast could rescue *pso2* mutant defects in ICL repair (Hazrati et al, 2008). Pso2 mutants show diminished survival in the presence of ICL agents, but not any other DNA damaging agents, suggesting that yeast Pso2 is uniquely involved in ICL repair (Hazrati et al, 2008). The ability of human SNM1A to complement *pso2* null strains and permit cell survival with exposure to ICL agents suggests that SNM1A is involved in the ICL repair response.

Of the three SNM1A knockout mice models, all homozygous *snm1a*-/- mice have modest sensitivity to ICL agent, mitomycin C (Ahkter et al, 2005, Dronkert et al, 2000, Hemphill et al, 2008). Homozygous *snm1a*-/- mouse embryonic fibroblasts are hypersensitive to mitomycin C, but this can be overcome with complementation of wild-type SNM1A (Ahkter et al, 2005, Dronkert et al, 2000, Hemphill et al, 2008). Human fibroblasts depleted of SNM1A with siRNA and treated with mitomycin C show

diminished cell survival and chromosome radial formation, an indicator of defective ICL repair (Hemphill et al, 2008). SNM1A-depleted cells treated with mitomycin C and SJG-136 also show an accumulation of MUS81-induced double strand breaks when SNM1A is not present (Wang et al, 2011).

Biochemically, SNM1A has been demonstrated to initiate replication-dependent ICL repair in conjunction with ERCC1-XPF *in vitro* (Wang et al, 2011). SNM1A can digest past SJG-136 ICLs in the presence of a 5' phosphate or nick produced by ERCC1-XPF (Wang et al, 2011). Only 5' phosphate-dependent exonuclease activity has ever been observed for SNM1A, thus it is thought that SNM1A uses this exonuclease activity to digest through the ICL (Wang et al, 2011). It is interesting to note that after the apparent translesional nuclease digestion though the ICL lesion, SNM1A pauses two nucleotides after the cross-link (Wang et al, 2011). It is unclear how or why SNM1A pauses after the cross-link, but the arrest of nuclease activity may be beneficial in protecting the DNA from further nuclease digestion during ICL repair. The postulated role of SNM1A in ICL repair is summarized in Figure 1.

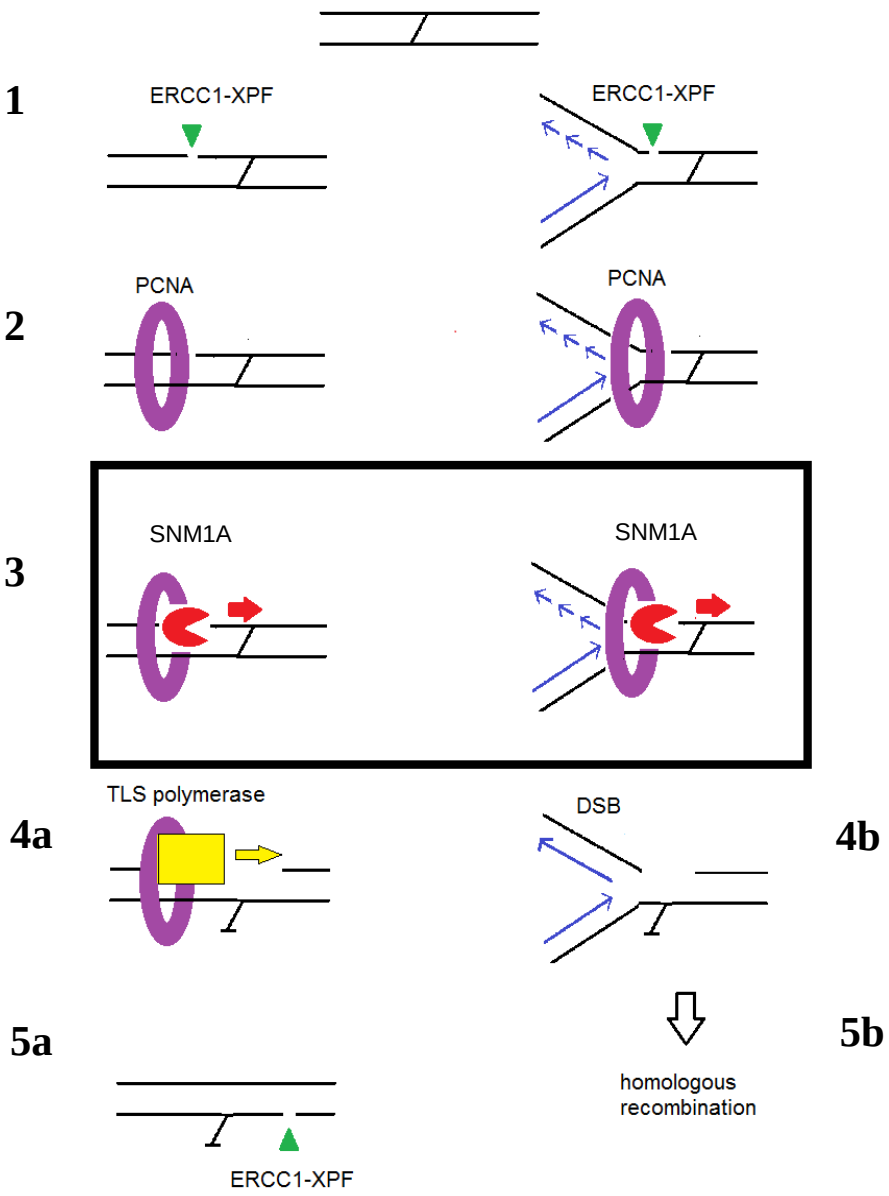


Figure 1 : Model of the postulated role of SNM1A in ICL repair.

During replication-independent repair (to the left), ERCC1-XPF makes the initial incision 5' to the ICL adduct (1). PCNA localizes to the repair site (2) and recruits SNM1A, which is thought to resect and unhook the ICL lesion (3). After resection, a translesion synthesis polymerase (likely polymerase zeta) bypasses the ICL (4a). Another round of unhooking (5a) and lesion bypass is required to restore the DNA. During replication-dependent repair, again ERCC1-XPF makes the first incision 5' to the ICL. PCNA-dependent recruitment of SNM1A allows for resection of the ICL containing strand. However, these incisions generate an ICL-induced double strand break (4b), which enables subsequent repair by homologous recombination (5b).

1.5 Additional functions of SNM1A

In addition to SNM1A's involvement in ICL repair, it has also been implicated in cell cycle regulation in the presence of other DNA damaging agents, despite not participating in the repair of these other lesions. SNM1A has been shown to regulate an early mitotic checkpoint induced by spindle poisons through a mechanism in which SNM1A negatively targets the anaphase promoting complex (APC) to arrest cell cycle progression upon exposure to mitotic stress prior to condensation of the chromosome (Akhter et al, 2004).

After DNA irradiation damage, SNM1A co-localizes with other DNA damage response proteins into distinct foci, such as p53-binding protein 1 (53BP1) (Richie et al, 2002). These foci are thought to be sites of DNA repair. SNM1A can be phosphorylated by ataxia telangiectasia mutated (ATM), a protein which initiates the activation of the DNA damage checkpoint (Akhter & Legerski, 2008). Moreover, the activation of both 53BP1 and ataxia telangiectasia mutated's (ATM's) downstream target, genomic guardian p53, is diminished in cells deficient in SNM1A (Akhter & 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 (Akhter & Legerski, 2008).

SNM1A knockout mice also exhibit defective cell cycle regulation. Although these mice are viable and fertile, they have accelerated tumorigenesis and decreased survival compared to wild-type mice (Ahkter et al, 2005). Interestingly, decreased survival in SNM1A knockout mice was a result of different causes between genders. Tumorigenesis was the cause of death for female mice, whereas the cause of death for male mice was from bacterial infections of the mandibular and preputial glands, suggesting SNM1A may be involved in immunity in addition to cell cycle regulation (Ahkter et al, 2005).

1.6 Structure and Function of SNM1A

The full-length protein of SNM1A is comprised of 1040 amino acids, with the metallo beta-lactamase and CASP homology domain residing at the C-terminal region of the protein, as shown in Figure 2. There are two localization regions towards the N-terminus of the protein. The nuclear localization region at 1-190 transports the protein into the nucleus, while the nuclear focus region at 396-614 localizes SNM1A from a larger nuclear body into smaller distinct foci at the sites of DNA repair (Richie et al, 2002).

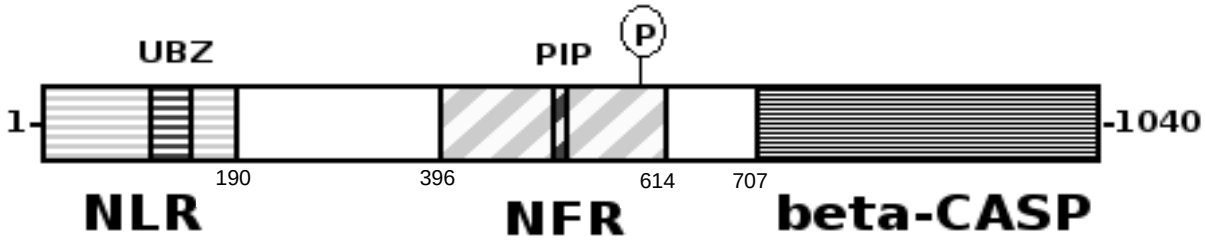


Figure 2: Domain boundaries of SNM1A.

The N-terminus of SNM1A encompasses localization regions for nuclear localization and foci formation. It also contains a functional UBZ and PIP box. The C-terminus encompasses the beta-CASP domain. NLR = nuclear localization region, NFR = nuclear foci region, beta-CASP = metallo beta-lactamase + CASP homology domain, UBZ = ubiquitin-binding zinc finger, PIP = PCNA interacting protein box, P = phosphorylated residue S590.

Also, there is a functional ubiquitin-binding zinc finger (UBZ) domain at amino acids 118-145 and PCNA interacting protein box (PIP box) at amino acids 556-563, both of which have been shown to act with PCNA and monoubiquitin (Yang et al, 2010). The PIP box is essential for SNM1A nuclear foci formation and mutations within these residues abolish SNM1A's ability to localize properly (Yang et al, 2010). Although SNM1A has a functional UBZ, monoubiquitination of PCNA is not essential for SNM1A's recruitment to PCNA, but it enhances SNM1A's ability to bind to it (Yang et al, 2010). Though it is clear that PCNA is an essential interacting partner, it is possible that other monoubiquitinated proteins other than PCNA may interact with SNM1A.

Functionally, full-length SNM1A has been shown to have 5' phosphate specific exonuclease activity, showing non-processive activity on single-stranded DNA (ssDNA) and minimal activity on double-stranded DNA (dsDNA) (Wang et al, 2011). However, N-terminally truncated SNM1A$^{608\text{-}1040}$ shows increased ssDNA exonuclease activity. This truncation has been reported to have a k_{cat} of 26 min^{-1} and 14 min^{-1}, with a K_m of almost 100nM and less than 5nM, for dsDNA and ssDNA oligos, respectively (Wang et al, 2011). With high molecular-weight DNA, however, SNM1A$^{608\text{-}1040}$ has increased processivity, reportedly being capable of digesting kilobases of plasmid DNA (Wang et al, 2011) . Hairpin-opening endonuclease activity has been tested for repeatedly for purified full-length SNM1A, but has not been observed (Hazrati et al, 2008, Hejna et al, 2007, Wang et al, 2011).

1.7 The beta-CASP domain

SNM1A is a member of the beta-CASP family, named after the representative members CPSF-73, Artemis, SNM1 and Pso2. They were originally identified using hydrophobic cluster analysis, and have subsequently been shown to share 5' phosphate dependent exonuclease activity (Cattell et al, 2010; Callebaut et al, 2002).Cleavage and polyadenylation specificity factor 73 (CPSF-73) is a 5' mRNA processing nuclease (Mandel et al, 2006). Artemis (SNM1C) is involved in processing V(D)J recombination intermediates and subset of ionizing radiation DNA damage. It has both 5'exonuclease activity and endonuclease activity, specifically acting on DNA hairpins, requiring constant association with and phosphorylation by DNA-PKcs (Bonatto et al, 2005). Apollo (SNM1B), Artemis' "twin brother", is a 5'exonuclease, functioning in telomere maintenance with TRF2 (Yan et al, 2010). It has also been implicated in the Fanconi anemia pathway of ICL repair via its interaction with SLX4 (Selewsky, 2012). The yeast homologue, Pso2, is required for ICL damage repair; however its specific role is uncertain (Cattell et al, 2010). Pso2 has also been shown to have hairpin-opening endonuclease activity in addition to its processive 5' exonuclease activity (Tiefenbach & Junop, 2012).

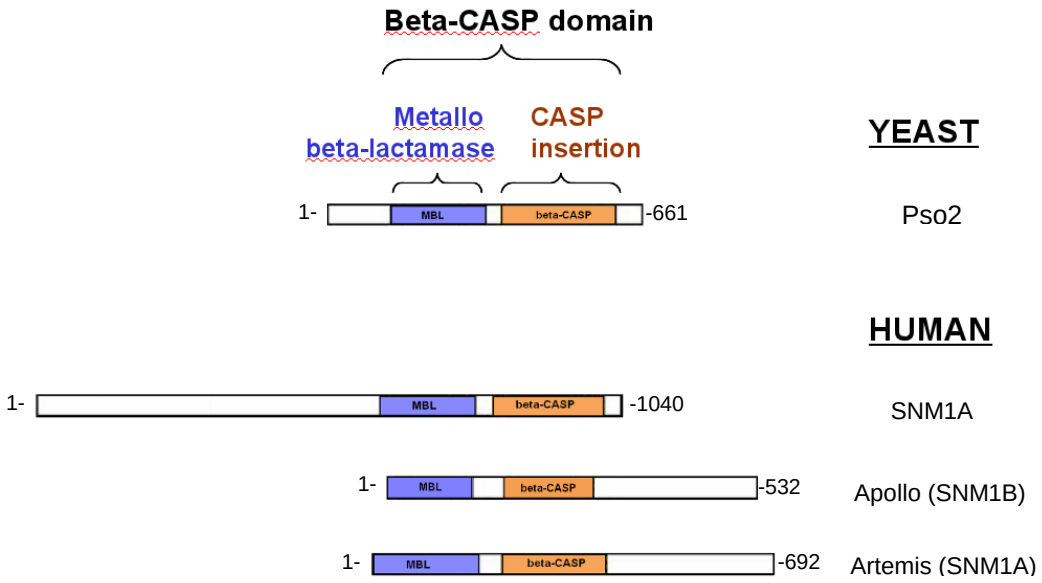


Figure 3: Alignment of beta-CASP DNA repair nucleases.

Alignment of beta-CASP nucleases involved in DNA repair. Yeast has only one beta-CASP DNA nuclease, Pso2. However, there are three homologues of Pso2 in humans: SNM1A, Apollo, and Artemis. The metallo beta-lactamase domain is in purple and the CASP homology domain is in orange.

I

All members of this family have a beta-CASP domain that is composed of the highly conserved metallo beta-lactamase (MBL) fold and a CASP homology domain, as seen in Figure 3 (Cattell et al, 2010). Proteins containing the MBL fold belong to a larger super-family which hydrolyzes a wide variety of substrates, most having an ester linkage and negative charge. The MBL fold is comprised of a four-layered beta-sandwich flanked by four alpha-helices, with zinc metal binding sites at the edge of the sandwich (Carfi et al, 1995). Despite the sequence diversity among MBL containing proteins, they all contain five sequence motifs of histidine and aspartic acid residues which appear to coordinate the bi-nuclear zincs within the active site required for the hydrolysis reaction (Callebaut et al, 2002). The entire super-family contains the highly conserved HxHxDH motif (Callebaut et al, 2002).

Amino acid sequence alignment of members within the MBL family of nucleases reveals the presence of only four of the five signature motifs, as shown in Figure 4. Also of interest, there is a region of shared homology (the CASP domain) with three additional conserved motifs. However, motif B of the CASP homology domain is likely to represent motif 5 of the canonical MBL motifs (Callebaut et al, 2002). Indeed, from the crystal structure of CPSF-73, six parallel strands surrounded by helices make the core of the CASP homology domain, but additional helices from the CASP domain complete the MBL fold, as seen in gold on the bottom domain in Figure 4 (Mandel et al, 2006). Thus, the CASP homology domain can be thought to act as an insertion cassette within the MBL fold (Mandel et al, 2006). From the structure of CPSF-73, it is unclear where a nucleic

acid substrate might be accommodated in order to interact with the MBL active site (Mandel et al, 2006). However, it is possible that the CASP domain may act as a lid for substrate specificity and/or regulation (Mandel et al, 2006).

Mutagenesis of these motifs and the corresponding phenotypes has helped facilitate understanding of this beta-CASP domain. Mutation of the aspartic acid residue in the motif 2 HxHxDH signature, such as in mutants SNM1A ^{D736A}, Pso2 ^{D252A}, and Artemis^{D35A}, results in either severely diminished or abolished nuclease activity (Hejna et al, 2007, Li et al, 2005, Ma et al, 2002). Mutations in motif A and B of the CASP domain in SNM1A results in protein mislocalization from protein inhibitor of activated STAT protein 1 (PIAS1), an E3 SUMOylation ligase. These mutations also cause decreased survival after exposure to cisplatin (Ishiai et al, 2004). Interestingly though, it does not abolish nuclease activity *in vitro* (Wang et al, 2011). Molecular mechanisms of these CASP motifs remain poorly understood, as is their precise role in DNA repair.

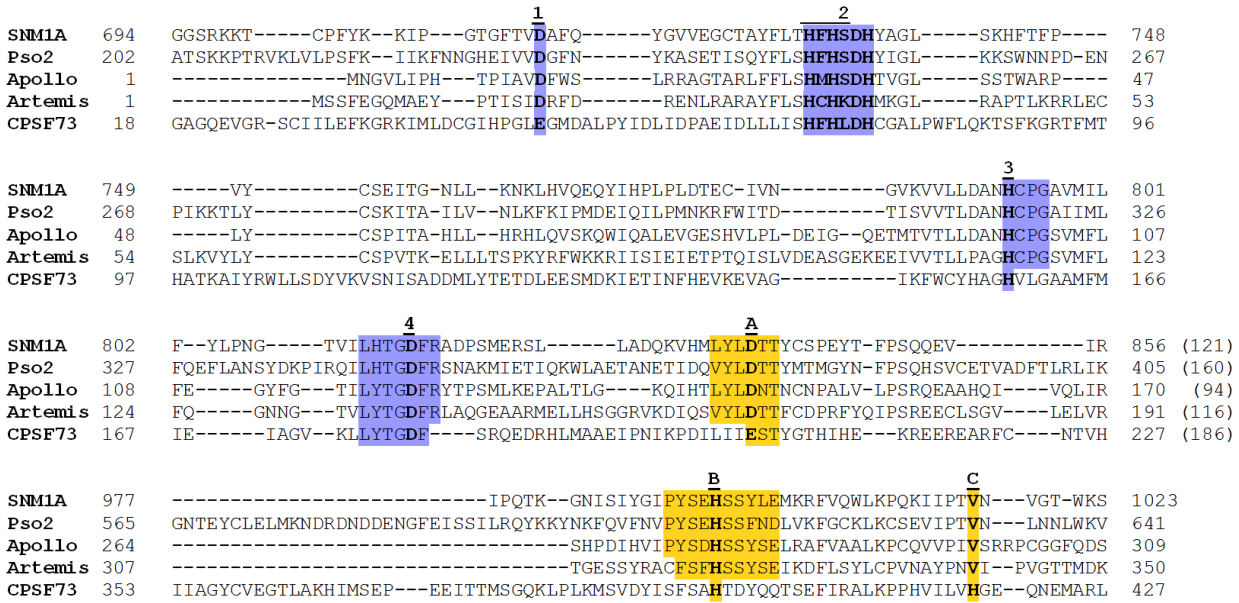


Figure 4: Amino acid sequence alignment of the beta-CASP domain of representative beta-CASP homologues.

Bolded residues are signature motifs of either the metallo beta-lactamase domain (1-4) or CASP homology (A-C) domains. Residues in purple are metallo beta-lactamase conserved motifs among the homologues. Residues in gold are CASP-specific conserved motifs among the homologues.

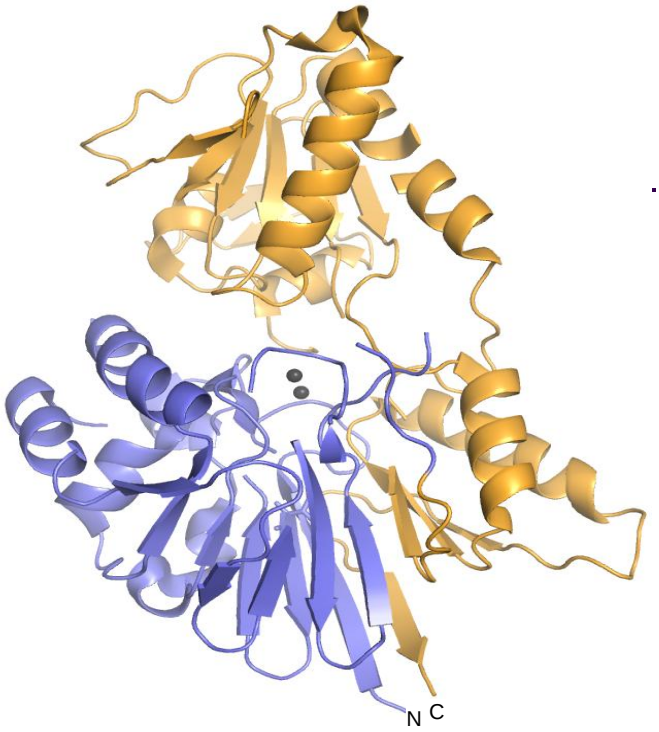


Figure 5: Crystal structure of beta-CASP homologue, CPSF73.

