Université de Montréal

Direction des bibliothèques

Université de Montréal

Characterization of the AP Endonuclease Enzyme APN-1 from *C. elegans*

par

Devang Patel

Programme de sciences biomédicales

Faculté de médecine

Mémoire présenté à la Faculté des études supérieures

en vue de l'obtention du grade de maîtrise (M.Sc.)

en sciences biomédicales

Décembre 2007

Université de Montréal

Faculté des études supérieures

Ce mémoire intitulé:

Characterization of the AP Endonuclease Enzyme APN-1 from *C. elegans*

présenté par:

Devang Patel

a été évalué par un jury composé des personnes suivantes:

Dr. Puttaswamy Manjunath
président-rapporteur

Dr. Dindial Ramotar
directeur de recherche

Dr. Ashok Srivastava
membre du jury

RÉSUMÉ

Les membres de la famille Endonuclease IV (Endo IV) d'endonucléases apuriniques/apyrimidiques (AP), sont des enzymes multifonctionnelles impliquées dans la réparation de l'ADN.   En effet, les enzymes   endo IV chez *E. coli* et apurinic/apyrimidinic endonuclease 1 (Apn1) chez *S. cerevisiae* sont capables de traiter des sites abasiques et des extrémités 3' bloquées. Des études récentes ont révélé que les AP endonucléases peuvent non seulement agir en tant que 3'→5' exonucléases, mais participent aussi à l'incision de certaines bases oxidées. Jusqu'à présent, peu de connaissances existent au sujet d'APN-1 de *C. elegans* (CeAPN-1); le premier homologue retrouvé chez un organisme multicellulaire. Nous avons développé un système d'expression chez la levure permettant de produire CeAPN-1 fusionnée à l'étiquette GST, et de la purifier à l'homogénéité pour une caractérisation approfondie. À l'aide d'essais *in vitro*, nous avons confirmé que, tout comme ses homologues, CeAPN-1 possède les activités AP endonucléase et 3'-diestérase. Pourtant, on a pu démontrer par la complémentation inter-espèces que les variantes E97G et E143G de CeAPN-1 ne confèrent aucune résistance à une souche de levure déficiente en AP endonucleases lors d'un traitement aux différents agents d'endommagement à l'ADN, tels que le MMS, le H_2O_2 et la Bléomycine. Ces résultats confirment des observations antérieures lors de la substitution de résidus chargés dans un des quatre domaines conservés d'Endo IV liée à l'activité. Cependant, on a révélé que l'activité enzymatique de CeAPN-1 semble être stimulée en présence d'une composante thermostable. Cela peut suggérer l'implication d'un inducteur ce qui différencie CeAPN-1 de ses homologues.

Mots Clés : CeAPN-1, Endo IV, AP endonucléases, site AP, réparation par excision de bases, réparation de l'ADN, *C. elegans*.

ABSTRACT:

Proteins of the Endonuclease IV (Endo IV) family of AP endonucleases have been well documented for their multifunctional capacities in DNA repair. Studies with the *E. coli* endo IV and *S. cerevisiae* apurinic/apyrimidinic endonuclease 1 (Apn1) have demonstrated that these enzymes have the ability to process abasic sites and 3'-blocked ends. Additional investigations have revealed a capability to act as $3' \rightarrow 5'$-exonucleases and, furthermore, have exposed a participation in the incision of certain oxidatively damaged bases. To date, very little is known of APN-1 (CeAPN-1), the corresponding *C. elegans* Endo IV protein and the first structural homologue to be found in a multicellular animal model. In this study, we have engineered an expression system in yeast to produce a CeAPN-1 fusion protein and subsequently, purified the polypeptide to homogeneity for characterization. We confirmed, through *in-vitro* assays, that CeAPN-1 also possesses AP endonuclease and 3'-diesterase activities as its counterparts found in unicellular models. Moreover, domain function analysis has revealed that the introduction of the CeAPN-1 variants, E97G and E143G, in a yeast strain lacking AP endonuclease activity is unable to restore parental resistance when exposed to the DNA damaging agents MMS, H_2O_2 and Bleomycin. These results coincide with previous observations made following the substitution of charged residues contained in one of the four conserved domains of Endo IV affecting activity. However, unlike its homologues, we showed that the enzymatic activity is enhanced by a heat stable component. This suggests a likely key feature which differentiates this protein from other counterparts, despite structural and functional homology.

Key Words: CeAPN-1, Endo IV, AP endonuclease, AP site, base excision repair, DNA repair, *C. elegans*.

TABLE OF CONTENTS

LIST OF TABLES:

LIST OF FIGURES:

LIST OF ABBREVIATIONS

8-OH-Gua	8-hydroxyguanine
AP	apyrimidic/Apurinic
Ape1	apurinic/apyrimidinic endonuclease 1 (human)
Apn1 and 2	apurinic/apyrimidinic endonuclease 1 and 2 (yeast)
ATP	adenine triphosphate
BER	base excision repair
BLM	bleomycin
C. elegans	Caenorhabditis elegans
CDC9	cell division cycle 9 (DNA ligase)
DHT	5,6-dihydrothymidine
DHU	5,6-dihydro-2'-deoxyuridine
DNA	deoxyribonucleic acid
hDR	direct reversal
dRPase	deoxyribophosphodiesterase
DSB	double strand breaks
E.coli	Escherichia coli
EDTA	ethylenediaminetetraacetate.
Endo III	endonuclease III
Endo VIII	endonuclease VIII
Exo IV	exonuclease IV
fapy-ade	4,6-diamino-5-formamideadenine
fapy-gua	4,6-diamino-5-formamideguanine
FEN1	flap endonuclease 1
Fpg	Formamidopyrimidine-DNA Glycosylase
GAD	GAL4 activation domain
GFP	green fluorescent protein
GG	global genomic
GST	glutathione-S-transferase
H_2O_2	hydrogen peroxide

HhH	hairpin helix domain
His	histidine
HR	homologous recombination repair
IPTG	isopropyl β-D-thigalactopyranoside
kb	Kilobase
kDa	Kilodalton
LIG1	ligase 1
LB	Luria Broth
Mg^{2+}	magnesium ion
MGMT	O6-methylguanine DNA methyltransferase
MMR	mismatch repair
MMS	methyl methane sulfate
NER	nucleotide excision repair
NHEJ	non homologous end joining repair
NIR	nucleotide incision repair
NLS	nuclear localization sequence
Ogg1	8-oxo-deoxyguanosine DNA glycosylase 1
PBS	Phosphate buffer saline
PCNA	proliferating cell nuclear antigen
PCR	polymerase chain reaction
Pol β, δ, ε	polymerase beta, delta, epsilon
RFC	replication factor C
RNA	ribonucleic Acid
RNAi	RNA interference
ROS	reactive oxygen species
S. cerevisiae	*Saccharomyces cerevisiae*
S. pombe	*Schizosaccharomyces pombe*
siRNA	small interfering RNA
SSB	single strand breaks
SV40	simian virus 40

tBH	*tert*-Butyl hydroperoxide
TC	transcription coupled
TIM	triosephosphateisomerase
UDG	uracil DNA glycosylase
UV	ultraviolet
UVDE	ultaviolet damage endonuclease
YPD	yeast extract, peptone, dextrose
XRCC1	x-ray cross-species complementing 1
αdA	α-anomeric 2-deoxyadenine
αdT	α-anomeric 2-deoxythymine

This thesis is dedicated in loving memory of my elder sister,

all that I am or ever hope to be, I owe to her.

ACKNOWLEDGEMENTS:

1. INTRODUCTION AND LITERATURE REVIEW

DNA is victim to continuous attack. Mutations, being products of DNA damage and subsequent genomic instability, have serious consequences to a given cell. From a human perspective, the inability to counter various insults affecting DNA integrity has been correlated to cancer, aging, and multiple disorders [1]. Interestingly, DNA damage can stem from both exogenous and endogenous sources. Thus, exogenous sources (environment and mutagenic chemicals) include ultraviolet and ionizing radiations as well as a plethora of toxins and chemicals. In contrast, endogenous sources of DNA damage primarily include reactive oxygen species (ROS) which arise from metabolic byproducts [2]. Each source manifests itself by causing specific types of modifications to ultimately harm the normal state of the nucleic acid macromolecule. Thus, common insults affecting the double helix can be observed via base/sugar damage (hydrolysis, deamination, alkylation, oxidation, strand breaks, abasic sites), odd base mismatches or pairing (pyrimidine dimers) and strand cross linking [2].

1. Alkylated bases
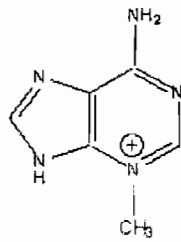
2. Deaminated bases

3-Methyladenine
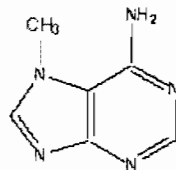
7-Methyladenine
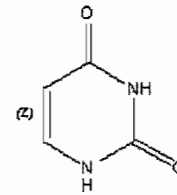
Uracil

3. Oxidized Bases
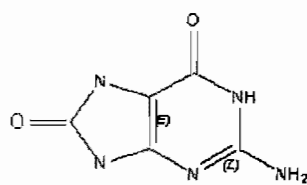
8-Oxoguanine (8oxoG)
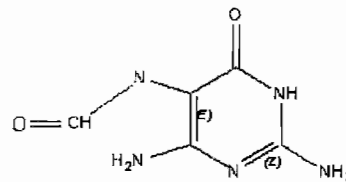
2,5-Amino-5-
Formamidopyrimidine (fapy)
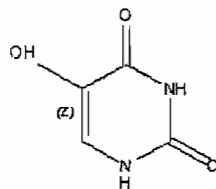
5-hydroxyuracil
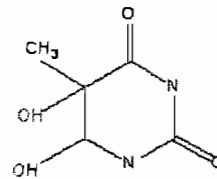
Thymine glycol

4. Other examples of Lesions
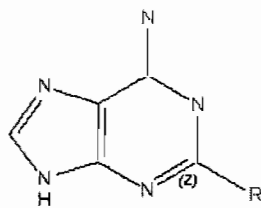
1,N⁶-Etheno-adenine
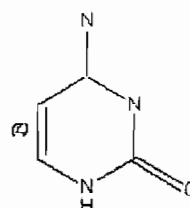
3,N⁴-Etheno-cytosine
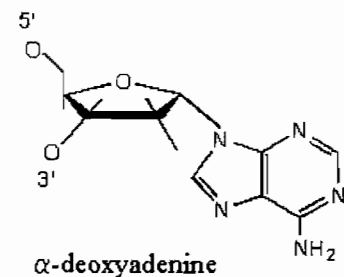
α-deoxyadenine

Figure 1-1: Examples of Lesions Stemming from Altered Bases (adapted from [2, 3])

A particular interest in this study would be devoted to the formation and repair of apurinic / apyrimidinic (AP) sites (shown below). These lesions (abasic) are characterized by a loss of the base component from the sugar phosphate moiety (seen in figures 1-2, 1-3, 1-4) [4]. These sites can occur spontaneously (hydrolysis) and also via enzymatic means to remove damaged bases [4, 5].
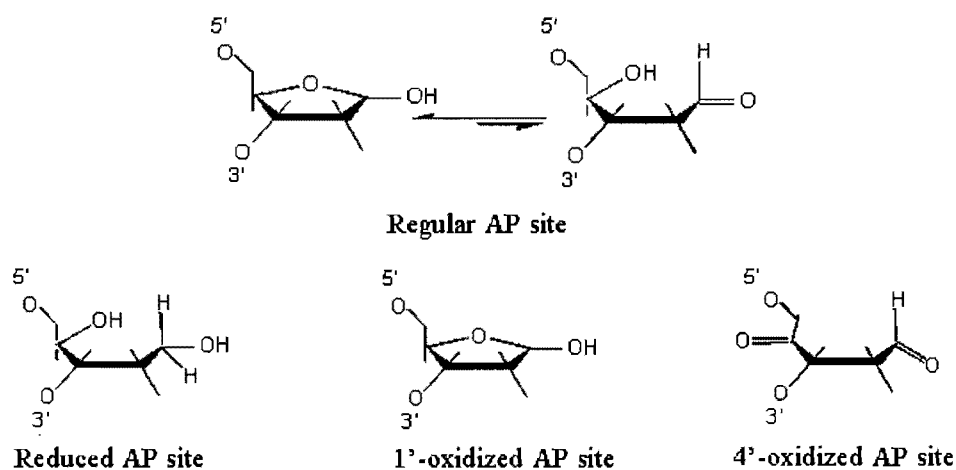
Figure 1-2: Different Forms of AP Sites (adapted from [3])

1.1. Reactive Oxygen Species

A primary source of endogenous cellular damage arises from oxidative stress. This results from an increased prevalence of oxidant molecules in the cell. Known as reactive oxygen species (ROS), these oxidant molecules are highly unstable due to an inherited unpaired electron [6]. Naturally, these reactive intermediates have an uncanny ability to react with electron rich components; notably, lipids, macromolecules and most importantly, DNA [7]. In fact, Fenton oxidants (i.e. $Fe^{2+} + H_2O_2 \rightarrow Fe^{3+} + \cdot OH + {}^-OH$)

are believed to be the most significant type of ROS causing DNA damage [2]. It should be stressed that changes made to the double helix structure require significant attention; hence, unlike other components, DNA cannot be turned over or replaced with normal type molecules. Lesions formed due to oxidative stress are associated with base alterations which can result in 8oxoG lesions, ring saturated pyrimidines, and also contribute to the formation of AP sites. ROS could also cause sugar fragmentation which can result in the formation of DNA strand breaks (with phosphoglycolate and phosphate blocked ends) [8].