Residues in purple form the MBL domain. Residues in gold are from the beta-CASP homology domain. In the centre of the structure is the bi-nuclear zinc active site (PDB: 2I7T**).**

Functionally, the beta-CASP domain appears to encompass the catalytic core of the nuclease. The beta-CASP domain of Artemis (1-385) has been shown to suffice for V(D)J recombination (Poinsignon, 2004). However, it is insufficient to repair DNA damage by ionizing radiation. This may be due to its requirement for phosphorylation and association with DNA-PKcs (Poinsignon, 2004). It is thought that the C-terminus of Artemis acts as a negative regulator of this enzyme and truncation of it permits hairpin-opening activity, however it is not clear if hairpin-opening activity is required for the repair of double strand breaks. Purified Apollo$^{6-328}$ demonstrates 5' exonuclease activity *in vitro*, but it is unknown if truncation of its C-terminus allows for telomere maintenance or Fanconi anemia-associated ICL repair (Poinsignon et al, 2004). The crystal structure of CPSF-73 was derived using CPSF$^{1-460}$ (Mandel et al, 2006). Although the biological consequences of truncating the unconserved C-terminal region of CPSF-73 have not been explored, CPSF$^{1-460}$ has both endonuclease and 5' exonuclease activity, sufficient for mRNA processing *in vitro* (Mandel et al, 2006).

SNM1A's functional homologue, Pso2, on the other hand, requires residues from its N-terminus for activity. The beta-CASP domain of Pso2 on its own (Pso2 $^{188-661}$) is unable to support DNA binding or nuclease activity (Tiefenbach, 2011). Moreover, Pso2$^{188-661}$ cannot participate in ICL repair (Tiefenbach, 2011). Protein-protein interaction studies of Pso2 have not demonstrated any proteins which associate with it. Overall, it appears that beta-CASP proteins have these domains under different means of regulation. Strict regulation of nucleases is vital to protect the integrity of the genome.

1.8 Thesis Objective

Work described in this thesis is focused toward answering the central questions: Does the beta-CASP domain of SNM1A retain nuclease activity *in vitro* and can this domain partake in ICL repair *in vivo*?

The characterization of SNM1A is important, given its specific role in ICL repair. Despite this, the studies on SNM1A within the literature are remarkably limited. In terms of the biochemical characterization, this is directly related to difficulties with SNM1A expression, solubility and proteolytic degradation. Overexpression of SNM1A is toxic in human cell lines. It is poorly expressed in insect and yeast cell lines, and essentially insoluble in *E. coli*. A recent publication reports SNM1A $^{608\text{-}1040}$ expressed in *E. coli* can be purified; however, it requires six different columns to purify protein from twenty-four litres of cultured bacteria to yield only microgram quantities of protein (Wang et al, 2011).

Many of the limitations in understanding SNM1A might be circumvented if greater amounts of active protein were readily available. It may be possible, however, to circumvent protein expression and purification problems of SNM1A by pursuing the minimal functional unit of the nuclease, encompassed by the beta-CASP domain, or by refolding the insoluble recombinant protein expressed as inclusion bodies.

It is hypothesized that the N-terminal domain of SNM1A acts as a negative regulator of the nuclease activity intrinsic to the beta-CASP domain. Accordingly, it can be postulated that the beta-CASP domain of SNM1A might retain sufficient nuclease activity for the repair of ICL DNA damage. Although Artemis demonstrates endonuclease activity when truncated solely to its beta-CASP domain, SNM1A's functional homologue, Pso2's beta-CASP domain does not exhibit any nuclease activity alone. Therefore, it must be determined whether the minimal beta-CASP domain of SNM1A behaves similarly to that of Artemis or Pso2.

Given the specificity of SNM1A to ICL repair, inhibition of this repair protein may be critical for effective anti-cancer therapies, especially in tumours that have become resistant to ICL agents, such as cisplatin. The beta-CASP domain could help elucidate SNM1A's exact function in ICL repair and its significance in the repair pathway as well as provide a suitable target for inhibitor development.

CHAPTER 2: MATERIALS AND METHODS

2.1 Cloning of SNM1A variants

Expression vectors were generated using Gateway cloning technology (Invitrogen). Briefly, DNA primers containing attB1 recombination sequences were used to amplify a target sequence from plasmid DNA. To generate an initial entry clone, the polymerase chain reaction (PCR) product was incubated with pDONR201 and BP Clonase II overnight to permit the recombination of the target gene into the entry vector. Following Proteinase K digestion for ten minutes at 37 ° C the next day, the resulting plasmids were transformed into TOP10 *E. coli* (Invitrogen). Colony PCR was used to confirm the correct integration of PCR product. Plasmid DNA was isolated from cells containing the correctly integrated PCR product using a high-speed mini-prep kit (Geneaid). In all cases, DNA sequencing was used to verify the sequence of the entry clone.

To generate the final expression vectors, entry clones were incubated overnight with the appropriate destination vectors and LR Clonase II. The resulting plasmids were again digested with Proteinase K for 10 minutes at 37 ° C and transformed into TOP10 cells. Recombination of the target gene into the destination vector was confirmed with colony PCR. Plasmid DNA was isolated using a high-speed mini-prep kit. A list of primers and destination vectors used for the construction of all expression vectors can be

found in the Appendix.

2.2 SNM1A expression in *S. cerevisiae* and purification

For protein expression in *S. cerevisiae*, a 100mL culture was grown at 30 ° C in uracil-lacking minimal media supplemented with 2% raffinose for 24 hours until full density was achieved. This culture was subsequently used to inoculate a 1L culture of uracil-lacking minimal media with 2% galactose for induction and 1% raffinose such that the OD_{600} was between 0.4-0.6. The culture was grown for 16 hours at 30 ° C, after which the cells were harvested by centrifugation at 33,000 x g.

For cell lysis, liquid nitrogen grinding was employed. After harvesting the cells, cell pellet was placed into a syringe and passed through slowly into liquid nitrogen, generating small drops of yeast. For lysis, the frozen small yeast pellets were mechanically ground using a pestle and mortar into a fine powder. The powder was resuspended in 20mM Tris pH 8, 500mM NaCl, and a protease inhibitor cocktail (1 mM benzamidine-HCl, 237 μM leupeptin, 1 μM PMSF, 1 mM pepstatin A, 1 mM aprotinin). Benzamidine, leupeptin and aprotinin were added again after lysis. Lysate was clarified first by centrifugation at 48,000 x g at 4 ° C and then filtration on ice prior to fast protein liquid chromatography (FPLC) purification.

The clarified lysate was loaded onto a 5mL HP Hi-Trap nickel column (GE Healthcare Life Sciences) equilibrated with nickel buffer (20mM Tris pH 8, 500mM

NaCl). The column was washed successively with 5 column volumes of nickel buffer containing 15mM, 30mM, or 45mM imidazole, and then eluted with nickel buffer with 210mM imidazole. Purification was monitored by analysis of fractions using sodium dodecyl-polyacrylamide gel electrophoresis (SDS-PAGE) and western blotting with anti-His monoclonal mouse antibodies (Abcam) + goat anti-mouse conjugated with alkaline phosphatase (Bio-Rad).

2.3 SNM1A and beta-CASP domain expression in *E. coli*

Full-length SNM1A (1-1040), SNM1A^{D736A}, beta-CASP domain (707-1040), and beta-CASP^{D736A} were expressed in BL21 (DE3) Star/pLysS *E. coli (*Invitrogen), harbouring the pRARE plasmid. Cells were grown at 37 ° C in Luria Burtani media (LB) supplemented with 50 μg/mL ampicillin and 25 μg/mL chloramphenicol to an OD_{600} between 0.4 and 0.6, induced with 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG), and then further grown for another 4 hours at 30 ° C. Cells were harvested by centrifugation at 33,000 x g for 15 minutes, and then stored at -80 ° C.

2.4 Inclusion body isolation

Frozen cell pellets were resuspended in resuspension buffer (50mM Tris pH 8, 50mM NaCl, 1mM Tris(2-carboxyethyl)phosphine (TCEP), 50μM EDTA, 5% glycerol) and lysed by sonication. 1% Triton X-100 was added to the lysate and the mixture was allowed to stir slowly for 20 minutes at room temperature. Insoluble inclusion bodies

were separated by centrifugation at 8,000 x g. The pellet fraction was then resuspended in wash buffer (50mM Tris pH 8, 50mM NaCl, 1mM TCEP, 50μM ethylenediaminetetraacetate [EDTA], 5% glycerol, 1% Triton X-100) and centrifuged as before. The resulting inclusion body fraction was resuspended in 5M urea containing 20mM dithiothreitol (DTT). Residual Triton X-100 and urea were removed by resuspending the pellet in resuspension buffer (above) for the final wash, and then centrifuged once more to collect the final inclusion body material. Isolated inclusion body material was resuspended in denaturation buffer (50mM Tris pH 8, 200mM NaCl, 2mM EDTA, 7M guanidine-HCl, 100mM TCEP) and stirred slowly for 2 hours to solubilize the inclusion body components. Any unsolubilized material was removed by centrifugation at 21,000 x g at 4 ° C. The supernatant containing the solubilized material was stored at -20 ° C.

2.5 Inclusion body solubilization assay

To ensure the same amount of inclusion body material was used for each assay, the inclusion bodies were resuspended in distilled water and aliquoted. The inclusion body material was then centrifuged at 21,000 x g for 5 minutes and the water was removed. The inclusion bodies were solubilized with guanidine containing buffer (50mM Tris pH 8, 200mM NaCl, 2mM EDTA, 7M guanidine-HCl), SDS buffer (0-2% SDS, 100mM Tris pH 8) or buffer composed of 2M urea and 50mM Tris, ranging from pH 1 to pH 12.

To determine the concentration of solubilized inclusion body protein, the resulting solutions were centrifuged at 21,000 x g for 5 minutes. The supernatant was transferred to a new 1.5 mL microcentrifuge tube and the concentration of inclusion bodies in the supernatant was determined using the Bradford assay (Bio-Rad).

2.6 Screening for aggregation of inclusion body material

For initial aggregation screening, solubilized inclusion body material resuspended in guanidine buffer (above) was rapidly diluted (1:50) using different buffering conditions contained within the iFOLD System 3. After gentle rocking for one hour, the OD_{340} was measured using a plate reader (Tecan). Readings of absorbance were correlated with aggregation.

For subsequent aggregation screening, solubilized inclusion body material was diluted (1:50) into 1mL of refolding solutions containing 50mM buffer, 0.1325g FoldACE4, and a reducing agent, as indicated in Table 2 on page 42. After gentle rocking for 30 minutes, the amount of aggregation was measured as absorbance at OD_{340}.

2.7 IMAC purification of protein from inclusion bodies under denaturing conditions

The solubilized inclusion body material was further purified using immobilized metal affinity chromatography (IMAC) nickel-nitrilotriacetic acid (Ni-NTA) agarose beads (Qiagen). Briefly, 1 ml of Ni-NTA agarose resin was washed twice with urea buffer (8M urea, 50mM Tris pH 8) and centrifuged for five minutes at 5,000 x g. Prepared

inclusion body sample was diluted with urea buffer to reduce the TCEP concentration to less than 5mM and incubated with Ni-NTA beads for one hour with gentle rocking. The beads were washed three times with urea buffer containing 10mM DTT, then once with guanidine buffer (6M guanidine, 200mM NaCl, and 50mM Tris pH 8). The purified inclusion protein was eluted with either guanidine buffer containing 500mM imidazole or guanidine containing buffer at pH 4.5. The amount of inclusion body protein that remained bound to the beads was determined by boiling 20μL of beads resuspended in SDS loading buffer. Protein was analyzed using SDS-PAGE and quantified using the Bradford assay.

2.8 Refolding of protein purified from inclusion bodies

Protein purified from inclusion bodies were refolded by rapid dilution (1:50) into 50mM N-Tris(hydroxymethyl)methyl-3-aminopropanesulfonic acid (TAPS) pH 8.5, 8mM reduced glutathione + 2mM oxidized glutathione, 1.325g FoldACE4 per 20mL (Novagen) and two molar equivalents of zinc acetate. After 1 hour incubation at 4 º C, the solution was centrifuged at 20,000 x g to remove any aggregates. The refolded protein was then concentrated in a centrifuge concentrator (Corning) of either 30 kDa molecular weight cut-off (MWCO) for the beta-CASP domain or 50 kDa MWCO for full-length SNM1A. An equal volume of glycerol was added to the concentrated refolded protein for storage at -20° C. Alternatively, 10% glycerol v/v was added for storage at -80 ° C. Protein purity was analyzed using SDS-PAGE and western blotting with anti-His tag

monoclonal mouse antibodies and goat anti-mouse conjugated with alkaline phosphatase. Protein concentration was determined using the Bradford assay.

2.9 Assays for determining nuclease activity of SNM1A

For 5' labeling of a polyT 20mer, DNA (4μM) was incubated for 3 minutes with PNK buffer (NEB) and 150uCi of gamma ^{32}P ATP (Perkin Elmer) at 90 ° C, then incubated on ice before adding 1μL of PNK (NEB). After incubation at 37 ° C for 30 minutes, the reaction was passed through a Biospin-6 column (Bio-Rad) to remove excess ATP.

To test exonuclease activity, 1μL of protein (0.5-1μM), 1μL of 5' labeled DNA (0.5-1μM), 2.5μL of 10X reaction Buffer F (50mM Tris acetate pH 7.2, 10mM magnesium acetate, 75mM potassium acetate, and 1mM DTT) and water were combined to a volume of 25μL and incubated at 37 ° C for 30 minutes. The reaction was stopped by addition of 25μL of denaturing DNA loading buffer (7M urea, 50% glycerol, 2% xylene cyanol, 2% bromophenol blue, 50mM EDTA). Reaction components were then resolved using 10% urea denaturing PAGE electrophoresed at 40mA. Reaction components within the gel were imaged using a phosphor screen (Kodak) and Typhoon imager (Amersham).

2.10 Assays for determining nuclease activity of the beta-CASP domain of SNM1A

Oligonucleotides for use in nuclease assays were purchased with a 3’ 6-fluorescin amidite (6-FAM) label (Biobasic) for detection by fluorescence. To test nuclease activity, 1μL of beta-CASP protein(2-4μM), 1μL of DNA (0.5 μM), 2.5μL of 10X reaction buffer F (above), 0.25μL of BSA (10mg/ml) and water were combined to a volume of 25μL and incubated at 37 ° C for up to 1 hour. The reaction was stopped by addition of 25μL of denaturing DNA loading buffer. Reaction components were resolved by electrophoresis on a 20% urea denaturing PAGE electrophoresed at 40mA. The gel was imaged with a Typhoon imager (Amersham) at 526 nm.

2.11 Enzyme kinetics determination

Enzyme concentrations for initial velocity rates were determined under conditions where 10% of the substrate was hydrolyzed at the 16 minute time point. 5-15 nM of beta-CASP protein was used to assay 25 to 10,000 nM of ssDNA and dsDNA containing 5' phosphorylation and 3' 6-FAM label modifications. Nuclease reactions consisted of 5-15 nM of beta-CASP protein, 25 to 10,000 nM of ssDNA and dsDNA containing 5' phosphorylation and 3' 6-FAM label modifications, 1X reaction buffer F, BSA (1mg/ml) and water were combined to a volume of 10 μL and incubated at 37 ° C for 16 minutes. Reaction tubes were pre-warmed to 37 ° C before the addition of enzyme, and stopped with the addition of denaturing DNA loading buffer. Products were electrophoresed on a 30% urea denaturing PAGE at 40mA. The gel was imaged with the Typhoon Imager at

526 nm.

The products were analyzed using ImageJ (NIH). To determine the non-processive exonuclease rate, the negative control and the first three nucleolytic products were quantified (n-1, n-2, n-3) so that each product could be converted into a percentage of total DNA. The empty lane adjacent to the first lane was measured to calculate the background and subtracted from each measurement. The percentage of DNA of each product was multiplied by the number of exonuclease events, then added and multiplied by the total amount of DNA to determine the exonuclease rate, where the rate = [%(n)x0 + %(n-1)x1 + %(n-2)x2 + %(n-3)x3] x (amount of DNA) (Cheng & Kuchta, 1993; Elisseeva et al, 1999). Kinetic parameters were determined by non-linear regression analysis (GraphPad) using standard Michaelis-Menten curve fitting.

2.12 Complementation assays

To transform yeast with expression vectors, one colony of *pso2* null *S. cerevisiae* used to inoculate 5 ml of yeast peptone dextrose (YPD) media and grown overnight in 30 ° C with shaking at 1,000 rpm. When the cells were in exponential phase, about 18 hours later, 1 ml of the culture was transferred to a new 1.5 mL microcentrifuge tube and centrifuged at 1,000 x g for 2 minutes at room temperature. After removing the supernatant, 20 μL of plasmid DNA (50-150 ng/μL) and 10 μL (10μg/μL) of single

stranded herring DNA were added to the yeast pellet then vortexed briefly. 1 mL of transformation buffer (40% polyethylene glycol 4000, 100mM lithium acetate, 10mM Tris pH 7.5, 1mM EDTA pH 8) was added. The mixture was incubated overnight at room temperature. The cells were centrifuged at 1000 x g for 5 minutes and 80% of the supernatant was removed. Cell pellets were resuspended in the remaining 20% of supernatant and plated on uracil-lacking minimal media agar with 2% glucose and grown for 3 days at 30°C.

A single colony of transformed yeast was grown to stationary phase in uracil-lacking minimal media with 2% glucose for 48-60 hours at 30 ° C. 500 μL of the culture was transferred to a microcentrifuge tube and centrifuged for 2 minutes at 1000 x g. The cells were washed three times with phosphate buffered saline (PBS) by resuspending cells in PBS, then centrifuging at 1000 x g and removing the supernatant. The cells were then resuspended in 500 μL of PBS. The cells were then treated with either 1 mM cisplatin or water, and incubated at 30 ° C with vigorous shaking at 3,000 rpm. After one hour, the cells were diluted into PBS and plated at various concentrations.

For yeast ICL damage sensitivity assays, 5 μL of each dilution of transformed ICL treated yeast was plated onto uracil-lacking minimal media agar supplemented with 2% galactose to induce heterologous protein expression. Plates were incubated for 3 days at 30 ° C. The plate was imaged at UV transilluminator (BioDoc-It System) under UV light.

For yeast ICL damage survival assays, 100µL of the transformed ICL-induced yeast were plated on uracil-lacking minimal media agar supplemented with 2% galactose to induced protein expression. Only plates containing 30 to 300 colonies were used to determine the number of colony forming units (CFUs). The number of CFUs was calculated for each expression vector was calculated and normalized to the number of CFUs formed by cells expressing positive control, Pso2, to determine % survival.

CHAPTER 3: RESULTS

3.1 Results for SNM1A

3.1.1 Expression and purification of SNM1A

SNM1A was expressed in *S. cerevisiae* by inducing cells with 2% galactose and harvesting cells after 16 hours, a time point experimentally determined to express detectable amounts of SNM1A (data not shown). As shown in Figure 6, comparison of induced and uninduced cells does not reveal a significant induction of full-length SNM1A protein, which should have been detected at 116 kDa. However, a very faint band at 37 kDa was detected in the induced cells using western blotting, suggesting possible proteolytic cleavage of the protein.

The wash containing 45mM imidazole and elution with 210 mM imidazole yields a protein of about 37 kDa on the Coomassie stained SDS-PAGE gel, but this band was not detected with western blotting. Although not detected via Coomassie staining, a band at 70kDa and two smaller proteins at 55 kDa and 45kDa were observed with western blotting. Neither SDS-PAGE nor western blotting showed the presence of full-length SNM1A (expected size of 116 kDa) for any fractions collected during the entire purification, suggesting poor expression of SNM1A that was subject to proteolytic cleavage.