ROS are produced through processes occurring during aerobic metabolism in the mitochondrion [6]. They can also be formed via a cell's interaction with exogenous sources, which may include ionizing radiation, UV radiation, carcinogenic compounds and anti-cancer drugs (Bleomycin, Neocarzinostatin) [9]. A cell can defend itself against ROS by exploiting antioxidant molecules and specific enzymes. A reduced capacity, if not an absolute failure, to counter DNA damage caused by ROS can easily hinder efficient replication and subsequently be highly mutagenic or otherwise, lethal [8].

1.2. DNA Repair Mechanisms

In order to counter DNA damage, the cell has multiple DNA repair mechanisms which ensure survival. Thus, these include mismatch repair (MMR), direct reversal (DR), double strand break repair (DSBR), nucleotide excision repair (NER), and also base excision repair (BER). Each repair pathway is responsible in countering specific types of lesions [2]. Undoubtedly, overlapping of the individual pathways is inevitable; thus, this phenomenon ensures substitute mechanisms in the repair of DNA lesions [10].

1.2.1. Mismatch Repair (MMR)

The fundamental role of the MMR system lies in the removal of mismatched bases and the insertion/deletion loops that form during DNA synthesis [11]. It is a process that is highly conserved from prokaryotes to eukaryotes. MMR, in essence, involves the removal of a patch of nucleotides (similar seen in excision repair), followed by repair synthesis and subsequent ligation [12]. It should be noted that two forms of mismatch repair systems have so far been characterized: long patch mismatch repair and short patch mismatch repair. As a result, long patch MMR is capable in repairing a vast array of mismatches by excising segments up to 1Kb long. On the other hand, short patches are characterized by an ability to handle only specific types of mismatches due to the fact that excision is restricted to a length of 1-10 nucleotides [12]. Recognition of mistmatches is carried through several enzymes which include MSH2 with MSH3 or MSH6. The complex allows the recruitment of a different group of proteins comprising of MLH1, MLH3, PMS1 or PMS2. These protein heterodimers allow the recruitment of other proteins implicated in the repair of the damaged strand [13]. Interestingly, defects in mismatch repair machinery in humans have been linked to the onset of hereditary nonpolyposis colorectal cancer [12].

1.2.2. Direct Reversal (DR)

The direct reversal system (DR) is thought to be the least complex of the DNA repair mechanisms. Unlike other pathways composed of multiple steps and subsequently, of multiple enzymes, the DR system involves a single step to remove the lesion in question. Hence, this type of quick reversal is observed primarily when thymidine dimers (caused by UV) or methylated guanines are formed [12]. Several

molecules have been characterized as DR enzymes; thus, these include photolyases and MGMT (O^6-methylguanine DNA methyltransferase) among others [14]. The functional role of photolyases in humans has not yet been clearly elucidated. However, the corresponding homologues in yeast and bacteria respectively, have shown a capacity to process damage caused by UV and cisplatin treatment. MGMT, on the other hand, is believed to be vital in processing lesions caused by alkylation [14]. Briefly, the alkyl group stemming from the lesion is removed and simultaneously shifted to the cysteine residue found in the active site of the protein. Although methyl transferases have shown to possess the ability to repair alkylated damage, the activity is known to be very energy costly. In fact, these DNA repair proteins are irreversibly inactivated following the enzymatic reaction and, subsequently, degraded by an ATP-dependent ubiquitin pathway [12].

1.2.3. Double Strand Break Repair (DSBR)

Double strand breaks (DSB), as the name implies, refers to the fragmenting of both opposing strands making up the DNA helix. These particular lesions are considered to be fairly dangerous and, consequently, may lead to harmful genome rearrangements or translocations, if not properly processed [15]. As a result, double strand breaks are countered via two principle systems known as non-homologous end joining repair (NHEJ) and homologous recombination repair (HR) [16]

Thus, in NHEJ, a specific complex of proteins joins the two fragmented ends directly without the presence of a homologous template. This process is thought to be most valuable in a pre-replication setting where no template is available [17]. NHEJ, in mammals, is initiated by a complex formed of MRE11, RAD50 and MRN amongst

other proteins. This complex enables the activation of DNA dependent protein kinase (DNA–PK) which phosphorylates several key proteins (histones). In the step, a particular DNA ligase, known as DNA ligase IV, associates with XRCC4 and XLF in order to seal the gap [18]. In contrast, homologous recombination does require the presence of a template (sharing homology to the damaged sequence) which is found on a sister chromatid (S or G2 phase) and allows the strand to be restored [15]. Undoubtedly, the HR process is known to be very efficient in preserving the integrity and identity of the sequence [19]. In yeast, HR repair involves an initial complex of Mre11, Rad50, Xrs2 as well as specific exonuclease enzymes; thus, the pathway is further mediated by Rad51 paralogs, Rad55 and Rad57. The final step (following resolution of the Holliday junction) requires, most notably, DNA polymerase and DNA ligase [20].

1.2.4. Nucleotide Excision Repair (NER)

The NER pathway is the major system utilized to counteract bulky DNA adducts or helix distorting modifications (i.e. 6-4 photoproducts and pyrimidine dimers) affecting long portions (2-20 bases) of a particular strand. Such lesions are formed via UV or specific chemical agents (aromatic hydrocarbons, cisplatin) [21]. The repair process is composed of 4 general steps starting with damage recognition, pre-incision, damage excision, gap filling and ligation [22]. Nonetheless, NER is categorized into two sub-pathways: Global genomic (GG) NER and Transcription coupled (TC) NER. These sub-pathways are differentiated by the fact that GG NER repairs all appropriate lesions in the genome; thus, including both transcribed and untranscribed segments of DNA. TC NER, on the other hand, processes damages encountered during transcription

where RNA polymerase becomes stalled at a helix distorting lesion (i.e. thymine dimer) [23].

1.2.5. Base Excision Repair (BER)

The BER pathway is a vital mechanism composed of multiple enzymatic steps that encompass the repairing of damaged DNA bases and AP sites [2, 24]. Briefly, the pathway is initiated via a DNA glycosylase enzyme which recognizes and removes the damaged base in question (Fig 1-3). The excision of the modified base from the deoxyribose-phosphate backbone forms an abasic site (AP site), which is itself a form of DNA damage that requires prompt attention to prevent further mutagenic consequences [25]. It should be noted that the resulting base loss (AP site) can also occur through spontaneous hydrolysis. In any case, the AP site is processed by either an AP endonuclease or apurinic/apyrimidinic (AP) lyase activity (processes 3' side of AP site via cleavage between C-O bond within deoxyribose) [2]. Thereafter, the single nucleotide gap is filled and sealed via the sequential action of DNA polymerase and DNA ligase [25].

The identity of the glycosylase acting on a particular base is crucial in determining subsequent steps. In essence, glycosylases are classified under two subsets: monofunctional and bifunctional. The former subset (monofunctional glycosylase) possesses a single activity which allows it to excise the damaged base [26]. The single function glycosylase primes the AP site for AP endonuclease to act on the lesion. Thereafter, AP endonuclease incises 5' to the AP site and forms a 3'-OH and a 5'-abasic sugar phosphate (deoxyribose-phosphate) [2]. In comparison, the bifunctional glycosylase is able to further process the AP site (AP lyase activity) itself by causing an

incision 3' to the lesion and forming a 3'-fragmented deoxyribose and a 5'-phosphate [27].
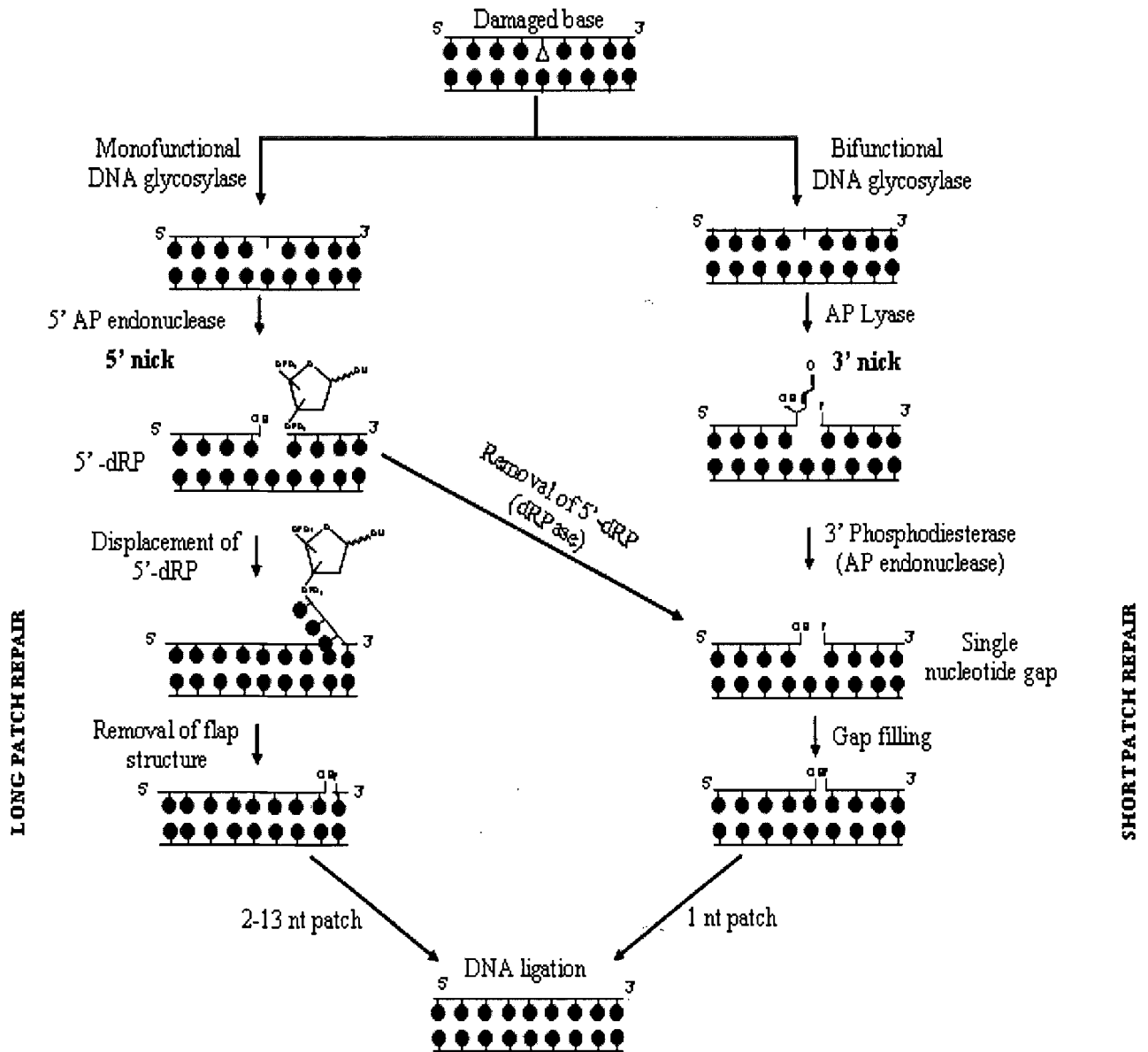
Figure 1-3: Base Excision Repair (BER) Pathway (adapted from [28])

Diagram highlights important steps in BER, including the sub-pathways (short and long patch) involved in mammalian cells. Damaged base is removed via either monofunctional or bifunctional glycosylase enzymes. The resulting AP site may be processed via AP endonuclease, forming a 3'-OH and a 5'-deoxyribose-phosphate (5'dRP) via a 5' nick. The AP

site may also be countered through an AP lyase activity, which results in a 3'-unsaturated aldehyde. The short patch repair pathway involves formation of a single nucleotide gap via the removal of the 5' dRP (following AP endonuclease activity through 5'dRPase) and in addition, the resulting repair of the unsaturated aldehyde (via 3'diesterase activity of AP endonuclease). The long patch repair pathway is associated with the formation of a flap structure which is later enzymatically removed. Both short and long patch pathways involve sealing the single nucleotide gap via DNA Ligase I or the action of DNA ligase III and XRCC1.