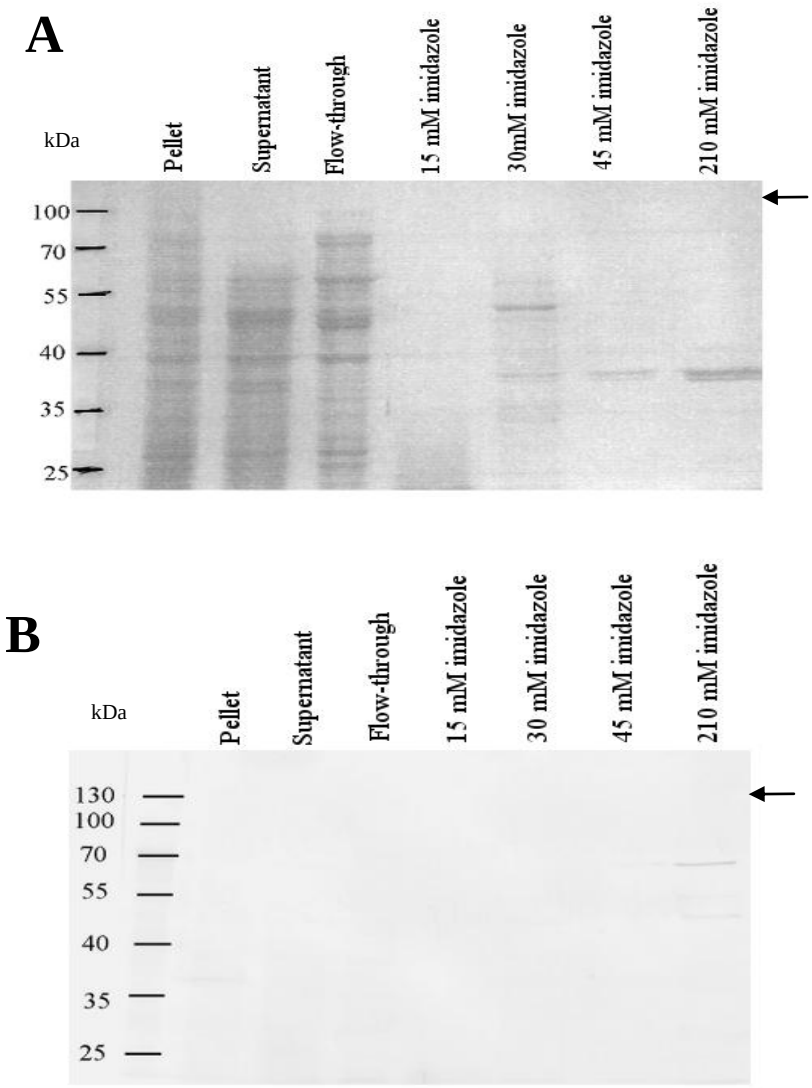


Figure 6: Analysis of fractions from a large-scale purification of SNM1A expressed in *S. cerevisiae.*

SNM1A was expressed in *S. cerevisiae*. One litre of cells was grown for 16 hours after induction using 2% galactose, and then lysed by manually grinding cells in liquid nitrogen. SNM1A was purified using IMAC nickel FPLC. Fractions were collected and electrophoresed on 10% SDS-PAGE gels. A) Coomassie-stained SDS-PAGE IMAC nickel purification. B) Western blot of IMAC nickel purification. Arrows indicate the expected size of full-length SNM1A.

Although SNM1A could be expressed in *E. coli*, results suggested that the expressed protein was subject to significant proteolytic degradation (see Figure 7B). Additionally, SNM1A was highly insoluble, localizing to the pellet as an inclusion bodies (see Figure 7A) at a range of temperatures. SNM1A expressed as inclusion bodies could be isolated with washes of 1% Triton X-100. Inclusion body isolation increases full-length SNM1A to about 70% purity (see 3rd lane, Figure 7A).

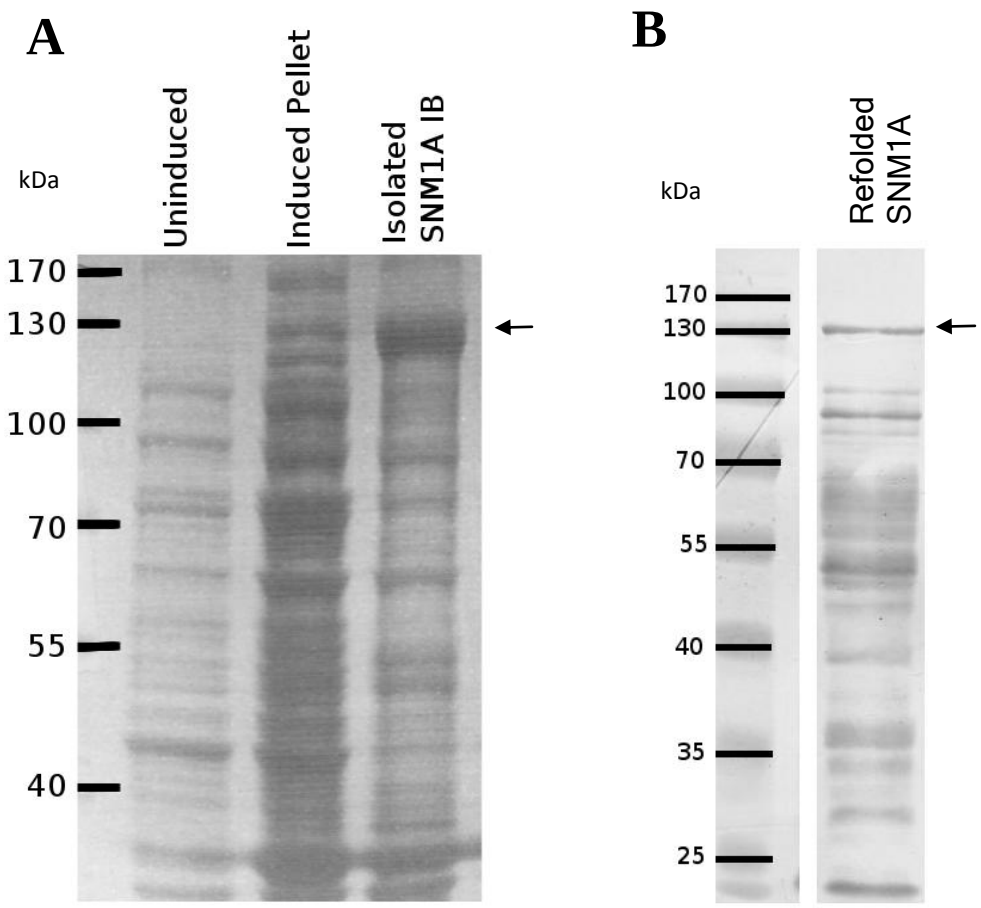


Figure 7: Expression and isolation of full-length SNM1A inclusion bodies in *E. coli.*

SNM1A was expressed as insoluble inclusion bodies in bacterial cells. These could be isolated and refolded. However, SNM1A inclusion bodies were subject to proteolytic degradation *in situ*. A) SDS-PAGE of uninduced whole cell sample, induced insoluble sample, and isolated inclusion bodies of full-length SNM1A after Triton X-100 washes on 10% SDS-PAGE gel, stained with Coomassie stain. B) Western blot of refolded inclusion bodies of SNM1A, detected with anti-His antibody. Arrows indicate the expected size of full-length SNM1A.

3.1.2 Refolding SNM1A from solubilized inclusion body preparations

SNM1A inclusion body material was solubilized by resuspension in a buffer containing 7M guanidine and 100mM TCEP. Solubilized inclusion bodies were used for screening for conditions able to permit refolding using the iFOLD Protein Refolding System 3 (Novagen). Refolding was inversely correlated to amount of aggregation detected at OD_{340}. Ninety-six different refolding conditions were tested for protein aggregation, as determined by spectrophotometric measurements at 340 nm, following rapid dilution of solubilized inclusion bodies into each condition. Table 1 summarizes the conditions which showed the least amount of inclusion body aggregation during refolding. 50mM TAPS, pH 8.5, FoldACE 4, 2mM oxidized glutathione and 8mM reduced glutathione had the lowest A_{340} measurement, suggesting that these condition best facilitate the refolding of SNM1A from solubilized inclusion bodies.

Initial conditions that yielded the lowest A_{340} measurements were retested in eighteen refined conditions at pH 7-8.5, in the presence or absence of FoldACE 4, with or without one of four reducing conditions as shown in Table 2. Inclusion body refolded in conditions buffered at pH 7.5 or 8.5 containing FoldACE4 and a reducing condition aside from 5mM TCEP showed increased solubility. A pH of 8.5 was observed to yield the optimal amount of refolded SNM1A.

Table 1: Comparison of folding conditions yielded lowest aggregation for SNM1A

Refolding condition	A_{340}
pH 7	0.366
pH 7 + 1 mM TCEP + FoldACE4	0.378
pH7.5	0.350
pH 7.5 + 6 mM GSH/4 mM GSSG + FoldACE5	0.350
pH 7.5 + 8 mM GSH/2 mM GSSG + FoldACE5	0.368
pH 8 + FoldACE4	0.374
pH 8.5 +8 mM GSH/2 mM GSSG + FoldACE4	0.310

GSH = glutathione (reduced), GSSG = glutathione (oxidized)

Table 2: Comparison of effects of refolding buffering components on SNM1A inclusion body aggregation.

Buffering condition	Buffer only	+FoldACE 4	1mM TCEP	5mM TCEP	6mM GSH 4mM GSSG	8mM GSH 2mM GSSG
MOPS, pH 7	0.875	0.248	0.368	0.781	0.630	0.731
HEPES, pH 7.5	0.857	0.055	0.087	0.333	0.161	0.114
TAPS, pH 8.5	0.971	0.062	0.041	0.388	0.016	0.042

Since 50mM TAPS, pH 8.5, FoldACE 4, 2mM oxidized glutathione and 8mM reduced glutathione produced the lowest A_{340} measurement, this refolding buffer condition was used to refold both full-length SNM1A as well as the beta-CASP domain.

3.1.3 Exonuclease activity assay for SNM1A

To verify that inclusion bodies of SNM1A were able to be refolded into a correctly folded state, a functional assay based on nuclease activity was performed, such as the one in Figure 8. 5' exonuclease activity was tested on inclusion bodies refolded in a range of buffers, with and without 1mM TCEP. The extent of nucleolytic degradation of the single 5' ^{32}P label was assessed.

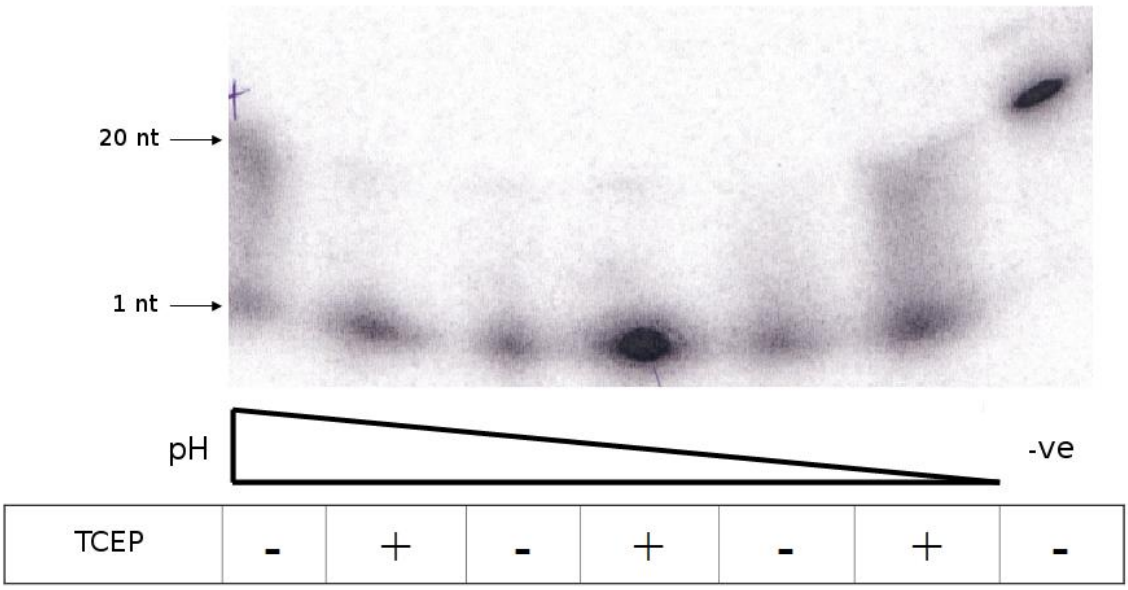


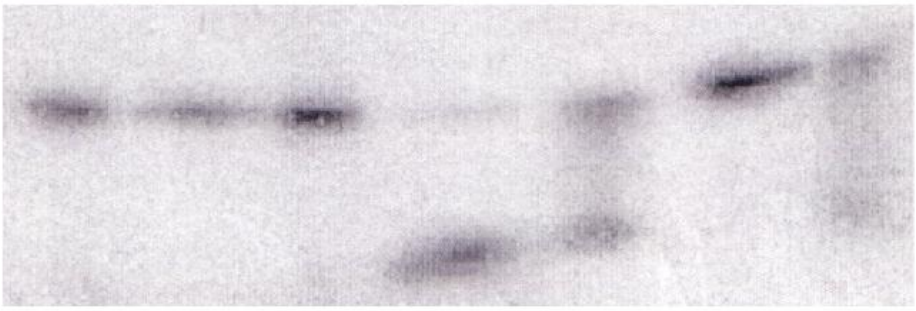
Figure 8: Functional exonuclease screen for active refolded SNM1A.

SNM1A inclusion bodies were refolded in the indicated conditions, and then tested for 5`exonuclease activity on 5' ^{32}P labeled ssDNA. Products were electrophoresed on 10% denaturing PAGE.

In order to ensure that the nuclease activity could be attributed only to properly refolded SNM1A and not a co-purified contaminating nuclease, the refolding components or nuclease reaction components were systematically omitted, as shown in Figure 9. Reactions lacking refolded protein, reaction buffer, or key refolding components with zinc showed no nuclease activity. Although reactions containing refolded protein and buffer with all components showed the most activity, degrading about 90% of 5' labeled DNA(see 4th lane, Figure 9), activity was detected in conditions either without a reducing agent (see 5th lane, Figure 9), or with only reaction buffer (see last lane, Figure 9). Under these conditions, however, nuclease activity was reduced to approximately 50% efficiency.

To confirm that observed 5' exonuclease activity was due to SNM1A and not another contaminating nuclease, SNM1A^{D736A} was prepared in the same manner and was tested for nuclease activity. Mutation of the aspartic acid of the HxHxDH signature motif within the metallo beta-lactamase domain abolished 5' exonuclease activity, as show in Figure 10. This result also demonstrates that the observed 5' exonuclease activity was phosphate-dependent, a property specific to the beta-CASP nucleases.

These results indicate that nuclease activity obtained from refolded SNM1A inclusion body material was indeed from the SNM1A and not from any other components of the refolding or reaction buffers. Moreover, the activity of the refolded protein was specific to refolded SNM1A and not any other contaminating nuclease.



IB	+		+	+	+	+	+
Refolding buffer	+	+	+	+	+		
Reducing agent	+	+	+	+			
Zinc	+	+		+	+	+	
Reaction buffer		+	+	+	+	+	+

Figure 9: Control assay of nuclease activity from refolded inclusion body components in SNM1A nuclease reactions.

SNM1A inclusion bodies were refolded in the indicated conditions, and then tested for activity on 5' 32P labeled ssDNA. Products were electrophoresed on 10% denaturing PAGE.

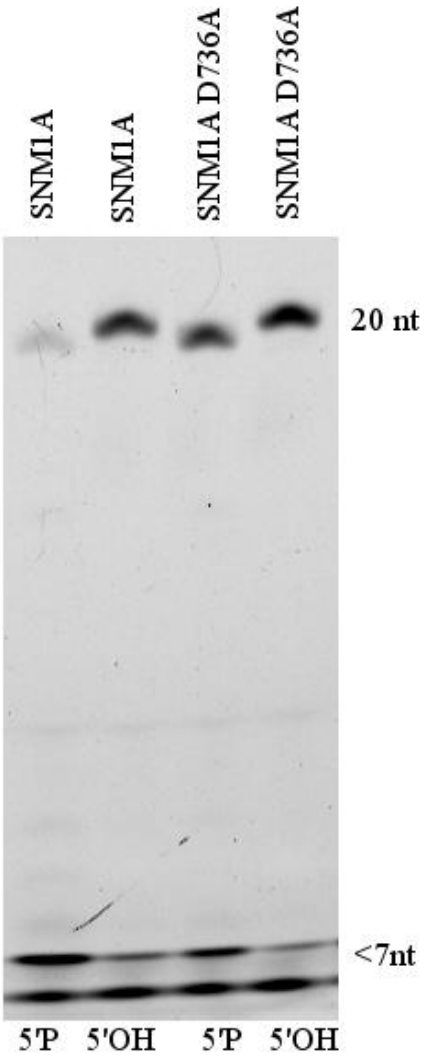


Figure 10: 5' Phosphate-dependent exonuclease activity of SNM1A and SNM1A ^{D736A}.

Refolded SNM1A and SNM1A ^{D736A} inclusion bodies were incubated with 3' fluorescently labeled ssDNA with either a 5' phosphate or 5' hydroxyl to determine phosphate-dependent activity. Products were electrophoresed on 20% denaturing PAGE.

3.2 Results for the beta-CASP domain of SNM1A

3.2.1 Expression of the beta-CASP domain

The beta-CASP domain of SNM1A could be stably expressed in *E. coli*, but once again, this protein was highly insoluble. The beta-CASP domain was expressed at higher levels compared to full-length SNM1A when comparing equivalent amounts of cells (Figure 7 vs. Figure 11). Inclusion body isolation using 1% Triton X-100 washes could enrich the amount of beta-CASP inclusion bodies to 85% purity.

3.2.2 Optimization of conditions for beta-CASP domain inclusion body isolation

For isolation of beta-CASP inclusion bodies, the wash conditions were first optimized. In this analysis, the supernatant of each wash was analyzed on SDS-PAGE to determine its efficacy at removing contaminating proteins from the beta-CASP inclusion bodies. Resuspension buffers lacking detergents or denaturants were not effective at removing contaminants, as seen in Figure 12, but did not reduced the amount of the beta-CASP inclusion bodies recovered. Wash buffers containing non-ionic detergent, 1% Triton X-100, was effective at removing some contaminating proteins, while at the same time leaving the beta-CASP inclusion bodies intact (see 2nd lane, Figure 12). Wash buffers containing 5M urea was observed to solubilize contaminating proteins that could not be fully removed with 1% Triton X-100, but under these conditions some beta-CASP inclusion bodies were also solubilized and therefore lost (see 3rd lane, Figure 12). Although this lowered the final yield of protein, it increased the purity of the recovered

inclusion bodies, removing protein either trapped in the inclusion bodies or insoluble protein. The final inclusion body isolation protocol included all three washes since this produced the greatest inclusion body purity and did not reduce the overall yield of recovered protein.

The force at which the inclusion bodies were centrifuged was also varied among different protocols. As reported by others, determining the optimal force increased the amount of inclusion bodies that are pelleted with the least amount of contaminating protein that is co-pelleted. A contaminant was observed migrating just below the beta-CASP domain on a SDS-PAGE gel in Figure 13. In 1% Triton, this contaminant was insoluble and co-pelleted with the beta-CASP domain. After the removal of detergent, however, this contaminant still co-pelleted with the beta-CASP domain at 16,000xg and 39,000xg. Thus, the optimal force for the isolation of the highest purity of the beta-CASP domain was determined to be 8,000xg.

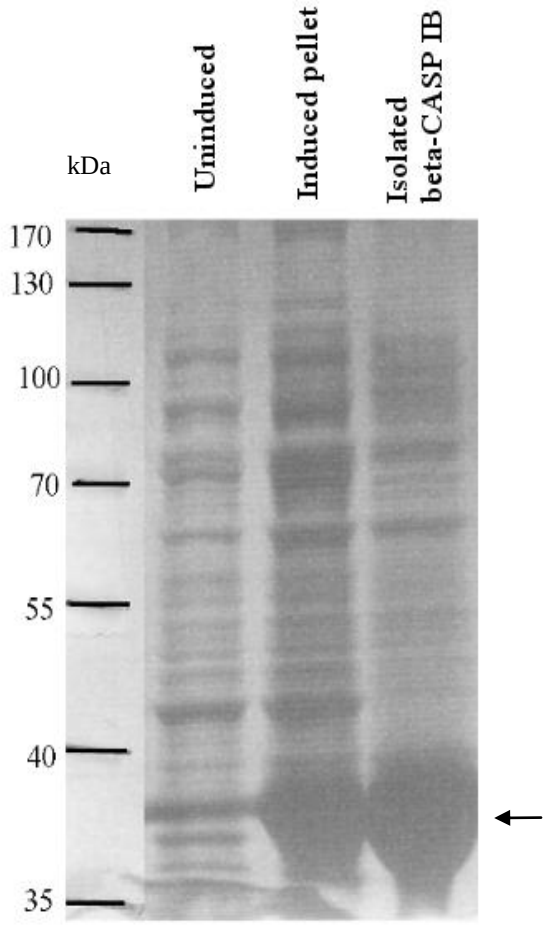


Figure 11: Expression and isolation of beta-CASP domain inclusion bodies in *E. coli.*

The beta-CASP domain of SNM1A was expressed as insoluble inclusion bodies in bacterial cells. A) SDS-PAGE of uninduced whole cell sample, induced insoluble sample, and isolated inclusion bodies of beta-CASP after Triton X-100 washes on 10% SDS-PAGE gel, stained with Coomassie stain. Arrow indicates the expected size of the beta-CASP domain.