Activity	E. coli	S. cerevisiae	Homo sapiens	S. pombe	C. elegans
Glycosylase	ung	UNG	UDG	ung1	ung-1
	alkA	MAG	MPG/AAG	mag1	-
	nth	NTG1, NTG2	NTH1	nth1	nth-1
	fpg	OGG1, OGG2	OGG1	-	-
	nei	-	NEIL1, NEIL2	-	-
	mutY	-	MYH	myh1	R13A5.9 (wormdatabase)
	mutt	-	MTH	pcd1	ndx-1, ndx-2
			TDG		-
AP Endonuclease	xth	APN2	APE1, APE2	apn2	exo-3
	nfo	APN1	not found yet	apn1	**apn-1**
Post-incision, Gap Filling & Sealing	polI	RAD27	FEN1	rad2	crn-1
	-	POL4	POLβ	pol4	-
	holC (χ subunit)	POL2, 3	POLδ/ε	cdc20/cdc6	F33H2.5, F10C2.4 (wormdatabase)
	-	MIP1	POLG	mip1	Y57A10A.15 (wormbase)
	polIII (β)	POL30	PCNA	pcn1	pcn-1
	polIII (γ)	CDC 44 (RFC1)	RFC	rfc1	rfc-1
	ligA	CDC9	LIG1, 3	cdc17	lig-1

Table 1-1: Homologous BER Genes in Different Organisms

Sequences were retrieved from gene databases (i.e. NCBI, wormbase, etc).

http://www.ncbi.nlm.nih.gov/
http://www.wormbase.org/
http://www.genedb.org/genedb/pombe/

1.3. BER in Mammalian Cells

The completion of the BER pathway in mammalian cells, following AP site processing, can either diverge into short patch repair or long patch repair (as shown in Fig 1-3) [28]. As a result, the short patch repair path is characterized by the replacement of a single nucleotide while its counterpart, the long patch repair system, is distinguished by the replacement of 2-13 nucleotides [29].

1.3.1 Short Patch Repair

The short patch mechanism can be completed following either AP endonuclease or AP lyase action upon the AP site. In essence, the 5'deoxyribose-phosphate terminus created by AP endonuclease is processed via the deoxyribophosphodiesterase (dRPase) component of polβ (DNA polymerase β) [29]. Thereupon, the 3'-blocked end created by the bifunctional glycosylase (AP lyase activity) is removed through the 3'-diesterase component of AP endonuclease (also bifunctional). The single nucleotide gap, formed in both situations, is filled in by polβ and the resulting nick is ligated via DNA Ligase I or the concerted action of DNA ligase III and XRCC1 [27]

1.3.2. Long Patch Repair

In contrast to short patch repair, the long patch mechanism is only initiated following the action of monofunctional glycosylase and AP endonuclease. Long patch repair, in fact, has been usually associated with the processing of oxidized or reduced bases [30]. Hence, the 3'-OH group that results from AP endonuclease is extended via polymerase beta or delta/epsilon (polβ or δ/ε), and the joint action of proliferating cell nuclear antigen (PCNA) and replication factor C (RFC) [27]. The synergistic action of

these enzymes allows strand displacement to replace the missing nucleotides and forms an extensive flap structure (Fig 1-3). As a result, the formation of this structure primes the entry of flap endonuclease (FEN1) which acts to cleave the flap [30]. FEN1 action causes a strand break which is in turn sealed by DNA ligase I (LIG1) or a complex of DNA ligase III and x-ray cross-species complementing 1 (XRCC1) [31]. Evidently, the long patch mechanism is differentiated from its counterpart (short patch repair) by the fact that it requires more wide-ranging machinery [27, 32].

1.4. BER in *Saccharomyces cerevisiae*

Although the components of the BER pathway are fairly conserved, there are notable divergences that should be accounted in *S. cerevisiae*. In fact, contrary to what is observed in mammalian cells, polβ homologues in *S. cerevisiae* have not been clearly documented in the literature. However, polβ-like nucleotidyltransferase domains are present in several yeast proteins (Pol4, Trf4, Trf5) [33]. It has been discovered that homologues of long patch repair proteins in mammalian cells such as Polε (DNA polymerase ε) (DNA repair synthesis) and Fen1 (5'-flap endonuclease) are found in yeast as Pol2 and Rad27 respectively [10]. These findings indicate the presence of the long patch base excision repair pathway in yeast [33]. Recently, the discovery of Trf4 in *S. cerevisiae*, a potential homologue of polβ in mammals, possessing the 5'dRP lyase activity suggests that the short patch repair pathway may also co-exist to complement long patch repair in yeast [33].

In addition, coordination between the various proteins in yeast and mammals also appears to differ. This phenomenon can be illustrated, for example, via Ape1 and its reported direct interactions with the BER proteins, UDG and MYH (glycosyslases)

[34]. The human AP endonuclease (Ape1) has also been revealed to play a role in regulating the function of several enzymes participating in the latter stages of the pathway; thus, these include FEN1, polβ, RPA, XRCC, PCNA and LIG1[34]. Studies have well exposed the complexities encompassing the Ape1 protein which go beyond its well characterized role in DNA repair. Thus, the mammalian protein has shown to be activated by thioredoxin which in turn, enables it to act as a redox factor. The redox component of Ape1 allows the protein to activate transcription factors such as activator protein 1 (AP-1) and nuclear factor κβ (NF-κβ) by maintaining specific cysteine residues in a reduced state [35]. Interestingly, Ape1 has also been reported to be a strong activator of p53. These distinctive features associated with the mammalian AP endonuclease have not been clearly shown to exist in the corresponding yeast homologous proteins. It should be noted, nonetheless, that the 3-phosphodiesterase and 3'→5' exonuclease activities of Apn2 have been revealed to be enhanced in the presence of PCNA [36]. Undoubtedly, differences associated between the yeast and mammalian homologues, become evident in the differing range of activities that they are composed of. Notably, Apn2, unlike its mammalian homologue Ape1, is not endowed with the ability to incise altered bases in a glycosylase independent manner (Nucleotide Incision Repair) [37].

1.5. BER in *Schizosaccharomyces pombe*

Fission yeast distinguishes itself from its evolutionary counterpart, budding yeast, in the processing of AP sites. Most notably, the classical AP endonuclease members, Apn1 and Apn2, have not been reported to being efficient in countering such lesions in fission yeast [38, 39]. In fact, it was found that the AP lyase activity of Nth

addresses this situation by producing blocked 3'-α,β unsaturated aldehydes [40]. The resulting 3'blocked ends finally promote the entry of Apn2 (major AP endonuclease) which repairs the blocked end via 3'-phosphodiesterase activity. It should be noted that Apn1 and Uve1 (minor AP endonuclease) act as a 'back-up' to Apn2 [41, 42]. Subsequent downstream events follow the traditional process of nucleotide gap filling and re-establishment of the phosphodiester bond [41].

1.6. BER Key Machinery

1.6.1 Glycosylases

As described earlier, glycosylases initiate the BER pathway by removing the modified base via cleavage of the N-glycosylic bond (linking base to sugar phosphate moiety). Glycosylases are known to be small monomeric proteins that, in general, function without the necessity of a particular cofactor [26]. The glycosylase family of enzymes is divided into two groups: monofunctional and bifunctional [10, 27].

1.6.2 Monofunctional Glycosylases

Monofunctional glycosylases are known to possess a single activity, which allows for the cleavage of the modified base. UDG, from *E. coli*, is reported to be the first DNA glycosylase to be characterized [26, 43]. Evidently, corresponding homologues can be found in *S. cerevisiae* and humans. The enzyme in question is fairly restricted in its specificity; thus, it has a strong ability to recognize uracil bases and closely related products (5-hydroxy-uracil) on DNA that may be formed as a result of the deamination of cytosine or misincorporation.[44].

Another example of a well characterized monofunctional family is the Alkylbase DNA glycosylases, which excise alkylated bases resulting from both exogenous and endogenous sources [45]. These enzymes are known to be of wider specificity than its UDG counterparts. Members have been found in *E. coli* (AlkA), *S. cerevisiae* (MAG) and in human (MPG) [46-49]

1.6.3. Bifunctional Glycosylases

Bifunctional glycosylases possess an additional AP lyase activity (Figure 1-4) that is responsible for the hydrolysis of the 3' phosphodiester bond of the AP site by a β or $\beta\delta$ elimination process generating a 3'α β unsaturated aldehyde [50]. This particular group of enzymes could be further divided on their ability to remove either oxidized pyrimidines or oxidized purines respectively [51]. The former sub-group (excising pyrimidine bases) is represented by the Endo III-like glycosylases and could be found in *E. coli* (endo III or nth), *S. cerevisiae* (Ntg1, Ntg2), *S. pombe* (Nth), and *H. sapiens* (hEndoIII) [26].

An additional bifunctional glycosylase known as endo VIII has also been documented in bacteria. Thus, despite being functionally overlapping, endo VIII shares no sequence similarities with endo III. However, sequence similarities are observed when comparing endo VIII to Fpg [26, 52]. The latter sub-group, known as the Fpg family, characterized by the ability to excise oxidized purines, also has homologues in *E. coli* (Fpg), *S. cerevisiae* (Ogg1) and human (α-hOgg1 and β-hOgg1). Fpg has been demonstrated to excise fapy-ade, fapy-gua and 8-oxoG (refer to Fig1-1 for representation of lesions) [53]. Interestingly, despite having common substrate

specificity, the eukaryotic Ogg1 is not closely structurally related to the bacterial Fpg [54] [26]

1.6.4. AP Endonuclease Action in the Processing of AP Sites

AP sites are known to be one of the most common spontaneous lesions that occur in DNA. 10 000 bases a day are lost per mammalian cell [10]. In addition to spontaneous circumstances, the lesion is also formed by the enzymatic removal of a base which may have been damaged or misincorporated. Undoubtedly, AP sites are highly mutagenic and cytotoxic if not properly repaired [5]. In essence, the presence of AP sites can promote the blocking of transcription and replication [2].
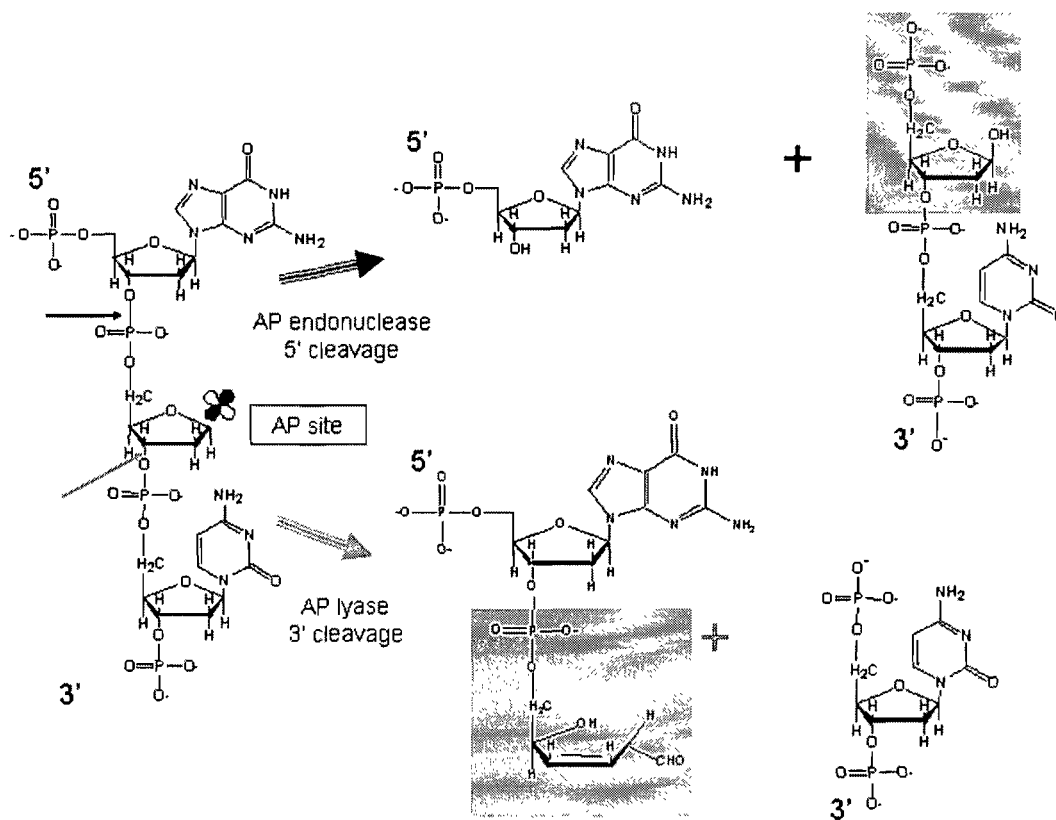
Figure 1-4: AP Site and Resulting Repair (adapted from [55])
Diagram illustrating an AP site (represented by star notation) and the immediate processing which occurs to counter the lesion; thus, this may

be achieved via a 5' cleavage stemming from an AP endonuclease activity or through a 3' cleavage provided by an AP lyase activity.

Repair of chromosomal AP sites in bacteria has been indicated to be performed by BER and to a smaller degree by a series of additional mechanisms: NER, translesion DNA synthesis and recombination [56]. It must be stressed that the BER pathway serves as the most predominant and effective route for processing AP sites [10]. The key enzymes in countering AP sites are AP endonucleases. This particular group of enzymes, as mentioned earlier, is multifunctional; thus, it is composed of AP endonuclease and 3'-phosphodiesterase components [57]. Specifically, the AP endonuclease activity is responsible for the nicking of the sugar phosphate bond at the 5' end of the AP site, producing 3'-hydroxyl group and 5'-deoxyribose phosphate. The 3'-phosphodiesterase component, in contrast, is responsible for countering 3' blocked ends produced by free radicals. It may equally come into play to counter the end products of lyase activity (3'-$\alpha\beta$ unsaturated aldehyde) to subsequently enable the formation of the polymerase responsive 3'-OH [4, 58, 59].