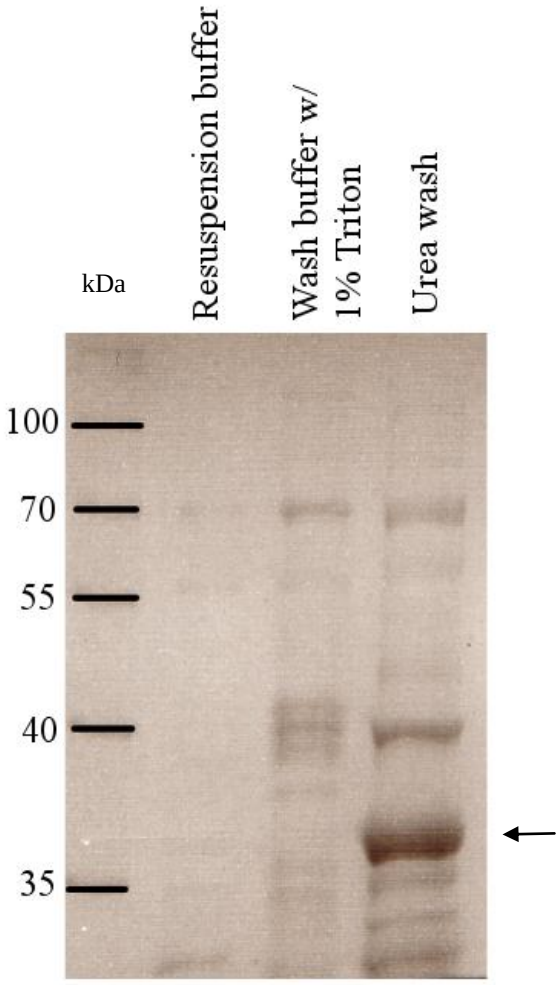


Figure 12: Washes for beta-CASP inclusion body isolation.

Beta-CASP inclusion bodies were resuspended in the indicated buffers then centrifuged to pellet the inclusion bodies. The supernatant was electrophoresed on 12% SDS PAGE. Gel was stained with Coomassie stain. Arrow indicates the expected size of the beta-CASP domain.

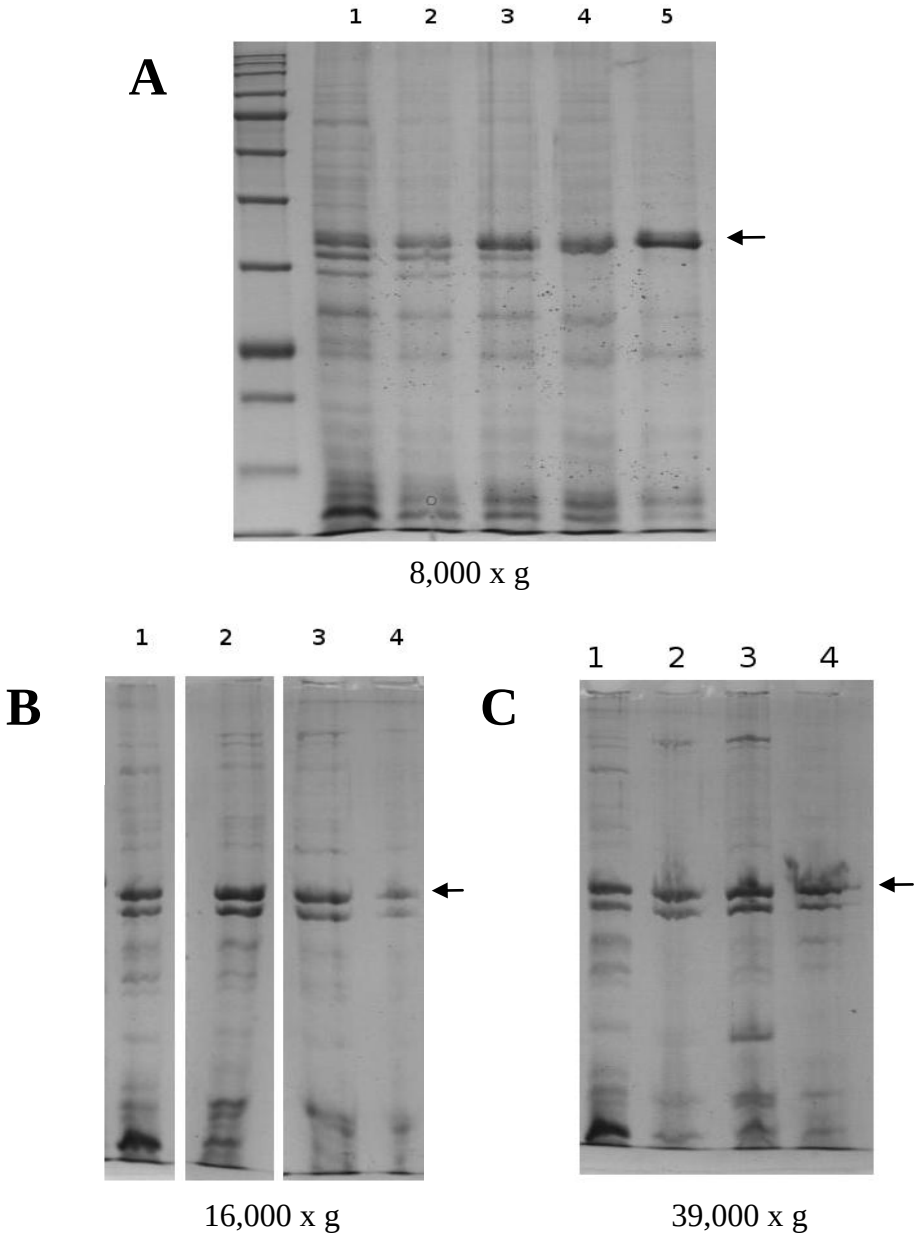


Figure 13: Centrifuge force optimization for beta-CASP inclusion body isolation.

12% SDS-PAGE gels of equivalent micrograms of solubilized inclusion bodies after each inclusion body wash step. Gels were stained with Blue Silver Stain. A) Isolation of inclusion bodies at 8,000 x g. B) Isolation of inclusion bodies at 16,000 x g. C) Isolation of inclusion bodies at 39,000 x g. The numbers above the gels correspond to the following washes: 1) 1% Triton X-100, 2)1% Triton X-100, 3) Resuspension buffer wash, 4) Second resuspension buffer wash. Arrows indicate the expected size of the beta-CASP domain.

3.2.3 Solubilization of the beta-CASP domain

The solubilization of the beta-CASP domain expressed as inclusion bodies was assessed at a pH range of 1 to 12. The inclusion body material was found to be insoluble at acidic and neutral pHs, but could be effectively be solubilized in basic conditions, with peak solubilization occurring at pH 11 (see Figure 14).

The solubilization of the beta-CASP domain inclusion bodies in SDS was also assessed from a range of 0 to 2% SDS. Solubilization of the inclusion bodies was significant at 0.5% SDS with > 50% of the inclusion bodies solubilized under these conditions. These inclusion bodies were >90% solubilized at 1% SDS (see Figure 15)

The efficacy of methods for inclusion body solubilization of the beta-CASP domain was compared using equivalent amounts of inclusion bodies. Solubilization of the inclusion body material was found to be almost 4-fold higher using guanidine as a denaturant compared to inclusion body solubilization using either SDS or alkaline conditions (see Figure 16)

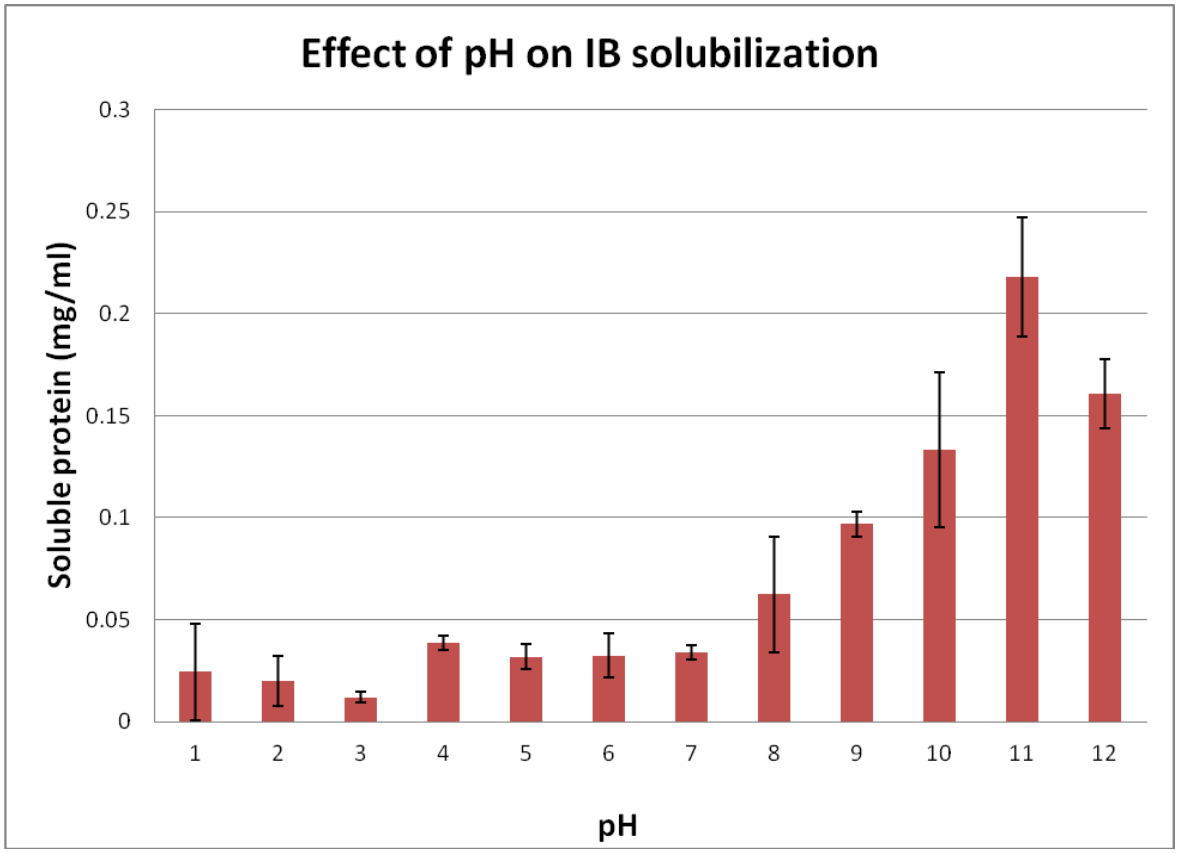


Figure 14: Comparison of recovery of soluble beta-CASP domain inclusion bodies buffered at various pHs

Inclusion body was resuspended in 50mM Tris pH 1 to 12. The amount of soluble protein after pelleting insoluble inclusion bodies was determined using the Bradford assay. Error bars represent the standard deviation from triplicate experiments.

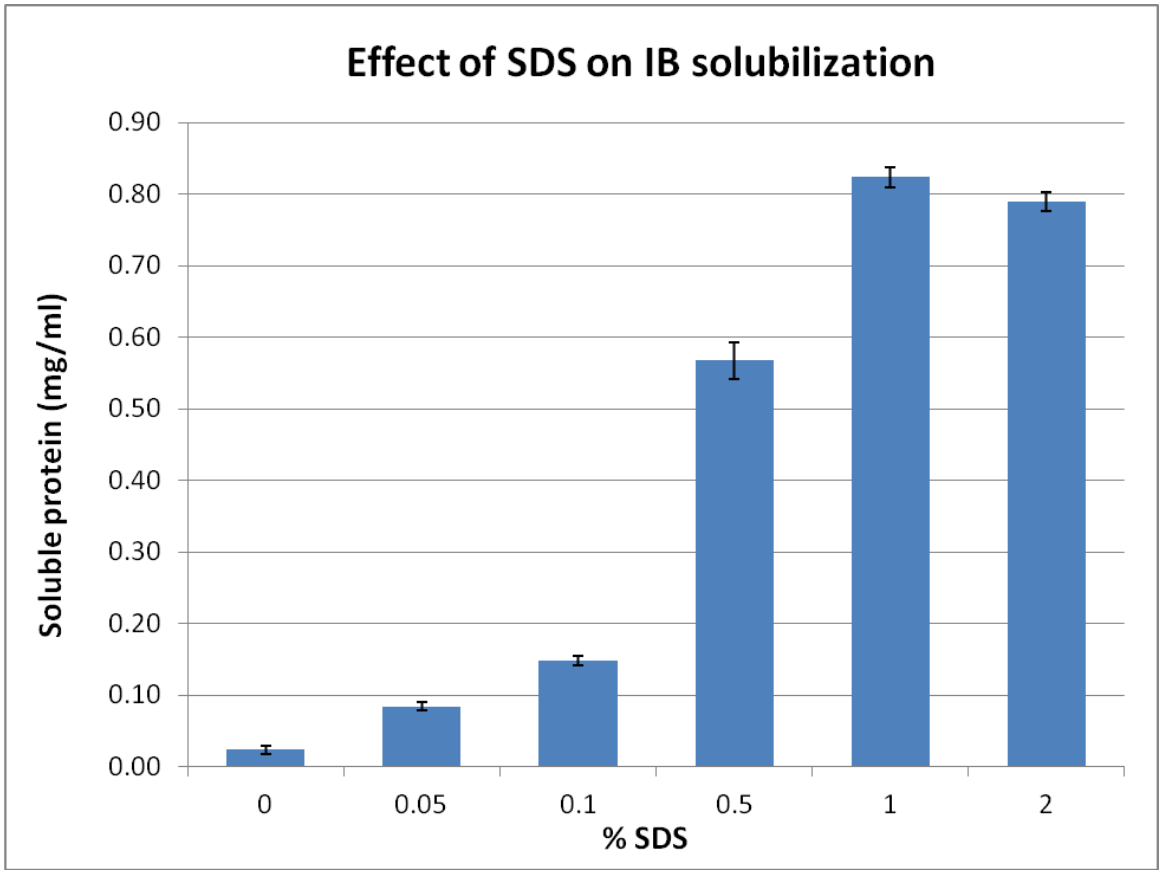


Figure 15: Comparison of recovery of soluble beta-CASP domain inclusion bodies resuspended in various concentrations of SDS.

Inclusion body was resuspended in 0.05% to 2% SDS. The amount of soluble protein after pelleting insoluble inclusion body was determined using the Bradford assay. Error bars represent the standard deviation from triplicate measurements.

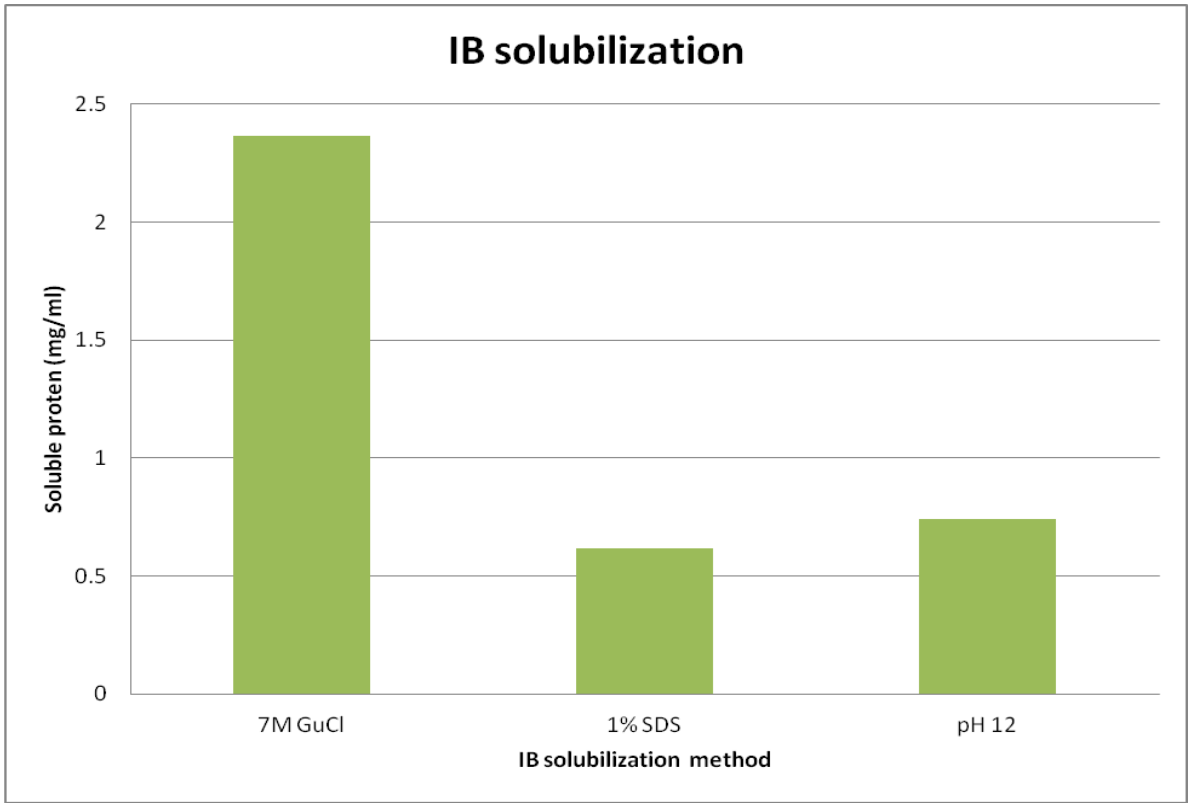


Figure 16: Initial determination of recovery of soluble beta-CASP domain inclusion bodies with different solubilization conditions.

Equivalent amounts of inclusion bodies were resuspended in the indicated solubilization buffers. The amount of soluble protein after pelleting insoluble inclusion bodies was determined using the Bradford assay.

3.2.4 Refolding of beta-CASP domain inclusion bodies

The beta-CASP domain of SNM1A could be expressed, isolated, and solubilized, yielding a concentration of solubilized inclusion bodies of 1mg/mL, or 24 μM. However, refolding these inclusion bodies consistently lead to aggregation, possibly indicating the refolding conditions required for refolding the beta-CASP domain differed for that required for full-length SNM1A. Refolding conditions for the beta-CASP domain was screened using the same conditions outlined in Table 2, but significant protein aggregation was observed in all conditions tested. Figure 17 indicates that higher concentrations of solubilized beta-CASP inclusion bodies resulted in aggregation as measured by the absorbance at 340nm. Aggregation was decreased with decreasing concentrations of solubilized inclusion bodies. At very low concentrations of solubilized inclusion bodies however, the amount of recovered refolded protein was beyond the lower limits of the detection range of the Bradford assay. Therefore, the concentration of solubilized inclusion bodies used for refolding was set to1μM, despite a small amount of aggregation that occurred during the refolding process. The final refolded beta-CASP domain from unpurified inclusion bodies is shown in Figure 18. This preparation can yield up to 3mg of beta-CASP domain of 85% purity.

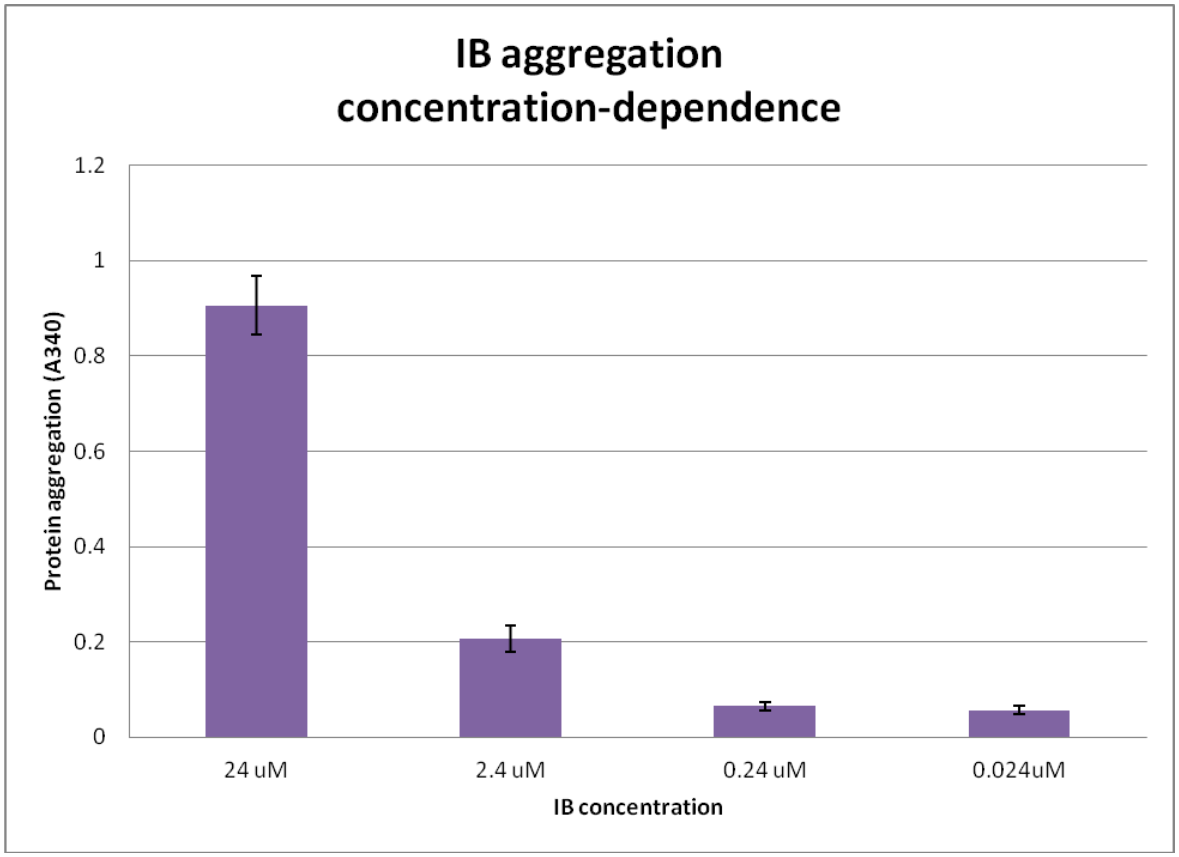


Figure 17: Effect of solubilized inclusion body concentration on beta-CASP domain aggregation

Inclusion body of indicated concentrations was refolded via rapid dilution. Solutions were tested for protein aggregation at 340 nm. Error bars represent the standard deviation of triplicate experiments.