AP Endonucleases have been traditionally categorized into two families, the Exo III and the Endo IV, which are structurally unrelated but generally perform the same function. The classic elements differentiating both families stem from the fact that Exo III homologues are Mg^{2+} dependent enzymes, readily inactivated via the addition of EDTA while, conversely, the Endo IV family is characterized by being Mg^{2+} independent [60]. Exo III members have been found in *E. coli, S. cerevisiae, S. pombe, drosophila, C. elegans,* mouse and human. The corresponding Exo III members in *drosophila*, mouse and human are thus far the sole AP endonuclease type found in these models. In comparison, Endo IV has been shown to exist in bacteria, yeast and worms

[55]. It should be emphasized that despite being associated in large part with the same function, both families of AP endonucleases have subtle variations in their effectiveness to counter certain lesions and differ in their overall activity within a given organism (highlighted in the next section) [61]. In particular, Endo IV enzymes, unlike the Exo III family, have an ability to repair lesions induced by specific oxidants (Bleomycin and tBH) [55, 61]. As a result, although not yet found, an Endo IV member in mammalian cells may exist due to its efficiency to process these specific lesions refractory to Exo III. The existence of a mammalian Endo IV member may stem from a functional homology as oppose to a strictly structural one. Thereupon, examples of this type of relationship may be illustrated when investigating Fpg in bacteria versus Ogg1 in yeast. Although both enzymes in the corresponding organisms are not closely related in terms of structure, they do, in fact, share functional homology [26]. This thesis will focus primarily on the Endo IV family.

1.7. Endo IV Family Members

1.7.1 *Escherichia coli* endo IV

Known to be a heat stable monomeric protein, endo IV is approximately 32 kDa in size [60],[62]. It is a multifunctional enzyme composed of, among others, well established AP endonuclease and 3-diesterase activities [53]. Thus, endo IV can readily process DNA lesions consisting of 3'- phosphoglycoaldehyde, 3'-phosphate, 3'-deoxyrobose-5-phosphate and 3'-hydroxy-2-pentanal [2]. Interestingly, it has also been comprised of endonuclease activity against DNA composed of urea residues [63]. However, endo IV is considered to serve as a 'back up' or the minor AP endonuclease to exo III in *E. coli*. Data has, indeed, showed that endonuclease IV was responsible for

only 10% of AP endonuclease activity in a bacterial cell. Naturally, the remaining proportion is attributed to exonuclease III [61]. Its expression was shown to be strongly induced in the presence of superoxide anions. In fact, it was reported that paraquat, plumbagen, menadione and phenazine methosulfate, all known to form superoxide radical anions, could induce endo IV by 10-20- fold and, thus, enable the protein to attain a similar level as that of its counterpart exo III [64]. Hence, this induction effect is under the regulation of an intricate system involving the SoxR and SoxS proteins. Elevated levels of superoxide anions are reported to stimulate SoxR to later induce the expression of the SoxS protein, which in turn causes an increase in endo IV expression. It should be noted that the SoxR/SoxS proteins stimulate, in addition to endo IV, a significant array of components capable in restoring potential damage caused by unstable radicals [65].

Studies have showed that endo IV is endowed with additional activities to go along with the traditionally recognized components highlighted above. Although initially not detected, it was discovered that this enzyme comprises of a 3'$\rightarrow$5' exonuclease activity acting predominantly towards 3' recessed ends [66]. In essence, it serves as a proofreading activity allowing the removal of matching/mismatching nucleotides and altered bases (8oxoG) at cleaved ends [67]. Contrary to AP endonuclease activity, the exonuclease component appeared to be readily inhibited in the presence of EDTA and reducing conditions (even in the presence of DNA substrate) [60, 66].

Structural analyses have revealed the presence of a tri-nuclear zinc cluster forming the active site which was demonstrated to play a crucial role in its biological activities. Two of the three zinc atoms are well buried and inaccessible to solvent. The

metal factors are coordinated to key amino acids which appear to be conserved across the Endo IV family [55, 68]. The presence of the zinc atoms was reminiscent of other similar clusters involved in phosphodiester cleavage. In fact, P1 nuclease in *Penicillum citrinum* and phospholipase C in *Bacillus cereus* have revealed an ability to exploit the corresponding zinc cluster to achieve hydrolysis of the phosphate bonds [68]. Structural studies involving endo IV have also disclosed the existence of a TIM (triosephosphateisomerase) barrel fold which was identified to play a significant role in DNA binding and the recognition of AP sites. The TIM barrel has been characterized by a parallel configuration of eight β strands surrounded by eight α helices. The dynamics involved in catalytic action are thought to arise from the enzyme to bend DNA 90° and in the process, flip the abasic (AP) site [69]. The underlying structural features of the bacterial AP endonuclease parallels many characteristics observed with UVDE (UV damage endonuclease) [70]. Interestingly, UVDE plays a significant role in DNA repair by its ability to process UV-induced lesions but also acts as an AP endonuclease in certain organisms (*N. crassa, S. pombe*) [70]

Phenotype characterization has shown that a mutant lacking *nfo* (gene encoding endo IV) is hypersensitive to the antibiotic Bleomycin (BLM). This particular anti tumor drug is an established DNA damaging agent which has been displayed to produce a series of single strand breaks forming 3' blocked ends (i.e. 3' phosphoglycolate DNA terminus), double strand breaks and the release of free bases from the sugar phosphate moiety [71-73]. Cunningham *et al* (1986) reported that the single mutant showed a slight increase in sensitivity to methyl methanesulfonate (MMS), an agent readily able to alkylate DNA bases (i.e. N3 position of adenine or N7 of guanine) and indirectly cause the formation of AP sites [5, 61, 74]. The partial sensitivity may be indicative of

the fact that exo III was still being expressed and, furthermore, endo IV is not the major AP endonuclease in bacteria. Interestingly, the strain was generally unaffected, in comparison to the wild type, to hydrogen peroxide, a potent oxidant known to cause single strand breaks with 3' phosphate blocked ends. The sum of these results contrasts to the situation where a single mutant lacking *xth* (gene encoding major bacterial AP endonuclease exonuclease III), in fact, displayed greater sensitivity towards MMS, hydrogen peroxide, near UV light and X-rays. However, sensitivity of the *xth* mutant was less dramatic in the presence of the oxidizing agent *tert*-Butyl hydroperoxide (tBH) and, furthermore, not reflected at all during BLM treatment [61]. The discrepancy in survival between both strains following BLM and *tert*-Butyl hydroperoxide (tBH) exposure illustrates that endo IV and exo III are not products of redundancy. Accordingly, a double mutant lacking *nfo/xth* demonstrated a synergistic susceptibility while under presence of alkylating agents (MMS) and oxidative agents (H_2O_2, BLM, tBH) [61]. As expected from the reported data, the overexpression of endo IV re-established drug resistance (BLM, tBH) in an *xth* single mutant bacterial strain. Conversely, the overexpression of exo III in a strain void of an Endo IV member has yet to demonstrate a similar consequence in the presence of identical agents [75]. Undoubtedly, endo IV is endowed with greater versatility in countering specific types of lesions (revealed by BLM and tBH exposure) that exo III is unable to match .

Additional studies have shed significant light to further the divergences that exist between endo IV and its exo III counterpart. For example, endo IV, unlike exo III, has been associated to a newly characterized mechanism complementing BER now known as Nucleotide Incision Repair (NIR). This pathway is differentiated by the fact that it is a glycosylase independent process, whereby the protein in question incises the

altered base and the deoxyribose component to subsequently promote eventual DNA synthesis. The NIR activity of endo IV allows the enzyme to process a diverse array of lesions stemming from oxidative stress: 5,6-dihydro-2'-deoxyuridine (DHU), 5,6-dihydrothymidine (DHT), 5-hydroxy-2'-deoxyuridine, 2 ,6-diamino-4-hydroxy-5-N-methylformamidopyrimidine and α-anomeric 2'-deoxynucleotides (αdA and αdT) [67]. Moreover, endo IV has also been associated with providing a protective effect to circumvent potential oxidative damage caused by UVB [76]. Unlike its counterpart exo III, endo IV is unable to directly repair the lesions caused by UVB. It should be noted that oxidative damage formed as a result of UVB irradiation is believed to be associated with AT to GC transitions and GC to CG transversions [76].

1.7.2. *Saccharomyces cerevisiae* Apn1 (ScApn1)

ScApn1 has been reported to be the major AP endonuclease enzyme in *S. cerevisiae*. Data has revealed that it accounts for 97% of the AP endonuclease and 3'-diesterase activity in the corresponding eukaryotic cell. The 42 kDa protein shares 41% amino acid identity with the bacterial endo IV and composed of a similar zinc cluster (3.3 atoms) found in the active site [77]. It should be stressed that like the bacterial endonuclease IV, ScApn1 does not necessitate an additional metal co-factor for its reported activities. Moreover, ScApn1 is able to act on a similar range of substrates as its counterpart in bacteria; thus, these include 3'-phosphoglycoaldehyde, 3'-phosphoryl groups and 3'-α,β-unsaturated aldehydes [2]. However, unlike endo IV, ScApn1 is endowed with a longer C-terminal end which is in fact shown to be composed of 3 important clusters (cluster 1, 2, and 3). These series of domains, rich in basic amino acids, have been elucidated to serve as an NLS. Cluster 1, in fact, was reported to being

essential for nuclear localization. Thus, the removal or replacement of cluster 1 may not affect enzyme activity in-vitro; however, it abolished the localization process to the nucleus. The substitution of cluster 1 with the NLS of SV40 (simian virus 40) did not appear to restore proper nuclear targeting. Furthermore, it was observed that neither cluster 1 or 2, on its own, was able to allow ScApn1 to be signaled to the nucleus and moreover, the concurrent removal of both produced the same effect. Thus, it was postulated that cluster 1 and 2 behave as a bipartite NLS [78]. Experiments have shown that the bipartite cluster also plays a role in targeting the protein to the mitochondrion. In fact, this phenomenon appears to be mediated via an interaction with the protein Pir1 [79]. In the absence of Pir1, ScApn1 accumulates in the nucleus and cytoplasm. In contrast, deletion of a significant portion of the C-terminal domain contributes to subsequent accretion of the protein in the cytoplasm only. [79].

Classical phenotypic analysis has revealed that ScApn1 single mutants are hypersensitive to MMS (alkylating agent) as well as to the oxidants tBH and hydrogen peroxide. As expected, DNA collected following treatment with the particular alkylants and oxidants has showed a striking accumulation of DNA lesions. Spontaneous mutation rates of Apn1 null mutants increased from 6 to 12 fold in comparison to the wild type [80]. Moreover, mutants have been correlated with a 59 fold increase in A.T to C.G transversion events [81]. The sum of these results displays the importance of Apn1 in protecting the genome against a wide array of lesions. Nonetheless, the single Apn1 mutant failed to show sensitivity following Bleomycin and Phleomycin treatment; thus, this piece of data contributed to the suggestion of a possible existence of an additional 3'-diesterase in *Saccharomyces cerevisiae* [80, 82]. Hence, Apn2 (Exo III family) was subsequently discovered and demonstrated to possess the same bifunctional

nature as its Endo IV counterpart. However, it should be noted that Apn2 remains the minor AP endonuclease in yeast [5]. As such, this conclusion rests on the evidence that an *apn2* single mutant fails to demonstrate a dramatic hypersensitivity towards the same agents (MMS, H_2O_2) as was observed with the *apn1* mutant. The increased mutagenesis associated with single *apn1* null mutants reveal that Apn2 lacks the ability to optimally substitute for its AP endonuclease counterpart [5]. Double mutants void of both *APN1* and *APN2* did, however, reveal an extensive sensitivity to MMS, H_2O_2 and Phleomycin that transcends what is observed with the respective single mutants [82] [5].

Recent investigations concerning Apn1 have found that akin to its bacterial homologue, it also is composed of a $3' \rightarrow 5'$ exonuclease component which complements the well established AP Endonuclease/3-diesterase activities. Thus, this newly associated function is characterized by a strong preference for duplex DNA with 3' recessed ends [67]. The exonuclease activity has been, interestingly, showed to permit the elimination of 8oxoG lesions (formed at the 3' end of gap) resulting from the direct oxidation of guanine or a misincorporation by DNA polymerase. The ability of Apn1 to process 8oxoG residues prevents G.C to T.A transversions to become prevalent in *ogg1* single mutants. As a result, it is believed that the Endo IV member in *Saccharomyces cerevisiae* may act as an alternative pathway to Ogg1 in the removal of 8oxoG from DNA. Apn1 has also shown to reflect its bacterial homologue in its ability to take part in NIR and process α-anomeric 2-deoxynucleotides (αdA and αdT) and other oxidative lesions [67]

ScApn1 has recently garnered important attention in its potential benefit to humans. Thus, the expression of Apn1 has showed an ability to counter damage instigated by ROS in the development of certain neurodegenerative diseases [83]. In

fact, the production of Apn1 in mammalian cells revealed a decrease in sensitivity following oxidative stress [83].