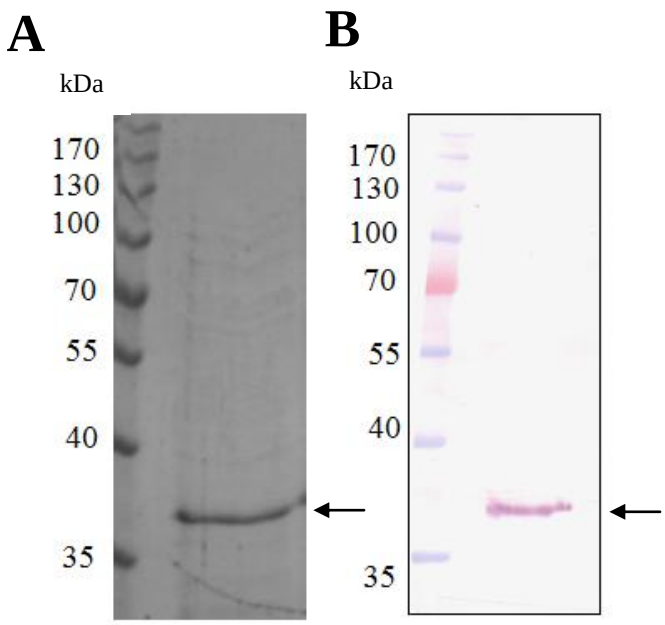


Figure 18: Refolded beta-CASP protein.

A) 12% SDS-PAGE of unpurified beta-CASP protein after refolding stained with Blue Silver stain. B) Western blot of refolded inclusion bodies of the beta-CASP domain, detected with anti-His antibody. The arrows indicate the expected size of the beta-CASP domain.

3.2.5 Purification of the beta-CASP domain under denaturing conditions

To increase the purity of the beta-CASP domain, further purification was required. Solubilized inclusion bodies were further purified using IMAC Ni-NTA agarose beads under denaturing conditions, as shown in Figure 19. As expected, a majority of the contaminants did not bind to the nickel resin. Further contaminants were removed during subsequent wash steps. The beta-CASP domain was first eluted with denaturing buffer containing 500mM imidazole, resulting in a significant yield of eluted inclusion bodies of >95% purity. A second subsequent elution using 50mM Tris (pH 4.5) was tested, which removed protein still bound the resin. Elution with Tris, pH 4.5 was volume-dependent, since the pH 4.5 is beyond the buffering capacity of Tris. The amount of residual beta-CASP domain still bound to the nickel resin after elution was then determined by boiling the Ni-NTA agarose beads in SDS loading buffer.

The beta-CASP inclusion bodies eluted with imidazole or pH 4.5 were refolded. A significant yield of refolded protein was obtained with inclusion bodies eluted with 500mM imidazole, whereas little protein refolded with inclusion bodies eluted at pH 4.5 was recovered (see Figure 20). The amount of trace contaminants in the final refolded protein was minimized in comparison to the unpurified refolded protein preparation shown in Figure 18.

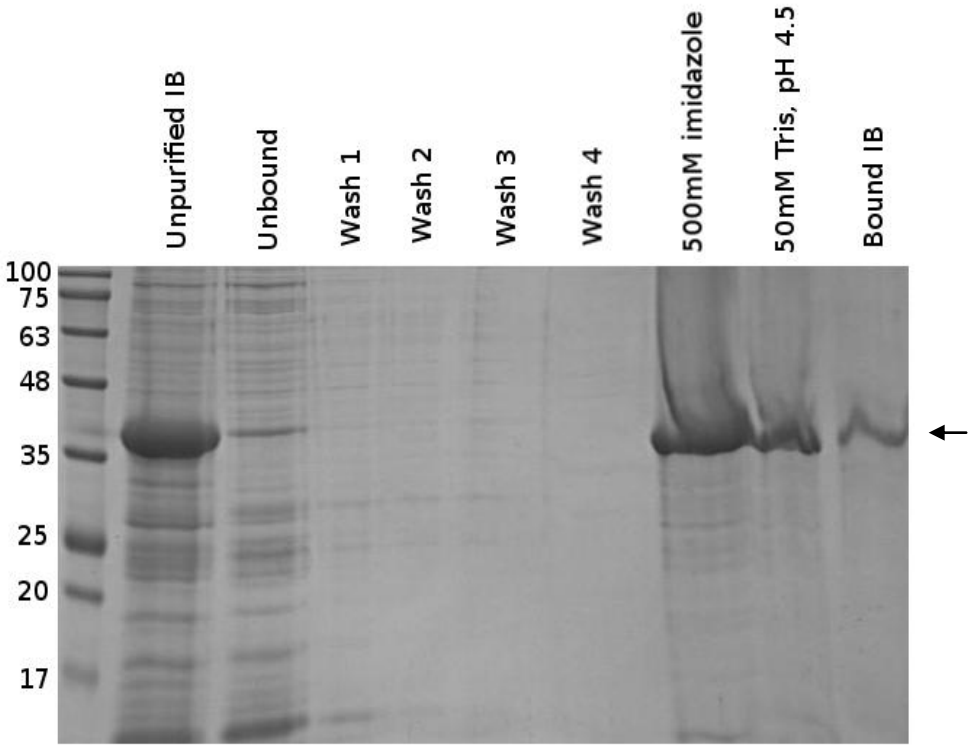


Figure 19: Analysis of fractions of the beta-CASP domain purified with Ni-NTA agarose beads under denaturing conditions.

Unpurified beta-CASP domain inclusion bodies were diluted and incubated with IMAC Ni-NTA agarose resin under denaturing conditions. Proteins which did not bind to the IMAC resin were removed and washed away. The purified beta-CASP protein was eluted successively with high concentrations of imidazole or low pH conditions. Bound inclusion bodies were protein released from the resin after the elution step by boiling in SDS loading buffer. Fractions were analyzed on 12% SDS-PAGE stained with Blue Silver stain. Washes 1 to 3 are comprised of 8M urea, 50mM Tris, pH 8 and 10mM DTT. Wash 4 is comprised of 6M guanidine, 50mM Tris, pH 8, and 300mM NaCl. The arrow indicates the expected size of the beta-CASP domain.

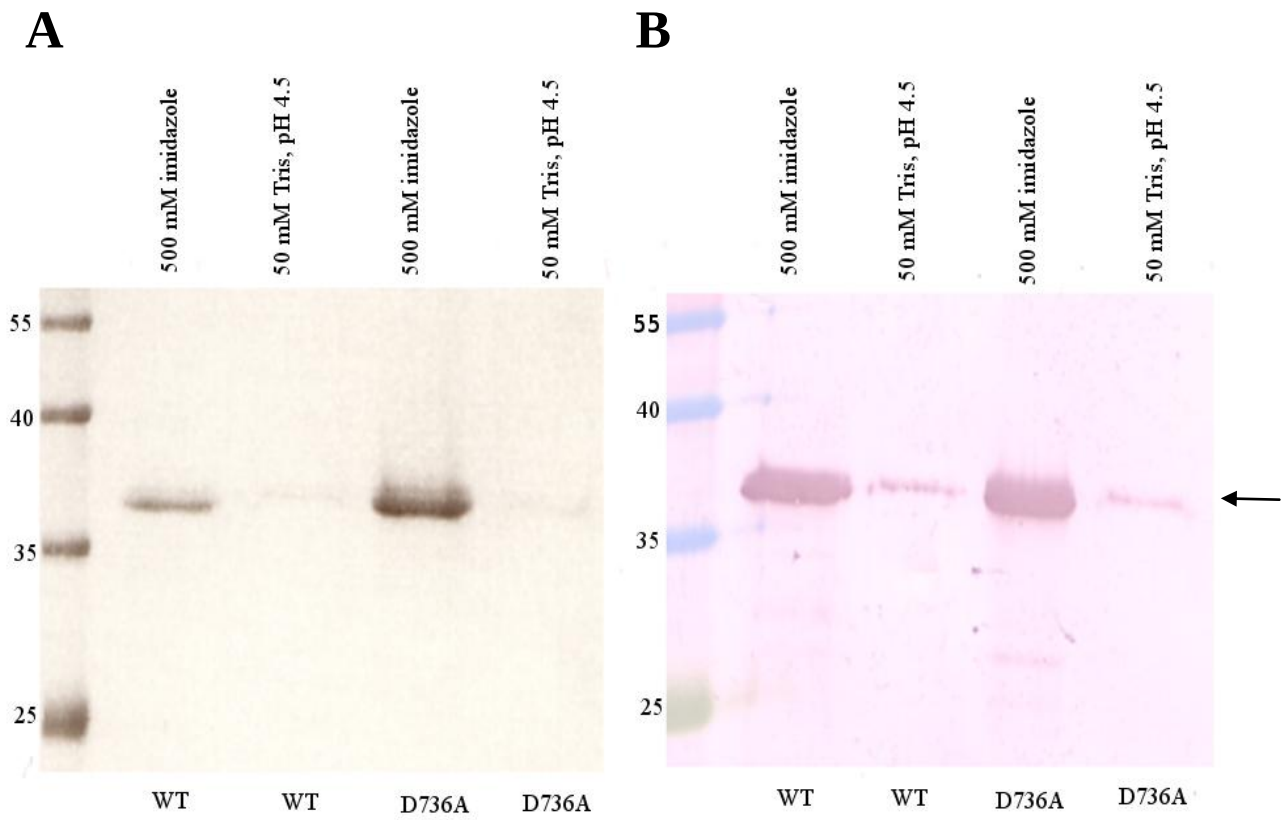


Figure 20: Effect of elution conditions on beta-CASP domain recovery and purity pH 4.5.

The beta-CASP domain purified under denaturing conditions was eluted with either high concentrations of imidazole or low pH conditions. The eluted protein was refolded and electrophoresed on a 12% SDS-PAGE gel. A) SDS-PAGE eluted inclusion bodies, refolded and concentrated. 12% SDS-PAGE gel was stained with Blue Silver stain. B) Western blot of refolded inclusion bodies of SNM1A, detected with anti-His antibody.

3.2.6 Nuclease activity of the beta-CASP domain

Refolding unpurified beta-CASP inclusion bodies yielded active protein, as shown in Figure 21. This refolded protein had 5' exonuclease activity on both ssDNA and dsDNA, seeming to have higher activity on ssDNA (see A vs. C, Figure 21). A mutation to the conserved HxHxDH motif abolished the 5' exonuclease activity of beta-CASP^{D736A}. This nuclease activity was also phosphate-dependent (see A vs. B), indicating that the 5' exonuclease activity was beta-CASP specific. However, product formation was observed with the both beta-CASP domain and beta-CASP ^{D736A} on both 5' phosphorylated and unphosphorylated dsDNA indicating a possible 3' exonuclease contaminant, necessitating a major purification step.

Nuclease activity assays were used to assess the activity of the refolded purified beta-CASP protein, eluted with either imidazole or pH 4.5. For 5’ exonuclease activity, only beta-CASP inclusion bodies eluted at pH 4.5 yielded active protein (see Figure 22). Inclusion bodies eluted with imidazole did not show any exonuclease activity. None of the refolded inclusion bodies exhibited 3’ exonuclease activity. Therefore, this 3’ contaminant was completely removed during IMAC purification.

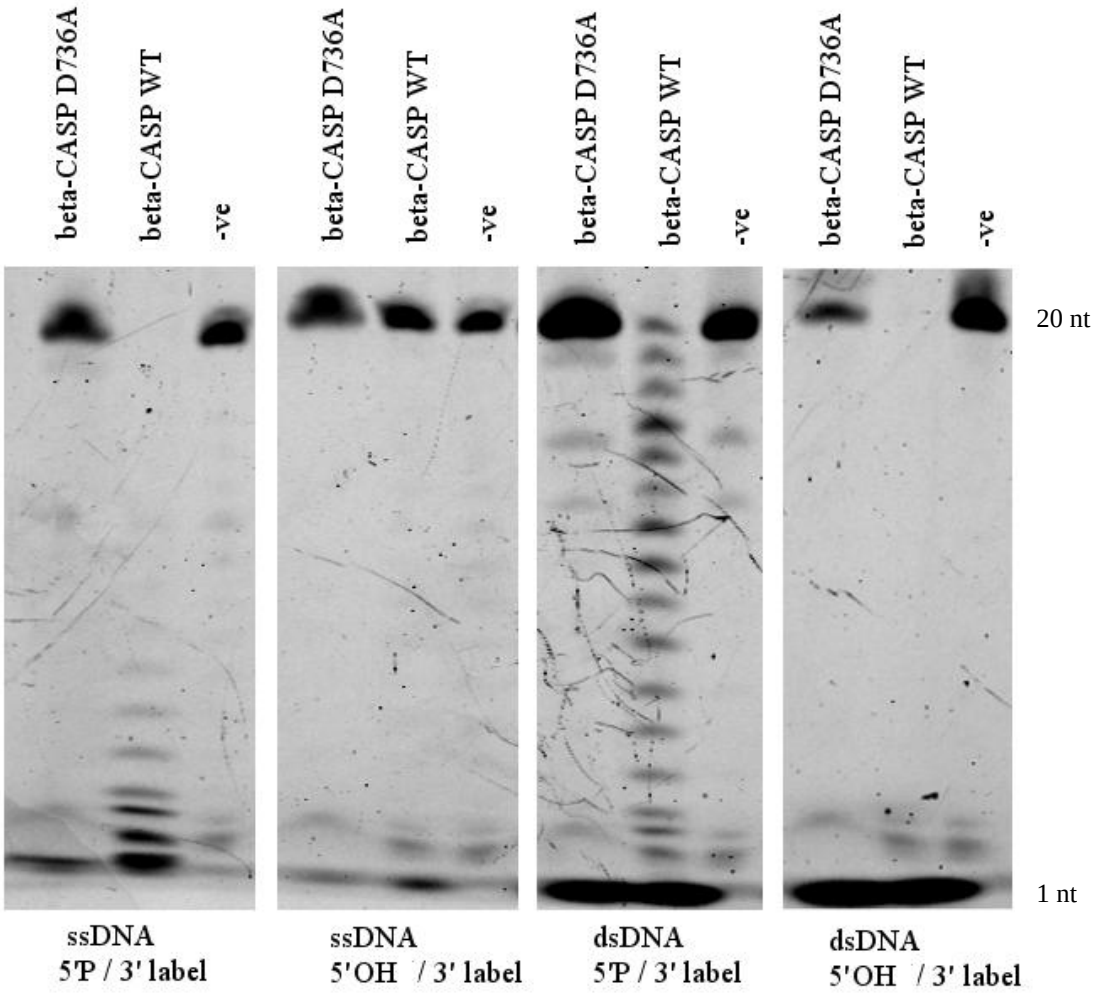


Figure 21: Analysis of 5' phosphate dependent exonuclease activity of the refolded unpurified beta-CASP domain.

Refolded unpurified beta-CASP domain and beta-CASP ^{D736A} were incubated with 3' fluorescently labeled ssDNA or dsDNA with either a 5' phosphate or 5' hydroxyl to determine phosphate-dependent activity. Products were electrophoresed on 20% denaturing PAGE.

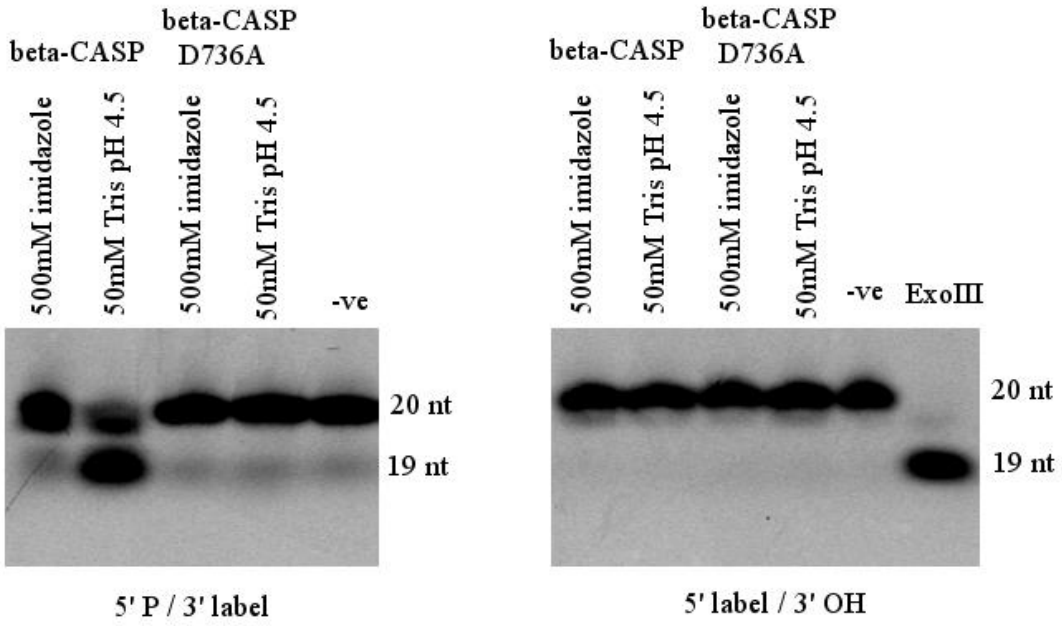


Figure 22: Effect of elution conditions on 5' phosphate-dependent exonuclease activity of the purified beta-CASP domain.

Purified beta-CASP domain and beta-CASP ^{D736A} inclusion bodies eluted with imidazole or at pH 4.5 were refolded and tested for nuclease activity. They were incubated with dsDNA with either a 5' phosphate and 3' label or 5' label to determine 5' or 3' exonuclease activity, respectively. Substrates had phosphorothioate bond after second nucleotide to observe only the first hydrolysis reaction. Products were electrophoresed on 20% denaturing PAGE.

3.2.7 Enzyme kinetics of the beta-CASP domain

Time course assays were done to determine the rate of substrate hydrolysis over time for progress curves (see Figures 23 and 24). The concentration of beta-CASP protein was adjusted such that only 10% of the substrate was hydrolyzed at the 16 minute time point. Progress curves of were generated for initial velocity calculations. This was calculated to be 1.9 $x10^{-11}$ M/s for ssDNA and 1.5 $x10^{-11}$ M/s for dsDNA.

Using this initial velocity, the beta-CASP domain was incubated with increasing amounts of either ssDNA or dsDNA. At 5μM, the reaction rate for ssDNA showed saturation of enzyme to substrate, but dsDNA required twice as much DNA to observe this. Using non-linear regression analysis, kinetic parameters were calculated, as shown in the table in Figure 25.

The beta-CASP domain has increased affinity for ssDNA over dsDNA, as reflected by the K_m values, but can generate twice as much dsDNA product than ssDNA, as indicated by the k_{cat}. Overall, the catalytic efficiency, as determined using k_{cat}/K_m, was more than two-fold higher for ssDNA compared to dsDNA.

Overall, the beta-CASP domain of SNM1A is a 5'phosphate dependent exonuclease with increased activity on ssDNA over dsDNA.