1.7.3 *Schizosaccharomyces pombe* Apn1 (SpApn1)

Contrary to the situation in *S. cerevisiae*, SpApn1 is not considered the major AP endonuclease in *S. pombe*; a title which is rather held by Apn2. Contrary to the *apn2* mutant, data has revealed that the *apn1* mutant failed to demonstrate sensitivity to the alkylating agent MMS [39]. An additional minor AP endonuclease activity is found in *S. pombe* stemming from the Uve1 enzyme, a previously established component of an alternate UV damage repair pathway to NER [42]. It has been reported by Fraser *at al.* that spontaneous mutations which occur in *uve1* null cells (*S. pombe*) are similar to the transversion mutations induced by 8oxoG in *S. cerevisiae*. This evidence first suggested the role of Uve1 in oxidative DNA damage repair in fission yeast [42]. However, SpApn1 and SpUve1 proteins were reported to possess AP endonuclease activity *in-vitro* but failed to initiate AP site repair *in-vivo* [41]. As discussed earlier (section 1.6), the direct processing of AP sites in *S. pombe* has been predominantly reported to be the domain of the Nth enzyme (bifunctional glycosylase) via its AP lyase activity (producing 3'α,β-unsaturated aldehyde ends) and not a member of either the Exo III or Endo IV families. The implication of such proteins, belonging to the respective AP endonuclease families, has thus far been restricted to the repair of 3' blocked ends (processed predominantly by Apn2, with Apn1 and Uve1 serving as back up components) induced in large part by Nth [41]. Undoubtedly, there are profound differences which exist between *S. pombe* and *S. cerevisiae* in the identity of the machinery having access to the lesions in question. These divergences are further

revealed through the expression and resulting complementation of ScApn1 in the repair of both AP sites and blocked ends in mutant fission yeast strains ($nth\Delta$ and $apn2\Delta$ single mutants) when exposed to MMS [41]. Interestingly, localization studies using GFP have revealed that SpApn1 is distributed in the nucleus and, in addition, the cytoplasm (not observed with SpApn2); thus, suggesting that it may possibly play additional roles than simply being involved in maintaining genomic stability [41].

1.7.4. *Caenorhabditis elegans* CeAPN-1

The gene encoding the *C. elegans* Endo IV AP endonuclease, CeAPN-1, was first identified following the utilization of a radiolabelled oligonucleotide (representing the sequence of yeast *APN1*) which was probed against a cDNA library of the multicellular organism [84]. The expressed protein was initially reported to be composed of 396 amino acids but subsequent findings have reported that it is even larger (discussed below). Nonetheless, CeAPN-1 shares considerable homology with Endo IV members found in the unicellular organisms highlighted. Hence, it shares 40.4% identity with the yeast *APN1* and 44.9% identity with the bacterial *nfo* [84]. Much of the homology is concentrated in 5 distinct domains, thought to represent, notably, portions of the DNA binding domain and active site. Interestingly, *C. elegans* is the first higher eukaryote, representing an animal model, which has been reported to have both Endo IV (CeAPN-1) and Exo III (CeEXO-3) homologues. The latter member (CeEXO-3) has been well characterized to being Mg^{2+} dependent and endowed with AP endonuclease/3'diesterase activities. However, it has also been found that the metalloenzyme is composed of a number of unique features which distinguishes itself from its Exo III counterparts; notably, a GST fusion CeEXO-3 protein failed to display

an important 3'→5' exonuclease activity following AP sites (opposite observed with bacterial and yeast homologues) and furthermore, it revealed to lack the ability to participate in Nucleotide Incision Repair (as observed with human Exo III, Ape1) [85]. However, contrary to CeEXO-3, very little is known about CeAPN-1. Numerous attempts have been made to express and purify the full length protein for characterization; however, these past efforts have not been as fruitful [55].

1.8. Research Project

1.8.1. Background and Objectives

Attempts have been made in developing an expression system in *E. coli* to allow for the isolation of CeAPN-1. This was performed by constructing a plasmid containing the Ceapn-1 gene under the control of an isopropyl β-D-1-thiogalactopyranoside (IPTG) inducible promoter. Unfortunately, the expressed product detected via immunoblot appeared to be smaller (10 kDa) than the expected size. Hence, a possible proteolytic process in bacteria may have been hindering the detection of the full length protein [55].

Due to the inability to detect the full length CeAPN-1 in bacteria and the fact that a sufficient yield (for characterization experiments) could not be obtained via *C. elegans*, an expression system was engineered in yeast [86]. This particular eukaryotic model has multiple advantages stemming from the fact that it is easy to manipulate and grow. Moreover, it should be recognized that a significant benefit arises from the fact that a mutant strain void of AP endonuclease, YW778 (*tpp1Δ::MET15 apn1Δ::HIS3 apn2Δ::KanMX4*), was readily available. Through the subsequent transformation of this strain, one could easily ascertain whether the introduction and expression of Ceapn-1 can functionally complement in the presence of DNA damaging agents [86].

It was shown that the yeast plasmid containing the Ceapn-1 gene, pNLS-GAD-Ceapn-1 (recovered from a *C. elegans* cDNA library), was able to substitute for yApn1 in the presence of MMS when introduced in the YW778 strain. As a result, a new construct was designed in order to equip the expressed product with a GST affinity tag for purification purposes [86]. Unfortunately, unlike the previous attempt, the introduction of pGST-Ceapn-1 failed to restore MMS resistance to the AP endonuclease deficient strain. An identical approach exploiting a GST tagged Ceexo-3 construct introduced into YW778 (*tpp1Δ::MET15 apn1Δ::HIS3 apn2Δ::KanMX4*) (void of AP endonuclease activity) did, however, display an ability to complement in the presence of MMS. Thus, it was hypothesized, at the time, that the failure of GST-Ceapn-1 to restore resistance was likely due to the GST tag that may have been hindering the activity domain of the protein [86]. It was, however, later discovered (during the course of the current project) that the Ceapn-1 sequence was more extensive than originally believed to be. Updated and better annotated sequences submitted to gene databases revealed that Ceapn-1 is composed of an additional 118 amino acids at the N-terminal which was not previously accounted for. Upon close inspection, the additional residues have been pointed out to being highly rich in basic residues (characteristic of NLS) which may be indicative that it may play a role in protein localization [87].

As a result, this project set out firstly to develop an effective expression system for CeAPN-1, secondly, to isolate a homogeneous fraction of the protein and, thirdly, characterize its precise biological function in an *in-vitro* setting. The goal in characterizing CeAPN-1 rests primarily on determining the types of lesions (AP sites, 3'-blocked ends, oxidized bases) it can process, evaluate its range of activities, assess whether additional co-factors are needed for its function and to differentiate it from its

counterparts. Fourthly, it was also our intention to conduct a domain function analysis, through site directed mutagenesis, in order to better understand the characteristics that make up CeAPN-1 and specifically, shed light upon some significant amino acid residues that contribute to its activities. Thus, we confirmed that the introduction of the tagged proteins, GST-CeAPN-1-HIS and GST-HIS-CeAPN-1, can respectively rescue a strain deficient in AP endonuclease activity while in the presence of DNA damaging agents (MMS, Bleomycin, H_2O_2). Partially purified fractions containing the recombinant CeAPN-1 protein revealed that the AP endonuclease and 3' diesterase components were retained. However, fractions obtained following full purification exposed a dramatic decrease in protein activity. Interestingly, it was demonstrated that the addition of crude extracts stemming from the YW778 strain (free of AP endonuclease activity) caused a significant increase in the activity level. Collected data has suggested that a potential heat stable factor may be stimulating the activities of the Endo IV protein under study; nonetheless, its true identity has yet to be determined. Furthermore, domain function analysis confirmed the significance of the E143 and E97 residues in governing the ability of the enzyme to remain functionally active. Thus, the variants, E143G and E97G, failed to demonstrate a capacity to substitute for yApn1 in the presence of DNA damaging agents. The results observed are in accordance with data previously obtained involving the study of charged residues contained in one of the four conserved Endo IV family domains affecting activity [88] [79].

2. METHOD AND MATERIALS

2.1. Yeast and Bacterial Strains

The *S. cerevisiae* strains utilized in this study were the parent YW465 (*MATalpha ade2Δ0 his3Δ-200 leu2Δ-1 met5Δ0 trp1Δ ura3Δ0*) and its isogenic mutant derivative, YW778 (apn1Δ::HIS3 apn2Δ::KanMX4 tpp1Δ::MET15). Both strains were obtained previously from Dr. Tom Wilson (University of Michigan) and preserved on YPD agar supplemented with adenine 20 μg/ml. The *E. coli* strain used in this study was DH10β which was maintained on LB agar.

2.2. Plasmid Construction

Name	Sequence (5'-3')
NLS-His-Ceapn-1-F1	AACTCCAAGCTTATGCCCAAGAAGAAGCGGAAGGTCTCGAG CGGCCATCATCATCATCATAAAAAAAAATCTCGTGAAAC G
Ceapn-1-R1	TTATCTTTTATCCATATTGTACAT
vector-NLS-Ceapn-1-F1	AACTCCAAGCTTATGCCCAAGAAGAAGCGGAAGGTCTCGAG CGGCAAAAAAAAATCTCGTGAAACGGTTGGCGTGGAAG
Ceapn-1-His-vector-R1	TTGATCATTCATCAGTAAAAAAGTTATGCAATTAAACAAGAA TTAATGATGATGATGATGATGTCTTTTATCCATATTGTACATG
GST-NLS-Ceapn-1-F	TTCGATGATGAAGATACCCCACCAAACCCAAAAAAAGAGAT CGAAATGCCCAAGAAGAAGCGGAAGGTC
GST-NLS-Ceapn-1-R	TTCAGTATCTACGATTCATAGATCTCTGCAGGTCGACGGATC CCCCGACGTCTTACTTACTTAGCGGCCG
E143G region2–F1	ATTATTTTAGTTCTCGGGACAATGGCTGGACAAGGG
E143G region2–R1	CCCTTGTCCAGCCATTGTCCCGAGAACTAAAATAAT
E97G region1–F1	CGAATGTCAACGGGCTGGGAAATTAGGAATCACG
E97G region1–R1	CGTGATTCCTAATTTCCCAGCCCGTTGACATTCG
Oligo U	GCTGCATGCCTGCAGGTCGAUTCTAGAGGATCCCGGGTACCT
Oligo G	AGGTACCCGGGATCCTCTAGAGTCGACCTGCAGGCATGCAGC

Table 2-1: Oligonucleotides Utilized in Study

pAS3A (pPC86 from Invitrogen *C. elegans* cDNA library) containing the *C. elegans* apn-1 gene was used as a template for PCR (refer to Fig 3-1). The primers (NLS-His-Ceapn-1-F1, Ceapn-1-R1) were designed to allow the insertion of a His tag at the flanking end of the gene and, thus, enable the simultaneous removal of the large GAD sequence (refer to fig 3-1). The resulting product was purified via electrophoresis (1% agarose TAE gel) whereby the corresponding band signal was excised and purified via gene cleaning kit (Amersham Piscataway, NJ). Additional pAS3A plasmid was meanwhile digested with *SalI* and *NcoI*. The linearized product was recovered through electrophoresis. Hence, the band representing the cut plasmid was excised and purified as mentioned above. The desired construct, pHis-Ceapn1, was obtained via gap repair whereby a wild type strain was co-transformed with the PCR generated His-Ceapn-1 product and the digested pAS3A plasmid. Resulting transformants were selected on – Trp agar plates (plasmid carries a Trp marker).

A second plasmid, pCeapn1-His, characterized by a C-terminal His tag was designed by exploiting pHis-Ceapn-1 as a PCR template for the primers, vector-NLS-Ceapn-1-F1 and Ceapn1-His-vector-R1 to obtain a final product consisting of NLS-Ceapn-1-His (refer to Fig 3-2). The PCR fragment was purified via conventional electrophoresis means (as mentioned above). Concurrently, pAS3A was digested using the same steps as described above. pCeapn1-His was obtained in the same fashion as pHis-Ceapn-1; thus, involving the co-transformation of the NLS-Ceapn-1-His fragment and the cut plasmid in the YW465 strain. Transformants were similarly selected on – Trp agar plates.

The third plasmid, pGST-His-Ceapn-1, which allowed the expression of two affinity tags was constructed using the previously obtained pHis-Ceapn-1 as the PCR template and the primers, GST-NLS-Ceapn-1-F and GST-NLS-Ceapn-1-R, to obtain a fragment consisting of GST-NLS-His-Ceapn-1 (refer to Fig 3-3). The pTW340 plasmid (bearing the gene *OGG1* with a GST tag) was digested (*SmaI* and *EcoRI* to remove *OGG1*) in parallel, and thereafter, used in the final gap repair step with the PCR product. It should be noted that transformants were selected on –Ura plates.

The final plasmid in question, pGST-Ceapn1-His, was designed in the manner as pGST-His-Ceapn-1 with one modification; hence, the pCeapn-1-His plasmid was exploited as the PCR template (refer to Fig 3-4). Therefore, the plasmid has an N-terminal GST tag and a C-terminal His tag (refer to Fig 3-4).

2.3. Isolation of Plasmid DNA and Transformation:

All techniques involving the extraction of plasmid DNA and transformation were performed by following the protocols given by Sambrook *et al* (1989) [89].

2.4. Expression System for Recombinant Proteins:

The respective plasmids (pCeapn1-His, pHis-Ceapn1, pGST-His-Ceapn1 and pGST-His-Ceapn1) were utilized to transform YW778. Resulting transformants stemming from the introduction of pCeapn1-His, pHis-Ceapn1 and pGST-His-Ceapn1, GST-His-Ceapn1 were selected on –Trp and –Ura plates respectively. The selection plates consist of minimal medium agar (lacking the corresponding tryptophan or uracil). The transformants were grown in the appropriate liquid minimal media for proliferation.

2.5. Gradient Plate Assay:

Gradient plate analysis was conducted as explained in previous trials [61, 78]. In essence, cultures were grown overnight in selective minimal media and thereafter, replicated on a plate (YPD media) containing the DNA damaging agent gradient. The plate was incubated for 48 hours at 30°C and later photographed.

2.6. Extraction of Proteins for Expression Analysis:

Cultures were grown overnight and subsequently, harvested from liquid minimal media via centrifugation at 4000 rpm (Sorval RT-7). The supernatant was discarded and the cell pellet was washed twice with water. The resulting pellet was thereafter resuspended in buffer A (50 mM Tris-HCl, 50mM NaCl, 10% glycerol) with protease inhibitor cocktail (Complete Mini, EDTA-free, Roche Diagnostics GmbH). Resuspended cells were beated with 425-600 μm glass beads (Sigma, St Louis, MO) via the utilization of a Mini-Beadbeater (BioSpec Products, Bartlesville, OK) at 4800 rpm for 20 seconds. It should be noted that the beadbeating step was repeated 5 times. Cell debris was separated through centrifugation which consisted of 5 minutes at 1000 rpm at 4° C using a tabletop microcentrifuge (Jouan, Inc. Winchester, VA). Extracts were isolated and quantitated via the traditional Bradford method [90].

2.7. Extraction of Proteins for Purification:

Cultures were grown over 3 days in minimal media. Thus, culture was initiated in 10 ml of minimal media, sub-cultured the second day in 100 ml of minimal media and finally added to 900 ml of fresh media. The 1 L culture was allowed to grow overnight. Thereafter, cells were harvested via centrifugation at 4000 rpm (Sorval RT 7)

and the resulting pellet was thoroughly washed three times with water. The cell pellet was aliquoted and stored at -80° C. However, for extraction purposes, 10 ml of cell pellet was resuspended in an equal volume with buffer A with protease inhibitors and bead beated via Maxi-beadbeater (BioSpec Products, Bartlesville, OK). It is important to note that the beating container was filled 70% with glass beads. Subsequently, the resuspended cells were added to the beads while ensuring no air space was remaining in the holder (added buffer A to ensure the container was filled). Cells were incubated in ice for 2 minutes and, thereafter, beated for 20 seconds; this process was repeated 20 times. Resulting extracts were centrifuged at 18 000 rpm for 10 minutes while at 4° C (Sorval SS34). Thereupon, the supernatant (protein extracts) was collected and stored at -20 C.

2.8. Site Directed Mutagenesis:

Two site directed mutations were performed by the Quickchange kit (Stratagene, La Jolla CA). The substitutions, E143G and E97G, were individually made via PCR using the corresponding polyacrylamide purified primers, (E143G region2–F1, E143G region2–R1 and E97G region1–F1, E97G region1–R1), and the pGST-His-Ceapn-1 plasmid as the template. The obtained PCR products were incubated with *Dpn1* and later utilized in the transformation of the bacterial DH10β strain. Transformants were selected on LB-Ampicilin plates and the plasmid DNA was later isolated and sequenced. Finally, the sequenced plasmids containing the corresponding mutations were used to transform YW778. Transformants were isolated on –Ura minimal media agar plates.

2.9. Protein purification Through Using a His-Affinity (Talon) Column:

Column resin was equilibrated with 20 volumes of Buffer A. Protein extracts were loaded into column and recycled 5 times through gravity. Following binding, resin was washed with 5 volumes of buffer A and 15 volumes of buffer B (50 mM Sodium Phosphate, 300 mM NaCl) at pH 7. Elution process was achieved by collecting five 200 µl fractions buffer C at pH 7 (Imidazole Elution buffer as given by Clontech resin manufacturer) which is composed of 50 mM Sodium phosphate, 300 mM NaCl, 150 mM Imidazole. Final elution fractions contained 20% glycerol.

2.10. Purification Using a Glutathione Sepharose Column:

To achieve purification of GST tagged proteins, collected extracts were diluted 50% with additional buffer A (containing protease inhibitors) and Triton-X100 before resin binding. The GST tagged proteins were purified using a glutathione-sepharose resin (GE-Healthcare) following general instructions specified in manufacturer handbook. Briefly, the resin was centrifuged at 500 x g to remove storage solution and subsequently washed with 1x PBS (140 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4) five times. The resin was centrifuged following each wash and subsequently incubated at room temperature (in order to resuspend) on a rocker for 5 minutes. Thereupon, protein extracts were added to the equilibrated matrix and subsequently rotated for 40 minutes (room temperature). Following incubation, extracts were spun down (500 x g) and the unbound supernatant was collected and saved at - 20°C. The resin was washed 5 times with 1x PBS + 1% Triton on a rotator and centrifuged (500 x g) at 4°C. Elution was performed on a column at 4°C. Hence, 5 x 1 ml fractions were collected following the addition of elution buffer (50 mM Tris-HCl

pH 8, 10 mM reduced glutathione). Glycerol was added to each fraction to a final concentration of 20%. The eluted fractions were stored at -20°C.

2.11. SDS PAGE Staining and Western Blot

Crude protein extracts and eluted fractions were separated on a SDS PAGE gel containing 10% acrylamide,/bisacrylamide through electrophoresis .

2.11.1 Sypro Staining:

Obtained gel (post electrophoresis) was fixed in solution D (10% MeOH, 7% Acetic acid) for 30 minutes and stained with Sypro-Ruby (Biorad, Hercules CA) for 3 hours. Thereupon, the gel was destained with solution E (10% MeOH, 7% Acetic acid) for 90 minutes while changing solution 3-4 times until appropriate resolution was achieved.

2.11.2 Western Blot:

Transfer was performed on a nitrocellulose membrane (Amersham, Piscataway, NJ). The membrane was incubated overnight with 0.1% anti-HIS (for single HIS fusion protein) (Santa Cruz Biotechnology, Santa Cruz, CA) and 0.05 % anti GST and later probed with a secondary antibody (anti-mouse for anti-His, anti-rabbit for anti-GST). Resulting image was developed following exposure with chemiluminescence solution (Perkin Elmer, Woodbridge, ON).

2.12. Preparation of AP Endonuclease Substrate:

The 42mer, oligo U (5'-GCTGCATGCCTGCAGGTCGAUTCTAGAGGA-TCCCGGGTACCT-3') (3.5 ng) was labelled with [γ-^{32}P]ATP using T4 polynucleotide kinase (USB, Cleveland OH) as performed previously [91]. The end labelled

oligonucleotide obtained following ethanol precipitation was subsequently, annealed to a complementary oligonucleotide G (3'-CGACGTACGGACGTCCAGCTGAGATCTCCTAGGGCCCATGGA-5' as previously performed [84]. The newly formed double stranded oligo was later treated with 25 units of Uracil DNA glycosylase (UDG) in buffer F (30 mM HEPES-KOH, pH 7.6, 50 mM NaCl, 2 mM EDTA) for 30 minutes at 37°C to form the AP site substrate. The reaction was terminated via a 10 minute heat treatment at 65°C [88, 92].

2.13. Preparation of 3' Diesterase Substrate:

The 3'diesterase substrate was prepared as previously described [93]. Briefly, the previously obtained oligonucleotide (AP endonuclease substrate) was treated with 25 units of Endo III at 37°C for 60 minutes. The reaction was terminated via a 10 minute heat treatment at 65°C [91].

2.14. Preparation of YW778 Extracts for Activity Assays:

Extracts were prepared in the same fashion as described earlier for expression analysis. For heat treatment experiments, however, small aliquots of collected extracts were heated in 1.5 ml eppendorf tubes at 95° C in a water bath for 5 minutes. Extracts were thereafter spun down on a tabletop centrifuge (Jouan, Inc. Winchester, VA) at 1000 rpm whereby the soluble fraction was collected. For dialysis experiments (evaluating AP endonuclease activity), the soluble fraction (100 μl volume) was dialyzed against 500 ml of buffer A (20% glycerol) at 4°C. The last step was repeated three times (to exchange buffer): twice following 4 hours and maintained overnight during third trial.

2.15. Enzyme Assays:

Reactions were performed using buffer G (50 mM Hepes-KOH at pH 7.6, 50 mM KCl, 1 mM EDTA, 100 µg of BSA) in a final volume of 12.5 µL. The substrate (1.5 ng) was incubated with 1 ng of protein (unless otherwise indicated). Reactions were terminated with the addition of 5 µL formamide loading buffer (76% formamide, 10mM EDTA, 0.3% bromophenol blue, 0.3% xylene cyanole). Thereafter, the samples were heated to 65°C for 5 minutes to separate any proteins which may have remained bound to the DNA. Reactions were, subsequently, loaded onto a 10% gel and 17% polyacrylamide-Urea gel respectively for the AP endonuclease and 3'-diesterase assays respectively. Following 1 hour migration, the gel was transferred to Whatman paper, wrapped in plastic and placed on a dryer (Biorad, Hercules CA) for 1-2 hours. Thereafter, the dried gel was placed in a cassette (GE Healthcare) with screen for overnight exposure. The image was developed via Storm Scanner (GE Healthcare).

2.16. Homology modeling of CeAPN-1

The deduced protein sequences of Ceapn-1 was submitted to the Swiss-Model server (Automated Comparative Protein Modeling Server, Version 3.5) [94, 95] for comparative protein structure modeling. All homology models were subsequently generated based on the template of *E. coli* endonuclease IV (PDB accession number: 1QTW and 1QUM; [68]) using Swiss-PdbViewer 3.7 [95, 96]. The electrostatic potential and molecular surface of each protein model was calculated through eF-surf server and visualized through PdbViewer (Version 3.0) [97].

3. RESULTS:

3.1. Plasmid Construction:

The pPC86 plasmid (refer to method and materials) carrying the *C. elegans* apn-1 gene (plasmid re-named pAS3A5) was utilized as a PCR template to amplify the gene in question. The primers, v-NLS-His-Ceapn1-F1 and Ceapn1-v-R1 were designed to allow the insertion of a His tag between the NLS and the gene and simultaneously 'loop out' the large GAD sequence (separating the NLS and apn-1 coding sequence) found within the plasmid. Thereupon, a resulting fragment composed of NLS-His- Ceapn-1 was obtained and subsequently co-transformed in YW465 with linearized pAS3A. The yeast system allows the fragment to recombine with the digested plasmid via gap repair.
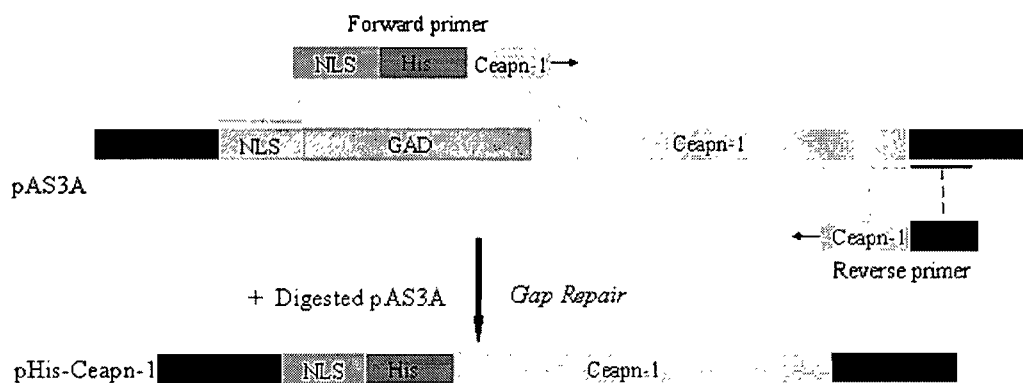
Figure 3-1: Construction of Ceapn-1 Expression Plasmid with N-Terminal His Tag

Ceapn-1 was retrieved from a cDNA library using the primers shown (v-NLS-His-Ceapn1-F1 and Ceapn1-v-R1) above. The gene was subsequently cloned into a construct where it was fused at the N-terminal to a His affinity tag (forming pHis-Ceapn-1).

A second plasmid was designed throughout the course of this project which is characterized for the His tag being at the C-terminal end of the gene. In essence, the

pHis-Ceapn1 plasmid was used as the template for the designed primers v-NLS-Ceapn-1-and Ceapn1-His-v. These primers would allow the simultaneous 'looping out' of the original N-terminal His tag and resulting incorporation of the tag to the C-terminal.
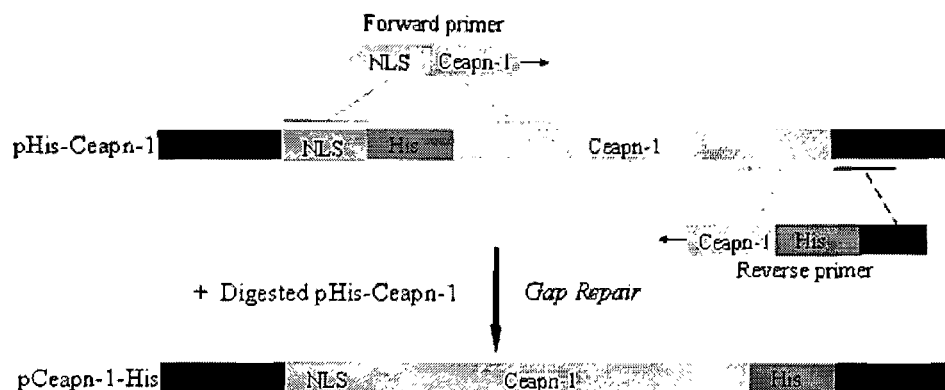
Figure 3-2: Construction of Ceapn1 Expression Plasmid with C-Terminal His Tag

A second plasmid was constructed using specifically designed primers (v-NLS-Ceapn-1-and Ceapn1-His-v) to allow the His affinity tag to be fused at the C-terminal of the gene. The plasmid, pHis-Ceapn-1, was used as a template for the primers utilized.