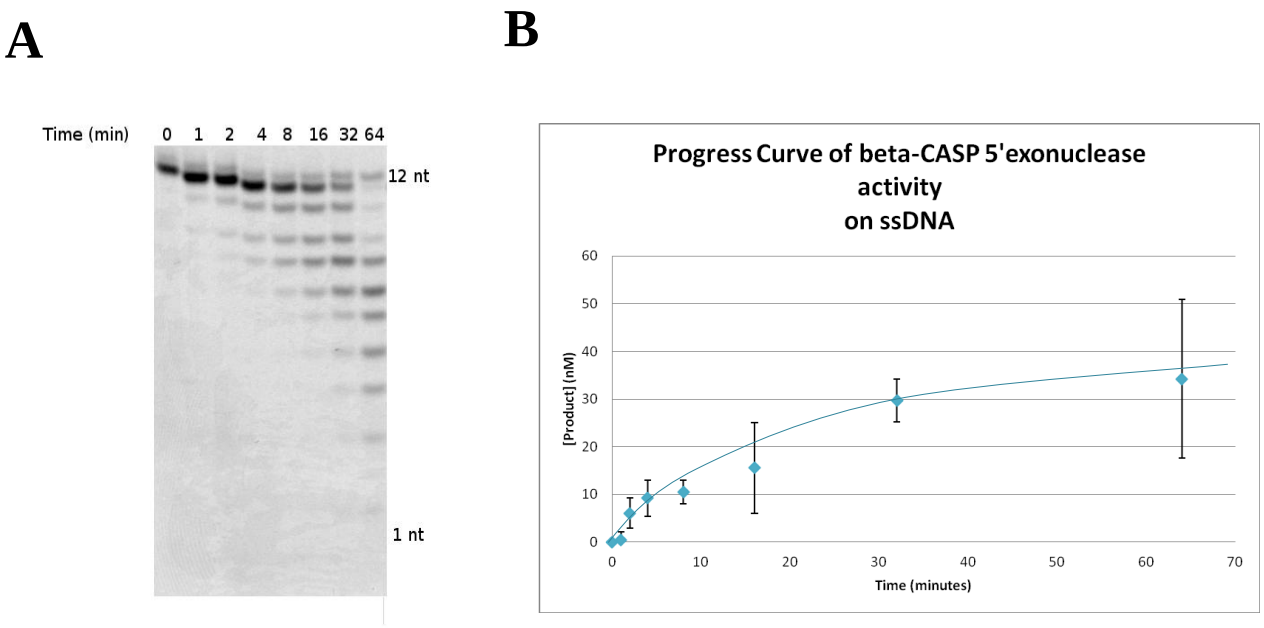


Figure 23: Progress curve of the beta-CASP domain 5' exonuclease activity on ssDNA.

A) Progress curve assay of 5nM beta-CASP protein on 200nM ssDNA. B) Graph of products of beta-CASP nuclease activity on ssDNA. Error bars represent the standard deviation from triplicate experiments.

A **B**

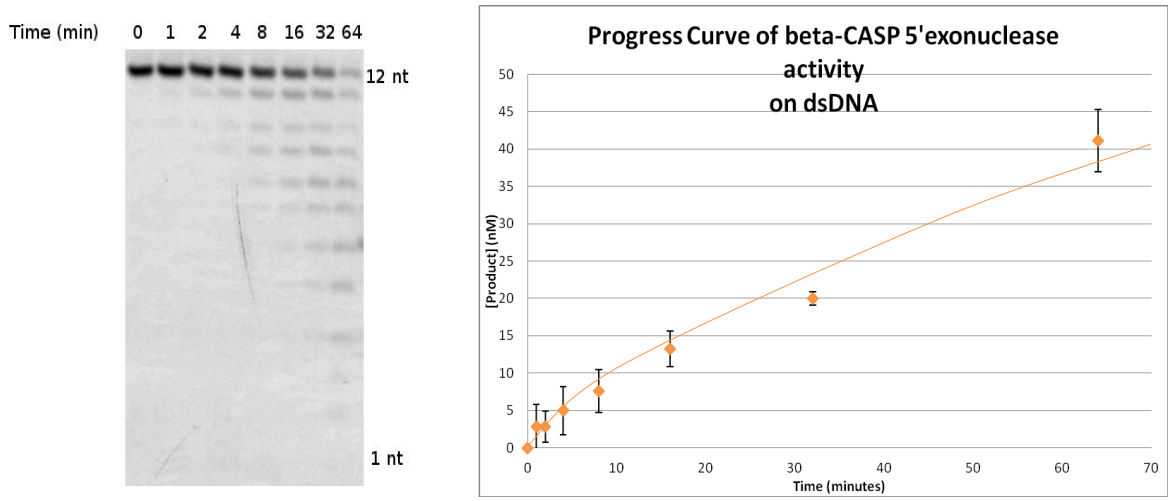


Figure 24: Progress curve of the beta-CASP domain 5' exonuclease activity on dsDNA.

A) Progress curve assay of 15nM beta-CASP protein on 200nM dsDNA. B) Graph of products of beta-CASP nuclease activity on dsDNA. Error bars represent the standard deviation from triplicate experiments.

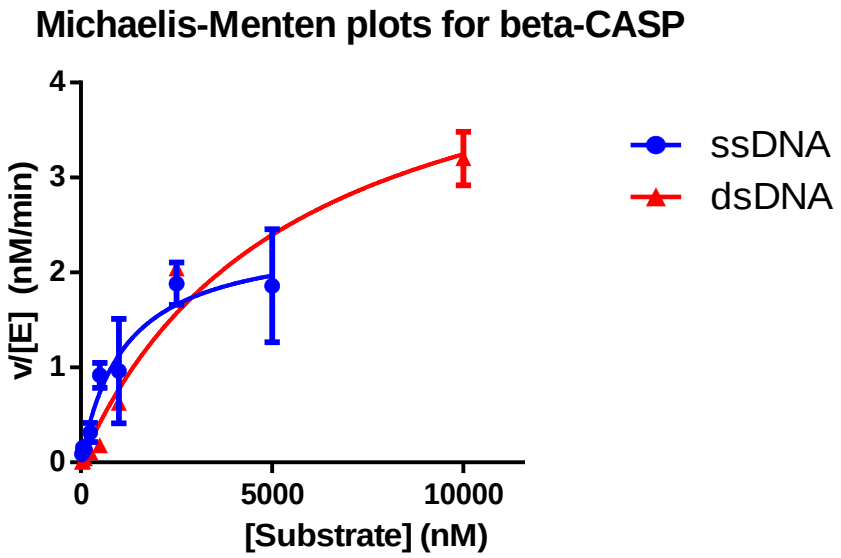


Figure 25: Michaelis-Menten plots for K_m and k_{cat} determination of the beta-CASP domain.

Purified refolded beta-CASP domain was incubated with 25nM up to 10µM ssDNA or dsDNA. The velocities of each reaction determined and plotted. Non-linear regression analysis and Michaelis-Menten curve fitting was performed with GraphPad. Error bars represent the standard error from triplicate experiments.

Table 3: Comparison of enzyme kinetic values for refolded beta-CASP domain.

Substrate	K_m (µM)	k_{cat} (s^{-1})	k_{cat}/K_m ($s^{-1} M^{-1}$)
ssDNA	1.1 ± 0.3	0.04 ± 0.01	3.5×10^4
dsDNA	5.5 ± 1.4	0.08 ± 0.01	1.5×10^4

3.3 Complementation of *pso2* null strains with SNM1A variants

To determine if the beta-CASP domain of SNM1A is sufficient to repair ICL adducts *in vivo,* complementation assays were used. The beta-CASP domain and beta-CASP ^{D736A} were transformed into *pso2* null *S. cerevisiae.* Full-length SNM1A and SNM1A^{D736A} were also analyzed under similar conditions for comparison. After treatment with 1mM cisplatin, the cells were plated onto 2% galactose containing agar plates to induce heterologous expression of the various protein constructs.

As shown in Figure 26, the beta-CASP domain of SNM1A shows similar sensitivity to cisplatin as beta-CASP ^{D736A} and SNM1A ^{D736A}, with only a slightly higher level of sensitivity to ICL damage compared to the empty vector. This isolated beta-CASP domain on its own displayed a ten-fold increase in sensitivity to ICL damage as compared to yeast expressing full-length wild-type SNM1A and nearly the same increase in sensitivity compared to the positive control, Pso2.

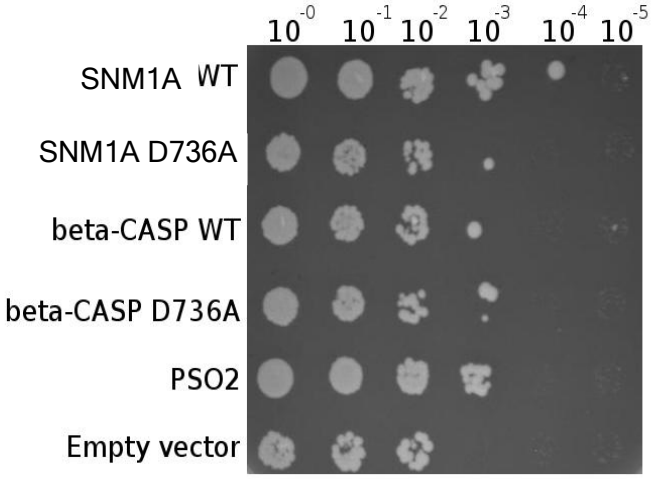


Figure 26: Cisplatin sensitivity assay of *pso2* null *S. cerevisiae* complemented with SNM1A and the beta-CASP domain.

SNM1A, beta-CASP domain, and their respective mutants were heterologous expressed after exposure to 1mM cisplatin. A dilution series of the ICL induced cells were plated to determine the extent of their sensitivity to cisplatin. Untreated cells were also plated to verify the same number of complemented cells was used. Results are representative of those observed for four independent experiments.

.

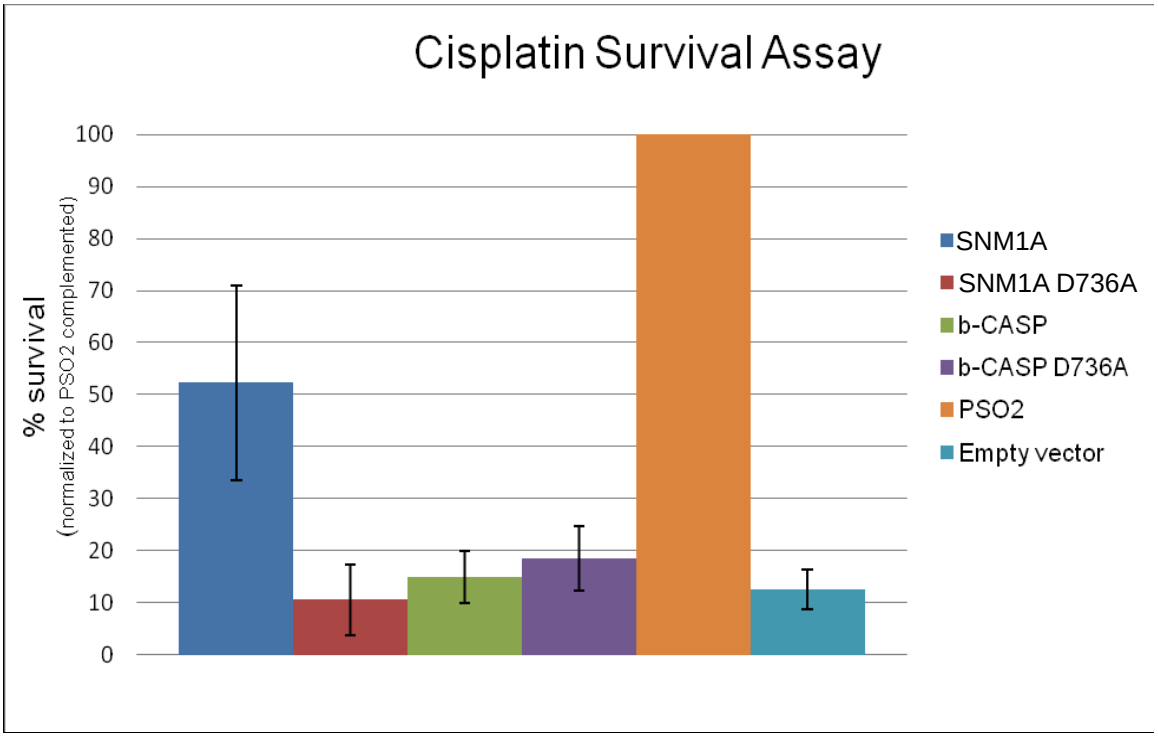


Figure 27: Cisplatin survival assay of *pso2* null *S. cerevisiae* complemented with SNM1A and the beta-CASP domain.

SNM1A, beta-CASP domain, and their respective mutants were heterologously expressed after exposure to 1mM cisplatin to determine the extent of their survival to cisplatin. Untreated cells were also plated to ensure no significant variation in cell survival. % survival was normalized to the survival of cells complemented with PSO2. Error bars represent the standard deviation from triplicate experiments

To quantify survival in response to exposure to ICL damage, resulting numbers of CFUs were calculated and normalized to CFUs of Pso2 expressing yeast, as shown in Figure 27. In the presence of 1mM cisplatin, only yeast complemented with Pso2 or wild-type SNM1A showed survival after cisplatin treatment. Wild-type beta-CASP domain had about a ten-fold reduction in survival under the same conditions, exhibiting the same cisplatin survival as the full-length and beta-CASP domain mutants. Complementation of the beta-CASP domain or any of the mutants did not increase survival after treatment with cisplatin.

Overall these results suggest that although the beta-CASP domain of SNM1A has nuclease activity *in vitro*, this activity is not sufficient *in vivo* to rescue *pso2* null defects following exposure to ICL damaging agent, cisplatin.

CHAPTER 4: DISCUSSION

4.1 Discussion of SNM1A

4.1.1Expression of human SNM1A in *S. cerevisiae*

In order to fully understand the function of the isolated beta-CASP domain, full-length SNM1A was required for comparison of activity to the beta-CASP domain. Heterologous expression of human SNM1A has been previously shown to rescue defects of the yeast ICL repair homologue *pso2* in *S. cerevisiae* (Hazrati et al, 2008). Prior biochemical characterization of human SNM1A employed protein expressed and purified from *S. cerevisiae* (Hejna et al, 2007). Therefore, expression of full-length SNM1A in budding yeast appeared to be a reasonable avenue to pursue for obtaining human SNM1A for subsequent biochemical studies.

Unfortunately, under conditions reported here, no significant induction of the full-length SNM1A protein was detected in an induction time point assay over 24 hours (data not shown). It was not clear if induction in yeast is generally poor or if this is construct-specific since the construct tested in this work varied slightly from that which was previously reported. Large scale purification was attempted to determine if expressed protein containing a hexa-histidine tag could be recovered using Ni-NTA IMAC chromatography, but no full-length protein was observed in any fraction from the purified protein. Elution with high concentrations of imidazole recovered a smaller fragment of 37

kDa, which was not detected using western blotting. This molecular weight was consistent with proteolytic degradation of the full-length SNM1A protein and may represent yielding the beta-CASP domain. Since the C-terminal hexa-histidine tag was not detected with western blotting, however, this interpretation may not be correct. Nonetheless, the serious proteolytic degradation of SNM1A made obtaining full-length SNM1A problematic. The yeast strain used in prior studies was protease-deficient and switching to this strain may have overcome proteolysis issues. Despite this, protein expression of human SNM1A in yeast was not likely to yield significant amounts of protein and was therefore abandoned.

4.1.2 Expression of human SNM1A in *E. coli*

It has been established that human SNM1A is expressed in an insoluble form in *E. coli* (Hejna, et al, 2007). It is possible that this insolubility may result from the presence of a long N-terminus, which has little predicted secondary structure. Although expression of insoluble protein leads to inclusion body formation, inclusion body expression can be beneficial. In particular, overexpression of protein which is toxic to cells is possible when expressed as inclusion bodies (Tsumoto et al, 2003). In addition, it has been observed that proteins expressed as inclusion bodies can be protected from endogenous proteases, generating final protein yields of generally higher homogeneity (Tsumoto et al, 2003). Finally, inclusion body material can be solubilized using strong denaturants, such as guanidine or urea to completely unfold the protein, and subsequently refolded back into

its native structure to yield active protein. That said, the refolding into solubility and recovery of activity are a highly empirical endeavor.

As shown in Figure 7B, the inclusion body fraction of full-length SNM1A was not protected from endogenous proteases. Research by Cubarsı´ et al (2001) showed that inclusion body material may be subject to proteolytic degradation. In the case of human SNM1A, this appears to be the case. In particular, expression at lower temperatures helped increase the amount of full-length protein recovered, as shown in Figure 7A, potentially by decreasing the activity of endogenous proteases. Despite limitations of proteolysis, relatively higher yields of SNM1A expression were achieved. This is significant to some degree since the level of recovery of SNM1A reported here has never been observed using other systems for expression, such as in yeast or insect cell lines (Hazrati et al, 2008; Hejna et al, 2007). Despite difficulty with proteolytic breakdown, given the yields of full-length protein, screening for refolding conditions was pursued with the hope of being able to further purify full-length SNM1A following successful refolding.

4.1.3 Refolding SNM1A from inclusion bodies

Finding conditions which permit protein is proper refolding from isolated inclusion bodies can be a labour intensive process. Buffering conditions, redox environment, and the presence of small molecule additives must be considered and

treated empirically. Commercially available kits facilitate this process by permitting wide sampling of a sparse matrix of conditions frequently observed to yield refolded protein. Using the iFOLD System 3 from Novagen, it was possible to systematically screen dozens conditions known for their ability to permit refolding of stable, active protein. In this assay, absorbance at 340 nm was used to measure the degree of light scattering from aggregated protein and therefore provide a suggestive indicator of protein folding into its native structure.

The iFOLD System contains numerous conditions not tested, thus more rigorous testing might yield improved conditions for refolding. Based on results from preliminary refolding screening using the iFOLD System 3, a smaller screen was conducted to determine the effect of fewer variables on inclusion body aggregation, as shown in Table 2. It was interesting that SNM1A inclusion bodies aggregated less at pH 8.5, since proteins are generally less soluble when they are buffered close to their isoelectric point. SNM1A has a calculated pI of 8.24. However, it is possible that the TAPS buffer, rather than the pH of 8.5, was responsible for the observed decrease in aggregation during refolding assays.

The family of FoldACE reagents is composed of proprietary refolding additives of organic salts less than 500 Da, which have been empirically determined to frequently facilitate refolding. Such refolding additives can be categorized as either refolding enhancers, which increase the stability of the protein, or aggregation suppressors, which

reduce unwanted side chain interactions. Current reports available for FoldACE reagents do not distinguish between these possibilities. As such, the mechanistic basis underlying the observation that FoldACE reagents were able to significantly enhance the recovery of soluble SNM1A refolded from denatured inclusion bodies is not yet understood.

4.1.4 5’ Exonuclease activity of SNM1A

Despite problems with insolubility and degradation of protein obtained from refolded inclusion body preparations, SNM1A could be successfully refolded from inclusion bodies. Thus represents the first report of a successful protocol to obtain SNM1A from inclusion bodies obtained from expression in *E. coli.* Importantly, 5’ exonuclease activity observed from the preparations could be attributed to proper inclusion body refolding, as shown in Figure 9, since omission of refolding agents or various reaction components resulted in decreased or abolished nuclease activity. Equally important, observed nuclease activity was 5’ phosphate-dependent, a unique biochemical property of SNM1A and other beta-CASP family members, suggesting that the 5’ nuclease activity was specifically from refolded SNM1A and not from some other contaminating nuclease.

4.2 Discussion of the beta-CASP domain of SNM1A

4.2.1 Expression of the beta-CASP domain of human SNM1A in *E. coli*

In comparison to expression of full-length SNM1A, the beta-CASP domain of SNM1A was highly overexpressed in *E. coli,* in which beta-CASP domain inclusion bodies comprised more than 30% of total cell mass. This increased expression may be simply a result of the shorter construct boundaries, whose messenger RNA may be at a lower risk of forming unproductive secondary structure for translation machinery. Indeed, it is possible that the messenger RNA for full-length SNM1A formed some secondary structure that compromised translation. Similar to SNM1A, however, the beta-CASP domain was almost completely insoluble. The beta-CASP domain was marginally soluble when expressed at 16 ° C; however, very low yields were obtained from large scale expression and purification attempts even when expression was conducted at lower temperatures. Given the insolubility of both proteins, it can be concluded that the N-terminus of SNM1A is not predominantly responsible for either protein's insolubility. The abundance of inclusion bodies and prior success in refolding full-length SNM1A made it reasonable to undertake a systematic search for conditions able to generate refolded, soluble and active beta-CASP domain protein.

4.2.2 Optimization of isolation of beta-CASP domain inclusion bodies

A number of inclusion body isolation protocols were compared in order to generate a protocol for inclusion body isolation of the beta-CASP domain (See

Appendix). The first parameter for inclusion body isolation was optimization of components of wash buffers. The use of Triton X-100 is standard in many inclusion body isolation protocols, as it facilitates the removal of membrane bound proteins (Singh & Panda, 2005). EDTA is also commonly used to chelate metals, which are reported to lead to non-specific disulphide formation from metal catalyzed air oxidation of cysteines (Singh & Panda, 2005). The protocol for the isolation of human interferon alpha 2B included a 5M urea wash containing 20mM DTT (Valente, 2006). This wash was tested for its efficacy to increase the purity of the isolated beta-CASP domain inclusion bodies. This treatment resulted in a loss of the beta-CASP domain, but increased the final purity of the isolated inclusion body material, with removal of contaminants having apparent molecular masses of 40 and 70 kDa in Figure 12.

The second parameter optimized for the isolation of inclusion bodies was the force at which the inclusion bodies were isolated. Low speed centrifugation is able to separate inclusion bodies from cellular components since inclusion bodies are generally denser than cellular debris. Of the protocols for inclusion body isolation reviewed (see Appendix), the centrifuge force varied. Therefore, this parameter was investigated systematically. Optimization of the force used for centrifugation balances inclusion body recovery and purity of recovered inclusion bodies. In Figure 13, a co-pelleting contaminating protein with the apparent molecular weight of 36 kDa was observed in the pellet in the presence of Triton X-100 at all centrifuge forces used to recover inclusion bodies. Fortunately, this contaminating species was successfully removed after the

removal of the detergent and subsequently centrifuging the beta-CASP domain inclusion bodies at a centrifugal force of only 8,000 x g, but not 16,000 or 39,000 x g.