Both plasmids shown above were separately utilized to produce a dual tagged construct (composed of both His and GST affinity tags). Hence, primers were carefully designed to allow utilization against both templates and concurrently produce the constructs pGST-His-Ceapn-1 and pGST-Ceapn-1-His.

Forward primer

Gst   NLS →

pHis-Ceapn-1 ▨▨ NLS ▨▨ His ▨▨ Ceapn-1 ▨▨ ▨▨

Reverse primer

+ Digested pTW340 | *Gap Repair*

pGST-His-Ceapn-1 ▨▨ Gst NLS His Ceapn-1 ▨▨
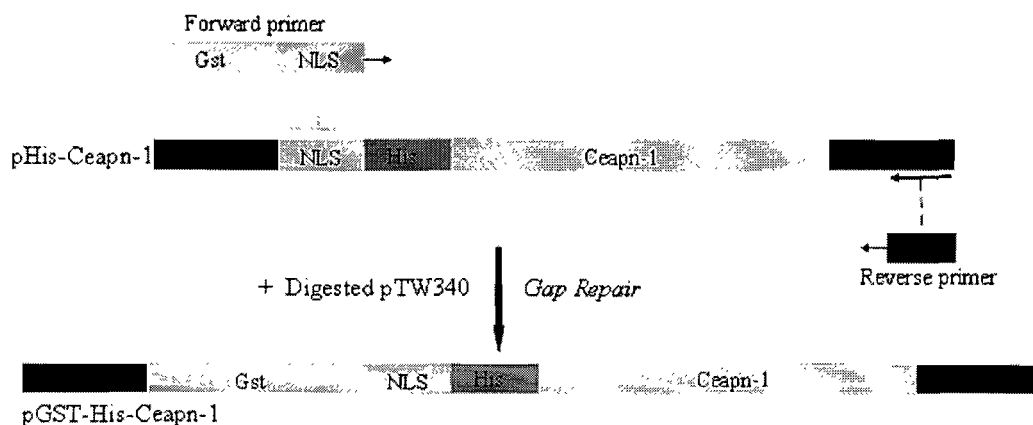
Figure 3-3: Construction of Ceapn-1 Expression Plasmid with Dual N-Terminal GST and His Tags

A dual tagged plasmid consisting of GST and His affinity sequences were inserted in frame at the N-terminal of the Ceapn-1 gene. The PCR product was obtained using the primers, GST-NLS-Ceapn-1-F and GST-NLS-Ceapn-1-F, and pHis-Ceapn-1 plasmid (section 2.1).

Forward primer

Gst   NLS →

pHis-Ceapn-1 ▨▨ NLS ▨▨ Ceapn-1 ▨▨ His ▨▨

Reverse primer

+ Digested pTW340 | *Gap Repair*

pGST-Ceapn-1-His ▨▨ Gst NLS Ceapn-1 His ▨▨
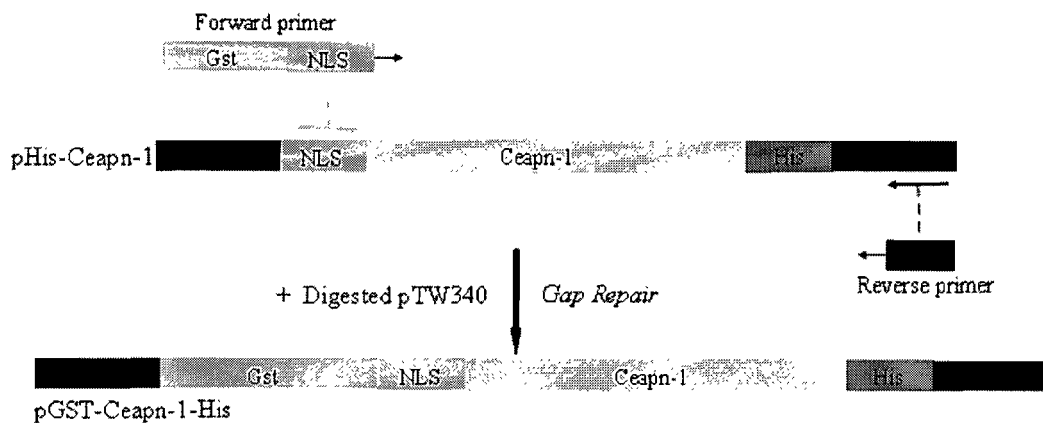
Figure 3-4: Construction of Ceapn-1 Expression Plasmid with dual tagged N-terminal GST and C-terminal His

A second dual tagged plasmid was concurrently constructed consisting of affinity sequences at opposing ends of the Ceapn-1 gene. Cloning was accomplished using the pCeapn-1-His plasmid and the GST-NLS-Ceapn-1-F and GST-NLS-Ceapn-1-F primers.

3.2. Cross Species Complementation:

pHis-Ceapn-1 was the first plasmid to be constructed and subsequently introduced into AP endonuclease deficient YW778. To assess whether the protein was expressed and able to rescue the mutant yeast strain (YW778) from DNA damaging agents, a gradient plate assay was performed. The YW778 strain (lacking both Endo IV and Exo III components) is hypersensitive to MMS (induces AP sites) and, thus, unable to process the created AP sites. Consequently, it grows only a very short distance along the increasing gradient composed of the DNA damaging agent. In contrast, a wild type strain (YW465), equipped with AP endonuclease activity, is able to resist damage induced by MMS and sustains growth along the entire gradient.

Figure 3-5 shows that the introduction of pHis-Ceapn-1 into YW778 rescues the mutant phenotype; thus, protecting the strain from MMS induced damage. Interestingly, the newly transformed strain, YW778 with pHis-Ceapn-1, maintains the same level of growth as the parent strain (YW465) when exposed to the alkylating agent (MMS). The assay confirms that the construct is able to encode a functionally active recombinant protein in the yeast expression system. It should be noted that this phenomenon was equally revealed with the introduction of the other plasmids made during the course of this investigation (pCeapn-1-His, data not shown; pGST-His-Ceapn-1, see Fig 3-27; pGST-Ceapn-1-His, data not shown) into YW778.

Lane	Host Strain	Introduced Plasmid
1	YW778	pCeapn-1-His
2	YW465	--
3	YW778	--
4	YW778	Empty vector

0%   [MMS]   0.008%
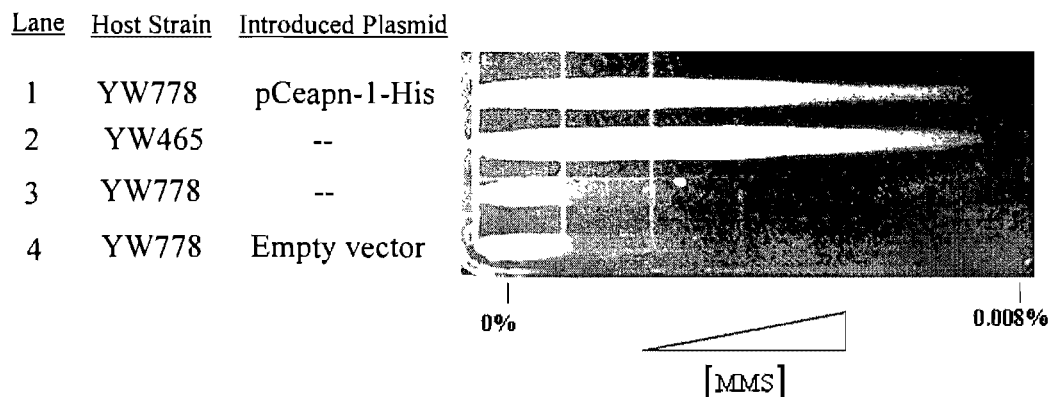
Figure 3-5: Cross Species Complementation Analysis

A gradient plate assay was conducted to observe the ability of Ceapn-1 to substitute for y*APN1* in yeast. Analysis was performed via an increasing gradient of MMS, an agent known for causing the formation of AP sites. As a result, a wild-type strain (Lane 2), equipped with AP endonucleases, is shown to grow the full length of the gradient. In contrast, a mutant strain (YW778), deficient in AP endonuclease activity, is only able to grow a very short distance, as shown in lane 3. The introduction of a plasmid allowing the expression of His-Ceapn-1 in the mutant strain (Lane 1) parallels the growth of the parent strain and consequuently, ascertains its ability to substitute for its homologue.

3.3. Sampling of Adopted Purification Schemes:

Multiple purification schemes were attempted throughout this project in order to produce a fully homogeneous fraction containing the protein of study (CeAPN-1-HIS). Figures 3-6, 3-7 & 3-8 show a sampling of the various steps, purification columns (Talon/His affinity, DNA affinity, anion exchange) and expression plasmids (described above) that were exploited in order to obtain a homogeneous protein fraction; however,

only a selected set would be discussed in this report. Hence, the experiments containing

an asterisk (*) are included in the results section.

I
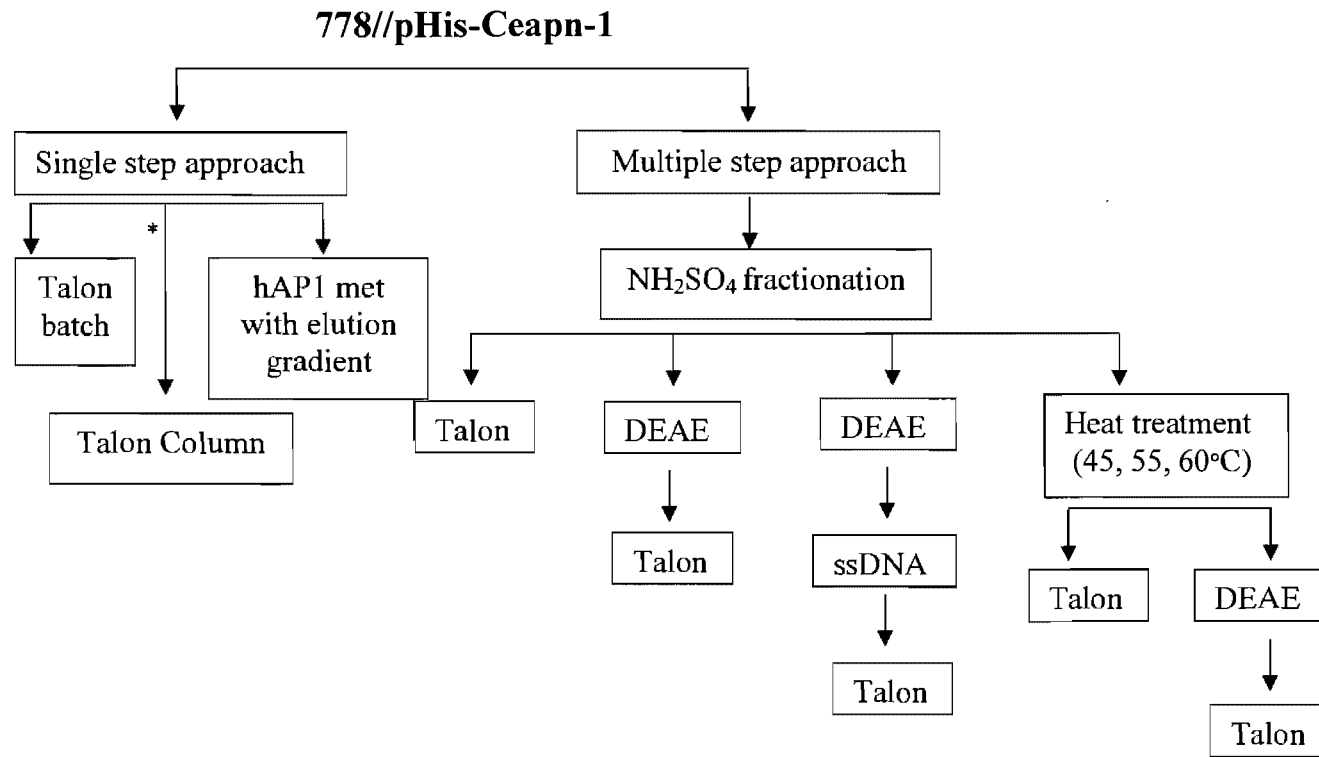
Figure 3-6: Purification of HIS-CeAPN-1

II 778//pCeapn-1-His

```
                        778//pCeapn-1-His
                 ┌──────────────┴──────────────┐
                 ▼                              ▼
        ┌─────────────────┐            ┌─────────────────┐
        │ Single step      │           │ Multiple step   │
        │ approach         │           │                 │
        └─────────────────┘            └─────────────────┘
```