4.2.3 Beta-CASP domain inclusion body solubilization

Protein aggregation is thought to occur largely as a combination of ionic and hydrophobic forces (Singh & Panda, 2005). Although most inclusion body solubilization methods use chaotropes such as guanidine or urea, which completely disrupt protein structure to unfold the protein, some reports have suggested that other methods of inclusion body solubilization are more effective in recovering bioactive protein when protein within the inclusion body are not fully denatured during solubilization (Li et al, 2012, Park et al, 2012, Valente et al, 2006). Despite the aggregation of proteins within inclusion bodies, these proteins are not without some secondary structure. Studies have shown that proteins within inclusion bodies exist as partially folded intermediates with some secondary structure; thus milder methods of solubilization keep this native-like secondary structure intact and in some cases, improve the yield of recovered protein (Singh & Panda, 2005). For example, human growth hormone expressed as inclusion bodies yielded almost three-fold more refolded protein from inclusion bodies solubilized in alkaline conditions compared to those solubilized in chaotropes (Singh & Panda, 2005).

Solubilization in alkaline conditions is thought to result from destabilization of ionic forces at a pH distant from the pI of the protein of interest, which does not disrupt native-like secondary structure. Such conditions are not sufficient to generate random coils during the refolding process, which are much more likely to occur when refolding inclusion bodies solubilized with high concentrations of chaotropes. Consistent with this proposal, the amount of beta-CASP domain which could be solubilized was significantly increased at high pH, as shown in Figure 14. The drop in protein solubilized at pH 12 may be a result of polypeptide degradation.

Solubilization of inclusion bodies can also be achieved quite efficiently with ionic detergents, such as SDS, Sarkosyl, or N-laurylsarcosine. This detergents disrupt ionic and hydrophobic interactions while, in some cases, maintaining some secondary structural elements (Singh & Panda, 2005). Solubilization with ionic detergents is thought to maintain some ordered structure, such as alpha helices (Singh & Panda, 2005). This in turn allows for the solubilization of protein, which, when the detergent is removed, can more readily assume its native conformation compared to protein refolded from conditions which fully removed all secondary structure. Since human chemokine receptor, CXCR1, has been shown to be successfully refolded after solubilization using SDS, solubilization of the beta-CASP domain using SDS was thus tested (Park et al, 2012). Preliminary data for the solubilization of the beta-CASP domain using SDS demonstrates the threshold at which the beta-CASP domain inclusion bodies are soluble, as shown in Figure 15.

After optimization of the above solubilization parameters, each was independently and additionally compared to determine the best method of solubilization (see Figure 16). Guanidine solubilization was shown to be more effective to denature the beta-CASP domain. This is not surprising since guanidine and urea are most commonly used for inclusion body solubilization (Singh & Panda, 2005).

4.2.4 Refolding of beta-CASP domain inclusion bodies

Protein aggregation during refolding of denatured inclusion body preparations can occur due to a variety of reasons. Refolding is a first order reaction, whereas aggregation is a higher order kinetic reaction (Alibolandi & Mirzahoseini, 2011). Rapid transfer from denaturing conditions into aqueous conditions may lead to misfolding and aggregation since intermediate folded structures are more unstable and cannot fluctuate to fold into a native structure (Alibolandi & Mirzahoseini, 2011). Consequently, it is a common observation that increased protein concentration can decrease, sometimes drastically, the yield of properly refolded protein (Alibolandi & Mirzahoseini, 2011). At high concentrations of protein, kinetic competition between refolding and aggregation favours aggregation, thereby decreasing the yield of protein correctly refolded protein (Alibolandi & Mirzahoseini, 2011). As such, it is often more productive to perform inclusion body refolding under extremely diluted conditions.

Completely unfolded inclusion bodies are also more prone to aggregation due to formation of non-native disulphide (Woycechowsky & Raines, 2000). The addition of a mixture of oxidized and reduced disulphides frequently lowers the extent of incorrectly formed disulphide bonds (Woycechowsky & Raines, 2000). Finally, impurities within the solubilized inclusion body preparation may also interfere with refolding and promote aggregation. Although inclusion bodies are composed mostly of insoluble recombinant protein, these entities also contains contaminants, such as RNA polymerases, membrane proteins, and other various enzymes, which may be adsorbed to the hydrophobic regions of the inclusion bodies (Valente et al, 2006). Such contaminants are solubilized with the inclusion bodies during solubilization procedures, and are then present and able to interact with the refolding recombinant protein, in many cases leading to aggregation.

For refolding the unpurified beta-CASP domain, decreasing its concentration to 1μM reduced the degree of observed aggregation to acceptable levels. Lower concentrations further reduced aggregation, but reduced yields, making it infeasible to assess the amount of the beta-CASP domain recovered via the Bradford assay for protein concentration determination. Purified beta-CASP domain, on the other hand, could be refolded at concentrations up to 10 μM. This suggested that the concentration-dependent effect of inclusion body aggregation from the unpurified beta-CASP domain in Figure 17 may have been attributable to the presence of other contaminating proteins.

4.2.5 Exonuclease activity of the beta-CASP domain

The beta-CASP domain of SNM1A was refolded and shown to be strictly dependent upon a 5' phosphate (see Figure 21). However, unexpected 3' exonuclease activity was observed, especially in the presence of dsDNA substrates. Nuclease reactions conducted with refolded beta-CASP ^{D736A} domain also showed the same surprising 3’ exonuclease activity. The most likely explanation for this observation was a contaminating nuclease present in the final preparation. A less likely possibility was that there existed another active site for this additional nuclease activity within the same SNM1A polypeptide fragment.

To exclude the possibility of a second active site, purification of the beta-CASP domain and its mutant was necessary. The presence of a hexa-histidine tag was exploited for protein purification using FPLC IMAC under denaturing conditions prior to refolding procedures. During FPLC purification however, a significant back pressure and a complete blockage of Ni-NTA columns was observed, quite likely due to serious aggregation, despite the protein being fully unfolded in 8M urea.

When Ni-NTA agarose was used to purify the beta-CASP domain under denaturing conditions, aggregation still posed a significant problem. An important observation, however, was that addition of a reducing agent, such as DTT, significantly reduced aggregation of protein bound to the nickel resin and increased the final yield of

protein during elution. Therefore, it is likely that aggregation was a result of random disulphide formation, possibly from metal-catalyzed oxidation. EDTA was an important component of buffers for inclusion body isolation and was included to inhibit metal-catalyzed oxidation. However, EDTA was omitted during inclusion body purification because it would chelate nickel from IMAC resin. Thus, during the IMAC stage of purification, the absence of EDTA increased the probability of metal-catalyzed oxidation of cysteines throughout the beta-CASP domain. Addition of 10 mM DTT was therefore found to be essential for inhibiting disulphide formation and increasing the amount of overall beta-CASP domain eluted from Ni-NTA resin.

Importantly, a significant amount of contaminating protein was removed as a result of incorporating IMAC purification. This can be clearly seen by comparing the unpurified beta-CASP domain fraction to the unbound fraction, containing proteins which did not bind to the nickel resin (first and second lanes of Figure 19). Given the efficacy of this protein purification, inclusion body isolation may not have required the optimized washes reported earlier, especially the 5M urea and 20mM DTT wash, which increased inclusion body purity at the expense of final yield. Rather, it may have been better to proceed directly to purification using IMAC under denaturing conditions, where the beta-CASP domain can be directly isolated from insoluble cell pellet.

For elution of the beta-CASP domain under denaturing conditions, more beta-CASP domain could be eluted with 500mM imidazole and yielded more refolded protein;

however this protein preparation was not active, as seen in Figure 20. Imidazole may have resulted in the chelation of zinc from the metallo beta-lactamase active site during refolding, disrupting the binuclear zinc active site. The interaction of imidazole with charged residues may have also inhibited refolding. Further nucleases assays supplemented with increasing concentrations of zinc may help eliminate the possibility of imidazole-mediated zinc chelation.

Elution at low pH (pH 4.5) resulted in the transition of histidine residues within the hexa-histidine tag from a pyrrole to an ammonium group. This transition is reported to greatly reduce the interaction between the hexa-histidine tag and Ni-NTA resin. After elution, the pH was simply adjusted to the pH required for refolding. However, pH 4.5 is out of the buffering capacity of Tris. As the concentration of protein eluted from the IMAC resin increased, the pH of the solution increased, eliminating the pH-dependent capacity of the buffer. In order to elute the protein, the volume of buffer was increased, thereby decreasing the concentration of purified solubilized protein. The use of another buffer with buffering capacity at pH 4.5, such as citrate or acetate, may effectively reduce the fluctuations in pH, thereby increasing the concentration of purified protein eluted and final yield of protein.

After refolding the purified beta-CASP domain eluted from Ni-NTA resin at a pH of 4.5, only wild-type beta-CASP domain demonstrated 5' exonuclease activity, while beta-CASP ^{D736A} demonstrated no 5' exonuclease activity (see Figure 22). Moreover, no 3'

exonuclease activity was observed for either preparation, excluding the possibility of another active site for 3' exonuclease activity within the same domain. Importantly, this confirmed the presence of a nuclease contaminant in inclusion body preparations that could be removed during the IMAC purification step.

One strategy that may be employed to increase the final yield of the beta-CASP domain may be immobilizing the inclusion bodies onto Ni-NTA resin under denaturing conditions, and then performing a batch purification protocol through a column rather than via centrifugation-based separation. Rather than refolding these purified inclusion bodies by rapid dilution, on-column refolding may be performed for greater overall success and recovery of active protein. Such steps may eliminate problems with aggregation, non-specific disulphide bond formation, and refolding of protein into an active state. Such refolded protein might then be eluted using standard imidazole elution buffers, which could be subsequently removed by dialysis or buffer exchange procedures to yield milligram quantities of beta-CASP domain protein.

4.2.6 Enzyme kinetics of purified, refolded beta-CASP domain

Purified beta-CASP domain was used to assess enzyme kinetic parameters. Initial velocities were calculated under conditions where after sixteen minutes following the initiation of the reaction, only 10% of the initial substrate was degraded. Under these

conditions, it was possible to apply the steady-state approximation for Michalis-Menten kinetics. This analysis indicated that about three times more enzyme was required to achieve the same rate of catalysis for dsDNA compared to that used for ssDNA.

Experiments for K_m determination were performed multiple times in order to determine the amount of substrate required to saturate the beta-CASP domain active site. Despite this, a significant degree of substrate hydrolysis was observed at ratios upwards of 1:2000 of enzyme to substrate. Under these reaction conditions, the calculated K_m for the refolded beta-CASP domain was significantly higher than that calculated for purified SNM1A $^{608\text{-}1040}$, with a difference of one to two orders of magnitude.

Kinetic parameters have not been calculated for full-length SNM1A, but the hydrolysis of full-length SNM1A and SNM1A$^{608\text{-}1040}$ on ssDNA and dsDNA has been compared qualitatively by Wang et al (2011). For both substrates, full-length SNM1A demonstrates about a ten-fold decrease in activity compared to SNM1A$^{608\text{-}1040}$. The catalytic efficiency for full-length SNM1A is therefore an estimate in Table 3 and Table 4.

The k_{cat} for the beta-CASP domain on ssDNA was two-fold lower compared to that for dsDNA, suggesting that SNM1A can turn over more dsDNA product in the same amount of time. This difference in substrate turnover for ssDNA or dsDNA for the beta-

CASP domain of SNM1A is consistent with the comparative substrate turnover difference of SNM1A$^{608\text{-}1040}$ (Wang et al, 2011). However, the calculated k_{cat} of the beta-CASP domain was significantly less than that reported for SNM1A $^{608\text{-}1040}$.

The specificity constant calculated by k_{cat}/K_m indicates that the beta-CASP domain is more catalytically efficient on ssDNA than dsDNA however. This is consistent with all studies on SNM1A, which have demonstrated that ssDNA is preferred over dsDNA (Hejna et al, 2007, Hazrati et al, 2008, Wang et al, 2011). However, in comparing the specificity constant for the beta-CASP domain to that for SNM1A$^{608\text{-}1040}$, the beta-CASP domain is less catalytically efficient than SNM1A$^{608\text{-}1040}$.

Table 4: Comparison of kinetic parameters of refolded beta-CASP domain, SNM1A$^{608-1040}$ and SNM1A$^{1-1040}$ for ssDNA

Substrate	Beta-CASP$^{707-1040}$	SNM1A$^{608-1040}$	SNM1A$^{1-1040}$
K_m (μM)	1.1	0.005	n/a
k_{cat} (s^{-1})	0.4	0.23	n/a
k_{cat}/K_m ($M^{-1}s^{-1}$)	3.2×10^4	4.6×10^7	5×10^6

Table 5: Comparison of kinetic parameters of refolded beta-CASP domain, SNM1A$^{608-1040}$ and SNM1A$^{1-1040}$ for dsDNA

Substrate	Beta-CASP$^{707-1040}$	SNM1A$^{608-1040}$	SNM1A$^{1-1040}$
K_m (μM)	5.5	0.1	n/a
k_{cat} (s^{-1})	0.83	0.43	n/a
k_{cat}/K_m ($M^{-1}s^{-1}$)	1.5×10^4	4.4×10^6	4×10^5

This difference in kinetic parameters may reflect the construct boundaries of the compared proteins. The beta-CASP domain encompasses amino acids residues from 707 to 1040, where SNM1A$^{608\text{-}1040}$ has an additional 99 residues at its N-terminus. These additional residues N-terminal to the beta-CASP domain may facilitate substrate binding. Substrate binding and substrate specificity can be significantly affected by interactions outside the catalytic core. Such is the case previously reported for two other well characterized human 5' exonucleases, FEN1 and EXO1. Although both of these 5' exonucleases share the same active-site structure, the DNA substrate has been shown to make additional contacts with the 3' flap pocket of EXO1 not found in FEN1 (Grasby, 2011). Therefore, amino acid residues N-terminal to the beta-CASP domain may similarly stabilize protein-substrate interactions and be responsible for at least some of the observed difference in K_m and k_{cat} for the minimal beta-CASP domain and the previously characterized SNM1A$^{608\text{-}1040}$.

Alternatively, additional residues N-terminal to the beta-CASP domain may be required for protein stabilization. For the construction of the beta-CASP domain in this study, the boundaries of this domain was based on sequence alignment of representative beta-CASP family members as well as predicted secondary structure of SNM1A. It is therefore possible that important residues for protein stability may have been truncated, resulting in this difference in kinetic values determined for the beta-CASP domain and SNM1A$^{608\text{-}1040}$.

The decrease in catalytic efficiency may also be a result of sub-optimal refolding. It is assumed that if the protein is soluble and active, that the protein is in its correct native conformation. Further assessment of secondary structural elements using methods, such as circular dichroism or melting point curves, can be used to determine the contribution of sub-optimal refolding to the decreased binding efficiency of the refolded beta-CASP domain. Alternatively, it is possible that only some inclusion bodies were refolded into its native conformation, while some did not assume this structure. Any misfolded protein would therefore be included in the total protein concentration, as determined by the Bradford assay, decreasing the k_{cat} and k_{cat}/K_m of the protein. The amount of aggregated protein could be possibly reduced by size exclusion chromatography.

The calculated K_m for the refolded beta-CASP domain is within the same order of magnitude as some enzymes which contain the conserved metallo beta-lactamase domain. For example, AmpC from *E. coli* stain HKY28 is an inhibitor-sensitive beta-lactamase that has K_m values for ampicillin, cefoxitin, and ceftazidime substrates ranging from 1.2 to 6.2µM, however, on the same substrates, this enzyme exhibits much lower k_{cat} values (Doi et al, 2004). This suggest that although substrate binding affinity may be similar for these two metallo beta-lactamase enzymes, the catalytic efficiency is greatly reduced for the refolded beta-CASP domain of SNM1A compared to some isolates of AmpC. This reduced catalytic efficiency may be a result of the previously described above factors.

The refolded beta-CASP domain of SNM1A can also be compared to other 5' exonucleases. Kinetic parameters for the refolded beta-CASP domain are similar to those obtained for RecJ-like protein Mpn140, a 5'exonuclease active on ssDNA (Wakamatsu et al, 2010). However, lambda Exo, a 5' exonuclease which has a strong preference for a 5'phosphate and dsDNA, has a calculated k_{cat}/K_m which is several orders of magnitude higher for its preferred substrates compared to the calculated k_{cat}/K_m for the refolded beta-CASP domain (Mitsis & Kwagh, 1999). Therefore, although the kinetic parameters reported in this study for the beta-CASP domain of SNM1A is similar to some 5' exonucleases, it is much significantly less efficient than comparatively similar enzymes. Again, this may be a result of the above factors.

Overall, these kinetic parameters for variants of SNM1A have interesting implications for ICL repair. Higher substrate turnover for dsDNA would be consistent with an ICL repair model in which SNM1A is active on DNA nicked by ERCC1-XPF to create a dsDNA substrate. However, it is not clear how the increase in catalytic activity on ssDNA could be understood with regards to SNM1A's role in ICL repair. In current models of ICL repair, ssDNA is not thought to be generated as an intermediate during the repair pathway. Further investigation will be required to elucidate if ssDNA intermediates are generated during ICL repair and if these are processed by SNM1A.

4.3 *In vivo* activity of a minimal beta-CASP domain

Although it can be demonstrated that the beta-CASP domain of SNM1A encompasses the minimal functional unit that has exonuclease activity thought to be required for ICL repair, *in vivo* survival and sensitivity assays were used to determine if this domain alone is sufficient to carry out necessary nuclease activities within the cell.

Various truncations of SNM1A have been reported within the literature, examining the ability of these truncations to form nuclear foci localization, but they have never been characterized in terms of their effects on survival following ICL damage.

SNM1A was originally identified to participate in ICL repair because of its ability to increase the survival of cells lacking human SNM1A's yeast homologue, Pso2, when treated with cisplatin and mitomycin C. Using the same assay system, the efficacy of the minimal beta-CASP domain was examined by complementing ICL repair defects in *pso2* null strains after cisplatin treatment.

Qualitative results from the cisplatin sensitivity assay and quantitative results from the cisplatin survival assay demonstrate that the beta-CASP domain of SNM1A cannot complement *pso2* null strains to permit survival in the presence of 1 mM cisplatin. Although this construct has been demonstrated to have activity *in vitro*, it is not sufficient

for ICL repair *in vivo.* Removal of domains found in the N-terminus, such as the nuclear foci formation region, PCNA interacting protein box and ubiquitin zinc finger (see Figure 3) likely affect SNM1A's ability to co-localize with other repair proteins in distinct foci at site of repair after DNA damage. It is therefore likely that without residues 1 to 706 of SNM1A, it cannot properly localize to the site of ICL repair to enable the nuclease activity of the beta-CASP domain to function appropriately. Insufficiency of the beta-CASP domain alone may also reflect possible decreased enzyme kinetics as a result of protein instability or reduced protein-substrate interaction stability. More work will be required to further elucidate such possibilities.

CHAPTER 5: CONCLUSIONS AND FUTURE DIRECTIONS

The work presented in this thesis focused on determining the sufficiency of the beta-CASP domain of SNM1A for nuclease activity *in vitro* and its ability to repair ICL damage *in vivo*. Biochemical studies required significant improvement to current available methods for isolation of recombinant SNM1A protein. As such, this thesis also centered on developing methods aimed at increasing the yield of bioactive protein. The ability to generate sizable amounts of active SNM1A will facilitate our understanding of this important and unique protein, integral in the repair of ICLs.

Worked outlined in this thesis demonstrates that the beta-CASP domain of SNM1A can be used as the minimal functional unit for nuclease activity *in vitro*. This domain can be overexpressed as inclusion bodies and successfully refolded. This is the first time that this protein or any of its domains have been to be refolded and to yield bioactive protein for *in vitro* studies. Although the beta-CASP domain is not sufficient for ICL repair *in vivo*, it is sufficient for 5' phosphate-dependent exonuclease activity *in vitro.* The expression and purification of the beta-CASP domain of SNM1A determined during this thesis significantly eases the ability to obtain active pure protein. Only one litre of cultured *E. coli* and one major purification step is required to yield microgram quantities, and possibly more, of active protein as opposed to twenty four litres and six purification columns, as reported previously (Wang, 2011). Although this refolded protein is less nucleolytically efficient compared to purified $SNM1A^{608-1040}$, the amount of protein

obtained through refolding of solubilized inclusion bodies makes it possible to conduct studies, such as high throughput screening for small molecule inhibitors and structural studies including x-ray crystallography.

Future direction will be focused on finding inhibitors of the beta-CASP domain of SNM1A. A fluorescence-based biochemical assay has recently been reported by Wang and colleagues for high-throughput screening (2011). With this assay and the abundance of protein now available, high-throughput screening for inhibitors with potentially anti-cancer activity can be conducted occur in the very near future. Current research focused of ICL repair has been largely limited to inhibition of ERCC1-XPF, but this protein is involved in DNA damage pathways other than ICL repair. As such, it is not an ideal anti-cancer drug target. Also, repair of ICL adducts is only one mechanism of cisplatin resistance. Given the specificity of SNM1A to only ICL repair, inhibition of this enzyme may specifically sensitize tumors which have become resistant to ICL-based chemotherapy without many of the off-target effects observed for inhibitors of other DNA repair proteins.

Additionally, future work must involve further optimizing the beta-CASP domain on-column refolding to increase yields further for further structural studies. Despite current structures of beta-CASP nucleases which hydrolyze RNA, there are no structures of beta-CASP nucleases which hydrolyze DNA to date. Moreover, none of the RNA beta-CASP nucleases were crystallized with substrate bound within the active site. As such,

structural determination of the beta-CASP domain, especially that of SNM1A, may help elucidate differences in substrate specificity for RNA versus DNA. Substrate bound SNM1A may also give insight into the catalytic mechanism of this strict-phosphate dependent nuclease activity. This is a defining property of all family members and may prove to be a key aspect of designing highly specific 5' exonuclease inhibitors. Finally, structural information will be critical for guiding attempts at structure-based drug design, increasing the specificity and affinity of a potential inhibitor of SNM1A.

Expansion of *in vivo* complementation studies described in this thesis would be highly beneficial. Importantly, heterologous protein expression levels must be determined in order to properly compare the ICL repair capacity of the proteins tested. Although the beta-CASP domain was not found to be sufficient for ICL repair, the question remains - is there a minimal functional unit for ICL repair *in vivo*? Longer truncations will need to be tested for their ability to complement *pso2* null yeast treated with cisplatin and possibly other ICL agents, such as mitomycin C or SJG-136, which result in minimal local distortion of the DNA backbone. The inclusion of the PIP box at amino acid positions 556-563 may permit association with PCNA, possibly allowing localizing of SNM1A to the site of ICL repair. Both Pso2 and SNM1A have been shown to possess UBZ domains in their N-terminal regions. Truncations of this domain have yet to be functionally characterized with regards to ICL repair capacity. Expansion of the complementation studies will provide more information about domain boundaries required by SNM1A for ICL repair.

Further analysis focused on careful evaluation of substrate specificity would be highly useful. Although hairpin-opening activity by SNM1A has been repeatedly tested, it may be possible that SNM1A harbours intrinsic structure-specific endonuclease activity that is highly regulated. Indeed, the beta-CASP domain of SNM1A may be similar to the beta-CASP domain of Artemis and its requirement for constant association and phosphorylation by DNA-PKcs. Interestingly, the beta-CASP domain of Artemis, which has endonuclease activity, has been shown to rescue *pso2* null defects. Additionally, SNM1A's homologue, Pso2, is known to possess hairpin-opening activity. Therefore, the structure-specific endonuclease activity of the beta-CASP domain of SNM1A must be thoroughly investigated. This is a particularly exciting avenue of future studies given the work reported here which provides easy access to workable quantities of pure SNM1A.

Mechanistically, it is thought that the nuclease activity of SNM1A involved in ICL repair is limited to its 5' exonuclease activity. However, this interpretation is inconsistent with the current model of unhooking damaged DNA. It is instead possible that the structure of the cross-linked DNA is hairpin-like in nature, and that SNM1A uses structure-specific endonuclease activity to create the second incision 3’ to the ICL adduct. This may only occur in association with an accessory protein or after some post-translational modification.

Proteins that interact with SNM1A may be vital to understanding the regulation and activity of SNM1A. As previously mentioned, other human beta-CASP homologues

only function *in vivo* in association with other auxiliary proteins. Although a number of proteins that interact with SNM1A have been identified, it is not clear if they modulate its activity. Like Artemis, SNM1A may require phosphorylation. Additionally, SNM1A may require SUMOlyation, since abolishment of its interaction with PIAS1, a SUMO E3 ligase, results in increased cisplatin sensitivity (Ishiai, 2004). It is interesting that this interaction is mediated by amino acids of the conserved CASP motifs, D838 and H994, which are thought to coordinate the binuclear zinc active site. Currently it is unknown if interaction with PIAS1 merely localizes SNM1A to the DNA damage repair sites or if it modulates the nuclease activity of it. Although not reported in this thesis, significant work has already been undertaken to investigate the possible binding partners of SNM1A. A yeast 2-hybrid analysis dedicated to search for binding partners needs to be completed in conjunction with studies aimed at defining functional domain boundaries.
Further studies aimed at understanding the domain boundaries, binding partners and regulation of SNM1A will be highly important to elucidating the mechanism of this unusual ICL repair protein.

In the past, low yields and difficult expression and purification of SNM1A made biochemical studies virtually impossible. The results presented in this thesis are a significant contribution to the field of ICL repair, and importantly, anti-cancer research.

APPENDIX

Table 6: Expression vectors of SNM1A, beta-CASP, and Pso2

Target DNA	**Template**	**Primer sequence**	**Destination vector**	**Expression**
SNM1A WT (1-1040)	pEBS7-hSNM1	SNM1A_F_1, SNM1A_R	pDEST17	*E. coli*
SNM1A WT (1-1040)	pEBS7-hSNM1	SNM1A_F_1, SNM1A_R	pYES-DEST52	*S. cerevisiae*
SNM1A D736A (1-1040)	pAC5.1-SNM1A D736A (from Rob Moses)	SNM1A_F_1, SNM1A_R	pDEST17	*E. coli*
SNM1A D736A (1-1040)	pAC5.1-SNM1A D736A	SNM1A_F_1, SNM1A_R	PYES-DEST 52	*S. cerevisiae*
Beta-CASP (707-1040)	pEBS7-hSNM1	SNM1A_F_707, SNM1A_R	pDEST17	*E. coli*
Beta-CASP (707-1040)	pEBS7-hSNM1	SNM1A_F_707, SNM1A_R	pYES-DEST52	*S. cerevisiae*
Beta-CASP D736A (707-1040)	pAC5.1-SNM1A D736A	SNM1A_F_707, SNM1A_R	pDEST17	*E. coli*
Beta-CASP D736A (707-1040)	pAC5.1-SNM1A D736A	SNM1A_F_707, SNM1A_R	pYES-DEST 52	*S. cerevisiae*
PSO2	pQE32-scSNM1 (from James Hejna)	PSO2_F, PS02_R	pYES-DEST52	*S. cerevisiae*

List of primers for cloning SNM1A, beta-CASP, and Pso2

SNM1A_F_1 : 5'-GGG ACA AGT TTG TAC AAA AAA GCA GGC TTA GAT TAC GAT ATC CCA ACG ACC GAA AAC CTG TAT TTT CAG GGC ATG TTA GAA GAC ATT TCC GAA GAC-3'

SNM1A_R: 5'-GGG GAC CAC TTT GTA CAA GAA AGC TGG GTC TCA TCA ATA TCC AGC TTC CAA TTT CC-3'

SNM1A_F_707: 5'GGG GAC AAG TTT GTA CAA AAA AGC AGG CCT AGA TTA CCC AAC GAC CGA AAA CCT GTA TTT TCA GGG CAT GAT ATG GGG TCC TGA TCC CCC-3'

PSO2_F: 5' GGGG ACA SGT TTG TAC AAA AAA GCA GGG TTA GAT TAC GAT ATC CCA ACG ACC GGA AAC CTG TAT TTT CAG GGC ATG TCA AGG AAA TCT ATA GTG C-3'

PSO2_R: 5' GGG GAC CAC TTT GTA CAA GAA AGC TGG GTC TTA TTT AGC CGC CCG CAT TTT CC-3'

PROTOCOL	RESUSPENSION	LYSIS	POST-LYSIS	CENTRIFUGE	WASH 1	WASH 2	WASH 3	WASH 4	WASH 5	SOLUBILIZE
iFOLD	50mM Tris	sonication 3x1min	Add 1% Triton, stir 15 min RT	8000g, 15 min, 8 degrees	1% triton					GuHCl
	50mM NaCl				50mM Tris	50mM Tris	50mM Tris			50mM Tris
	1mM tcep,				50mM NaCl	50mM NaCl	50mM NaCl			200mM NaCl
	.5 mM EDTA,				1mM tcep,	1mM tcep,	1mM tcep,			2mM EDTA
	5% glycerol,				.5 mM EDTA,	.5 mM EDTA,	.5 mM EDTA,			1mM TCEP
	benzonase				5% glycerol,	5% glycerol,	5% glycerol,			
	lysozyme									
Valente et al, 2006	50mM NaCl	sonication 6x1min		5000g, 10min, 4 degrees	1% triton	5M urea				6M GuHCl
	1mM EDTA				50mM NaCl	20mM DTT				50mM Tris
										100mM BME
Strong Lab	50mM Tris	Sonication	10mM MgSO4	6000rpm, 15 min	50mM Tris	50mM Tris	50mM Tris	50mM Tris		100mM Tris buffer at pH 8.0
	100mM NaCl		Dnase .01 mg/mL		100mM NaCl	100mM NaCl	100mM NaCl	100mM NaCl		50mM Glycine and
	5mM EDTA		lysozyme to about 0.1mg/ml		5mM EDTA	5mM EDTA	5mM EDTA	5mM EDTA		8.5M urea
	0.1% NaN3		incubate 15min RT		0.1% NaN3	0.1% NaN3	0.1% NaN3	0.1% NaN3		
	0.5% Triton				0.5% Triton	0.5% Triton	0.5% Triton	1mM DTT		
	1mM DTT				1mM DTT	1mM DTT	1mM DTT	SONICATE		
					SONICATE	SONICATE	SONICATE			
Bao et al, 2008	50 mM NaH2PO4, pH 8	ice for 30 min		15000g, 20 min, 4 degrees	Dnase + lysozyme					8 M urea,
	300 mM NaCl,				2 M urea,					100 mM NaH2PO4,
	1 mg/ml lysozyme				100 mM NaH2PO4,					10 mM Tris-HCl, pH 8.0
					10 mM Tris-HCl, pH 8.0					
					then ice 30 min					
Brown et al , 1992	100mM NaCl	sonication 4x30 sec		10000 rpm, 15 min, 4 degree	0.2 M Tris-HCl (pH 8.0)					7.2 M urea
	2.5 mg/ml lysozyme				500mM NaCl					Tris
					1mM DTT					1mM DTT
					and 0.1% Triton X-100					
					SONICATE					
Beassoni et al, 2011	50 mM Tris–HCl,	French press		12,000g, 10 min, 4 degrees	10 mM MgCl2,	0.2 mg mL−1 lysozyme	0.5 % v/v Triton X-100	distilled water		25 mM Tris–HCl,
	100 mM NaCl,				20 µg mL−1 DNase I	50 mM Tris–HCl,	10 mM EDTA			6.5 M urea,
	1 mM EDTA				50 mM Tris–HCl,	100 mM NaCl,	50 mM Tris–HCl,			1 mM DTT,
					100 mM NaCl,	1 mM EDTA	100 mM NaCl,			0.5 mM EDTA,
					1 mM EDTA					5.0 mM glycine, pH 8.5
Protocol to be test	**50mM Tris pH8**	**sonicate 5x1min**	**2mM MgSO4**	**test for best rpm**	**1% Triton**	**1% Triton**	**50 mM Tris–HCl,**	**50 mM Tris–Idistilled water**		
	100mM NaCl		**Benzonase**	**15 min**	**50 mM Tris–HCl,**	**50 mM Tris–HCl,**	**100 mM NaCl,**	**100 mM NaCl,**		
	1mM EDTA		**incubate 15min RT**	**4 degrees**	**100 mM NaCl,**	**100 mM NaCl,**	**1mM EDTA**	**1mM EDTA**		
	2.5 mg/ml lysozyme				**1mM EDTA**	**1mM EDTA**		OR		
			check for compatibility of triton and dnase					**5M urea**		
			Benzonase only 30% active in 1mM Triton					**20mM DTT**		

REFERENCES

Ahkter S, Richie CT, Zhang N, Behringer RR, Zhu C, & Legerski RJ (2005) Snm1-deficient mice exhibit accelerated tumorigenesis and susceptibility to infection. *Mol Cell Biol* **25**: 10071-10078

Akhter S & Legerski RJ (2008) SNM1A acts downstream of ATM to promote the G1 cell cycle checkpoint. *Biochem Biophys Res Commun* **377**: 236-241

Alibolandi M & Mirzahoseini H (2011) Chemical assistance in refolding of bacterial inclusion bodies. *Biochem Res Int* **2011**: 631607

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

Carfi A, Pares S, Duee E, Galleni M, Duez C, Frere JM, & Dideberg O (1995) The 3-D structure of a zinc metallo-beta-lactamase from Bacillus cereus reveals a new type of protein fold. *EMBO J* **14**: 4914-4921

Deans AJ & West SC (2011) DNA interstrand crosslink repair and cancer. *Nat Rev Cancer* **11**: 467-480

Doi Y, Wachino J, Ishiguro M, Kurokawa H, Yamane K, Shibata N, Shibayama K, Yokoyama K, Kato H, Yagi T, Arakawa Y (2004) Inhibitor-sensitive AmpC beta-

lactamase variant produced by an Escherichia coli clinical isolate resistant to oxyiminocephalosporins and cephamycins. *Antimicrob Agents Chemother* **48:**2652-2658.

Dronkert ML, de Wit J, Boeve M, Vasconcelos ML, van Steeg H, Tan TL, Hoeijmakers JH, & Kanaar R (2000) Disruption of mouse SNM1 causes increased sensitivity to the DNA interstrand cross-linking agent mitomycin C. *Mol Cell Biol* **20**: 4553-4561

Dronkert ML & Kanaar R (2001) Repair of DNA interstrand cross-links. *Mutat Res* **486**: 217-247

Gatsby JA, Finger LD, Tsutakawa SE, Atack JM, Tainer JA (2011) Unpairing and gating: sequence independent substrate recognition by FEN superfamily nucleases. *Trends Biochem Sci* **37:** 74-84

Galluzzi L, Senovilla L, Vitale I, Michels J, Martins I, Kepp O, Castedo M, & Kroemer G (2011) Molecular mechanisms of cisplatin resistance. *Oncogene*

Guainazzi A & Scharer OD (2010) Using synthetic DNA interstrand crosslinks to elucidate repair pathways and identify new therapeutic targets for cancer chemotherapy. *Cell Mol Life Sci* **67**: 3683-3697

Hartley JA, Spanswick VJ, Brooks N, Clingen PH, McHugh PJ, Hochhauser D, Pedley RB, Kelland LR, Alley MC, Schultz R, Hollingshead MG, Schweikart KM, Tomaszewski JE, Sausville EA, Gregson SJ, Howard PW, & Thurston DE (2004) 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. *Cancer Res* **64**: 6693-6699

Hazrati A, Ramis-Castelltort M, Sarkar S, Barber LJ, Schofield CJ, Hartley JA, & McHugh PJ (2008) Human SNM1A suppresses the DNA repair defects of yeast *pso2* mutants. *DNA Repair (Amst)* **7**: 230-238

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

Hemphill AW, Bruun D, Thrun L, Akkari Y, Torimaru Y, Hejna K, Jakobs PM, Hejna J, Jones S, Olson SB, & Moses RE (2008) Mammalian SNM1 is required for genome stability. *Mol Genet Metab* **94**: 38-45

Hirsch J (2006) An anniversary for cancer chemotherapy. *JAMA* **296**: 1518-1520

Ho TV & Scharer OD (2010) Translesion DNA synthesis polymerases in DNA interstrand crosslink repair. *Environ Mol Mutagen* **51**: 552-566

Hopton SR & Thompson AS (2011) Nuclear magnetic resonance solution structures of inter- and intrastrand adducts of DNA cross-linker SJG-136. *Biochemistry* **50**: 4720-4732

Ishiai M, Kimura M, Namikoshi K, Yamazoe M, Yamamoto K, Arakawa H, Agematsu K, Matsushita N, Takeda S, Buerstedde JM, & Takata M (2004) DNA cross-link repair protein SNM1A interacts with PIAS1 in nuclear focus formation. *Mol Cell Biol* **24**: 10733-10741

Kohno T, Sakiyama T, Kunitoh H, Goto K, Nishiwaki Y, Saito D, Hirose H, Eguchi T, Yanagitani N, Saito R, Sasaki-Matsumura R, Mimaki S, Toyama K, Yamamoto S, Kuchiba A, Sobue T, Ohta T, Ohki M, & Yokota J (2006) Association of polymorphisms in the MTH1 gene with small cell lung carcinoma risk. *Carcinogenesis* **27**: 2448-2454

Li M, Fan H, Liu J, Wang M, Wang L, & Wang C (2012) High pH solubilization and chromatography-based renaturation and purification of recombinant human granulocyte colony-stimulating factor from inclusion bodies. *Appl Biochem Biotechnol* **166**: 1264-1274

Li X, Hejna J, & Moses RE (2005) The yeast Snm1 protein is a DNA 5'-exonuclease. *DNA Repair (Amst)* **4**: 163-170

Liu X, Lao Y, Yang IY, Hecht SS, & Moriya M (2006) Replication-coupled repair of crotonaldehyde/acetaldehyde-induced guanine-guanine interstrand cross-links and their mutagenicity. *Biochemistry* **45**: 12898-12905

Ma Y, Pannicke U, Schwarz K, & Lieber MR (2002) Hairpin opening and overhang processing by an Artemis/DNA-dependent protein kinase complex in nonhomologous end joining and V(D)J recombination. *Cell* **108**: 781-794

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

Mitsis PG, Kwagh JG (1999) Characterization of the interaction of lambda exonuclease with the ends of DNA. *Nucleic Acids Res* **27**:3057-63

Niedernhofer LJ, Daniels JS, Rouzer CA, Greene RE, & Marnett LJ (2003) Malondialdehyde, a product of lipid peroxidation, is mutagenic in human cells. *J Biol Chem* **278**: 31426-31433

Noll DM, Mason TM, & Miller PS (2006) Formation and repair of interstrand cross-links in DNA. *Chem Rev* **106**: 277-301

Park SH, Casagrande F, Chu M, Maier K, Kiefer H, & Opella SJ (2012) Optimization of purification and refolding of the human chemokine receptor CXCR1 improves the stability of proteoliposomes for structure determination. *Biochim Biophys Acta* **1818**: 584-591

Poinsignon C, Moshous D, Callebaut I, de Chasseval R, Villey I, & de Villartay JP (2004) The metallo-beta-lactamase/beta-CASP domain of Artemis constitutes the catalytic core for V(D)J recombination. *J Exp Med* **199**: 315-321

Rahman KM, James CH, & Thurston DE (2011) Observation of the reversibility of a covalent pyrrolobenzodiazepine (PBD) DNA adduct by HPLC/MS and CD spectroscopy. *Org Biomol Chem* **9**: 1632-1641

Richie CT, Peterson C, Lu T, Hittelman WN, Carpenter PB, & Legerski RJ (2002) hSnm1 colocalizes and physically associates with 53BP1 before and after DNA damage. *Mol Cell Biol* **22**: 8635-8647

Sayre RM & Dowdy JC (2008) The increase in melanoma: are dietary furocoumarins responsible? *Med Hypotheses* **70**: 855-859

Sengerova B, Wang AT, & McHugh PJ (2011) Orchestrating the nucleases involved in DNA interstrand cross-link (ICL) repair. *Cell Cycle* **10**: 3999-4008

Singh SM & Panda AK (2005) Solubilization and refolding of bacterial inclusion body proteins. *J Biosci Bioeng* **99**: 303-310

Tiefenbach T & Junop M (2012) Pso2 (SNM1) is a DNA structure-specific endonuclease. *Nucleic Acids Res* **40**: 2131-2139

Tiefenbach, TE (2011) Functional analysis of Pso2 reveals a novel DNA hairpin endonuclease activity: Implications for interstrand crosslink repair. *Open Access Dissertations and Theses*. Paper 6286.

Tsumoto K, Ejima D, Kumagai I, & Arakawa T (2003) Practical considerations in refolding proteins from inclusion bodies. *Protein Expr Purif* **28**: 1-8

Valente CA, Monteiro GA, Cabral JM, Fevereiro M, & Prazeres DM (2006) Optimization of the primary recovery of human interferon alpha2b from Escherichia coli inclusion bodies. *Protein Expr Purif* **45**: 226-234

Wakamatsu T, Kim K, Uemura Y, Nakagawa N, Kuramitsu S, Masui R (2011) Role of RecJ-like protein with 5'-3' exonuclease activity in oligo(deoxy)nucleotide degradation. *J Biol Chem* **286:**2807-2816

Wang AT, Sengerova B, Cattell E, Inagawa T, Hartley JM, Kiakos K, Burgess-Brown NA, Swift LP, Enzlin JH, Schofield CJ, Gileadi O, Hartley JA, & McHugh PJ (2011) Human SNM1A and XPF-ERCC1 collaborate to initiate DNA interstrand cross-link repair. *Genes Dev* **25**: 1859-1870

Woycechowsky KJ & Raines RT (2000) Native disulfide bond formation in proteins. *Curr Opin Chem Biol* **4**: 533-539

Yang K, Moldovan GL, & D'Andrea AD (2010) RAD18-dependent recruitment of SNM1A to DNA repair complexes by a ubiquitin-binding zinc finger. *J Biol Chem* **285**: 19085-19091