Single step approach		Multiple step		
Talon	hAP1 purification method	Heat treatment (60°C)	Talon with elution gradient	Mono Q
		Talon	ssDNA / Mono Q	Talon
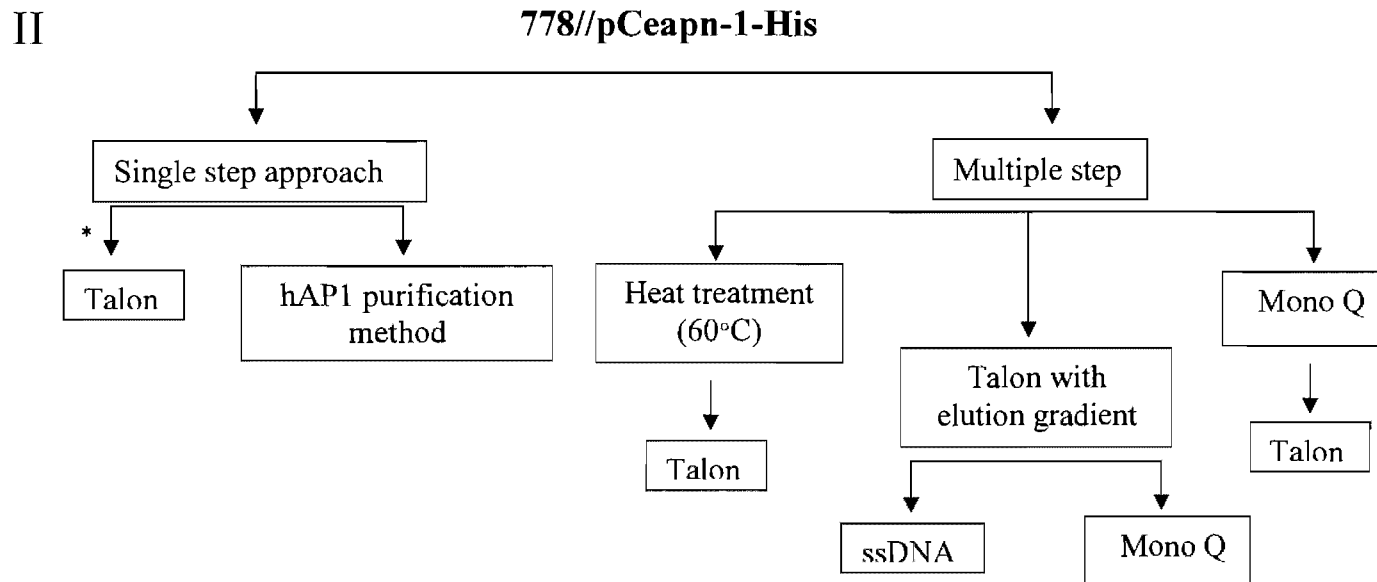
Figure 3-7: Purification of CeAPN-1-HIS

III 778//pGST-His-Ceapn-1

* Single step approach

Glutathione-sepharose

IV 778//pGST-Ceapn-1-His

Single step approach

Glutathione-sepharose
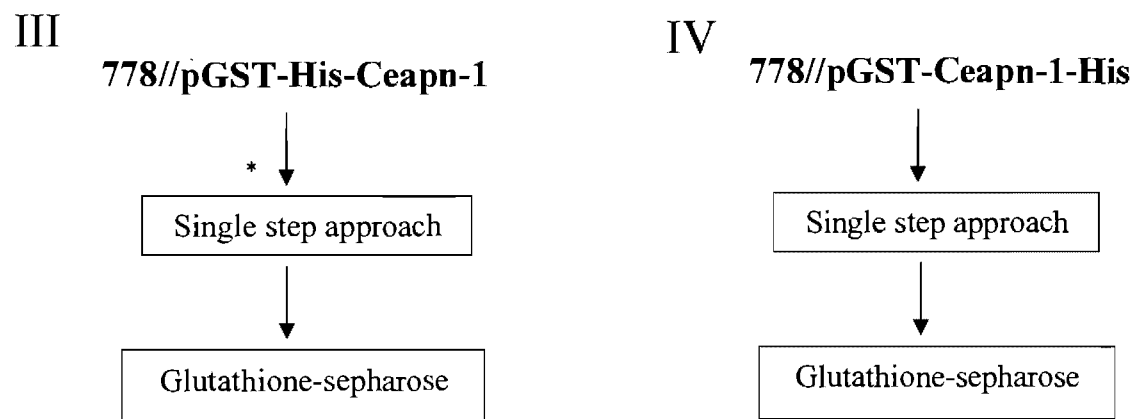
Figure 3-8: Purification of Dual Tagged Protein

48

3.4. Purification Attempts of CeAPN-1 Fusion Protein

Preliminary attempts adopted a single step approach, thus, exploiting a His affinity column to isolate the His tagged recombinant protein. The failure to capture a significant and fully pure fraction had forced the utilization of additional techniques to complement the affinity resin and achieve full purification. Some of these methods comprised the utilization of an ammonium sulfate treatment to progressively fractionate extracts, heat treatment to isolate thermo stable components, anion exchange matrices in order to separate proteins on the basis of charge and additional affinity techniques (i.e. DNA binding affinity). The design of the construct and the nature of the resulting folding of the expressed product were as well scrutinized. Moreover, the choice and number of affinity tags (single versus dual tagged) were also parameters that were pondered and experimented with to increase efficiency.

It should be noted that not all of the purification schemes that were carried out are illustrated in this thesis. The results section, in fact, only highlights the single purification step approach that was performed in the isolation of HIS-CeAPN-1, CeAPN-1-HIS and the dual tagged recombinant protein, GST-HIS-CeAPN-1.

3.4.1. Gravity-Flow Column purification of the HIS-CeAPN-1 Recombinant Protein

Figures 3-9A and 3-9B reveal the data obtained following a column purification attempt of HIS-CeAPN-1. It should be noted that the protein extracts were passed through the His affinity column via gravity. It is believed that this technique would increase binding efficiency by allowing more time for the appropriate proteins to interact with the affinity resin without forced agitation which may disrupt the desired association.

Figure 3-9A represents an AP assay conducted to monitor protein activity throughout the different steps of the purification scheme. Thus, a 42 bp ^{32}P labeled oligonucleotide containing an AP site at the 21st bp of the strand was utilized as substrate for AP endonuclease. In the presence of enzymatic activity, the 42mer (upper band) is cleaved 5' to the AP site to produce a 20mer (lower band). As a result, lane 4 shows that the addition of crude extracts generates substrate turnover; thus, confirming the presence of active AP endonuclease (HIS-CeAPN-1) in the sample. Similarly, of the 7 subsequent eluted fractions, it is observed that only fraction 2 (F2) in lane 4 shows significant substrate turnover; suggesting the definite presence of HIS-CeAPN-1. Figure 3-9B demonstrates the western blot, probing with 0.1 % anti His antibody, that was performed to analyze the active fraction, F2 (500 ng), collected from column purification. Thus, lane 1 (corresponding to elution F2) reveals the presence of multiple bands which may be likely indicative of contaminating proteins in the eluted fraction. The inability to detect the recombinant protein via the anti-His antibody in the crude extracts could be due to the fact that an excessive concentration of protein (0.5 mg) was loaded. In fact, this phenomenon is also observed in Figure 3-15C whereby at very elevated levels, protein detection becomes less efficient in crude extracts due to over-saturation. Figures 3-9B and 3-9C reveals that the band signal representing HIS-CeAPN-1 (lane 1) is not the predominant species of intensity in the purified sample.
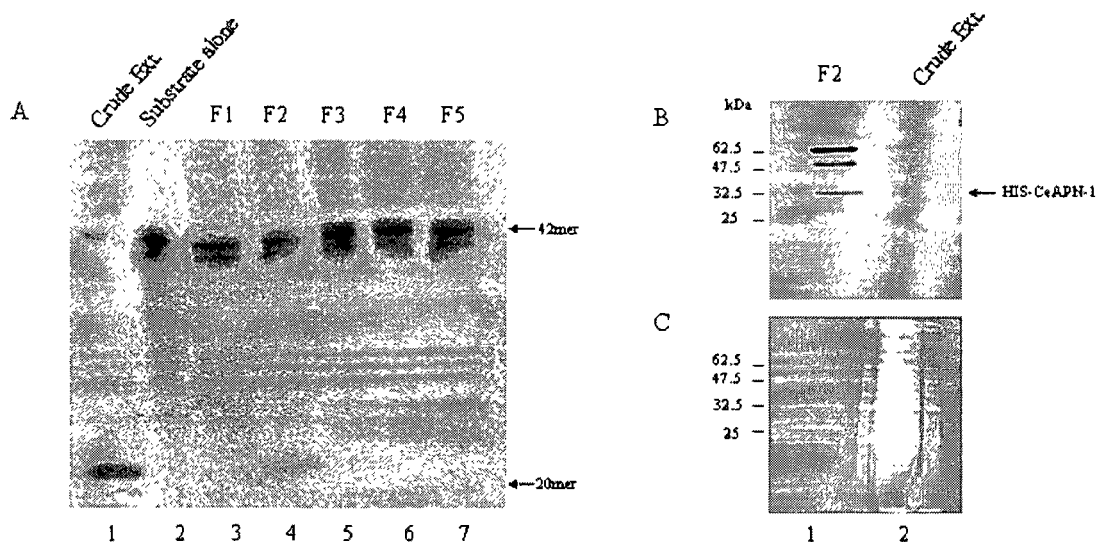
Figure 3-9: Evaluation of single step His Affinity (Talon) column purification

A) 42 bp oligonucleotide substrate $U^{21}\cdot G$ was incubated at 37°C for 30 minutes with collected elution fractions (lanes 3-7) from the His affinity Talon column. Substrate turnover was evaluated with a phosphorimager following migration on a 10% polyacrylamide/urea gel. 1.5 ng of substrate was utilized for each reaction. Lane 1, substrate alone; lane 2, substrate with 5 µg crude extract (pre-column loaded); lanes 3-7, collected elution fractions (fractions 1-5). 2 µl of each fraction (200 µl) was added into each reaction. B & C) The purity of the elution fraction F2 (demonstrating activity) was ascertained via western blot analysis using 0.1 % anti-His (B) and sypro staining (C). Lane 1 (500 ng), F2; lane 2, crude extract (50 µg) for both western blot and sypro staining. Lane 1 demonstrates that the fraction lacks purity as indicated by the presence of the contaminating band signals. Figure 3-9B shows that HIS-CeAPN-1 is found in the investigated fraction, as indicated by the arrow. Although active, F2 only shows a very small yield. Also, the recombinant HIS-CeAPN-1 protein is unable to be detected from the crude extracts due to the elevated concentration of proteins in the lane,

as similarly seen in Fig 3-15C-lane 4. Figure 3-9C (sypro staining) fails to clearly distinguish the presence of HIS-CeAPN-1.

3.4.2. Evaluation of Detected Proteins Following Purification

Repeated efforts in isolating a homogeneous protein sample have been mildly successful; hence, as opposed to a fully purified fraction, a partially purified but active fraction containing CeAPN-1 was obtained. The bands observed via sypro staining may indeed be the result of non-specific bands. However, it is plausible that CeAPN-1 may be interacting very tightly with other components in the extracts and subsequently making the task of obtaining a fully pure protein more challenging. The additional bands observed on the corresponding stained SDS PAGE gel following His affinity purification could stem also from a form of protein modification affecting CeAPN-1.

To assess whether the bands commonly observed following elution from the Talon column are in fact, representative of interacting proteins, modifications of CeAPN-1 or simply non specific, a parallel purification (using Talon resin) was performed. Hence, this was performed using extracts from the original mutant background (YW778) and YW778 expressing HIS-CeAPN-1.

Figure 3-10 reveals a western blot whereby the corresponding eluted fractions stemming from the two sets of extracts (YW778 and YW778/pHis-Ceapn-1) are probed with anti-His. Hence, lanes 1-2 represent an increasing concentration of F2 from the YW778 column while lanes 3-6 show an increasing concentration of F2 stemming from the YW778/pHis-Ceapn-1 column. It is clear that most of the bands can be accounted as non-specific. The protein of interest is represented at 34 kDa in lanes 5-6. The uppermost band found across lanes 5-6 can be suggested to be the result of a possible interacting component with CeAPN-1. The remaining bands of strong intensity,

especially that found at 62 kDa, are thought to be non-specific, the result of contaminating proteins.
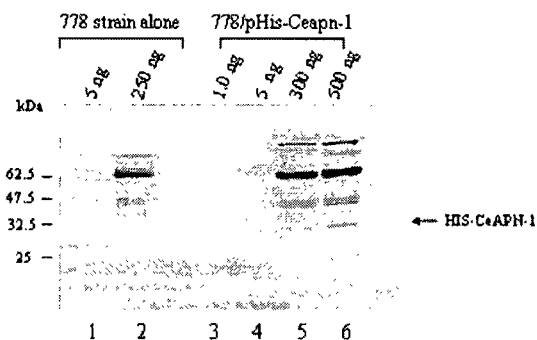
Figure 3-10: Purification Comparison Using Extracts from YW778 and YW778 Expressing HIS-CeAPN-1 via Western Blot

Extracts stemming from both strains were loaded onto individual His affinity (Talon) columns and purified in parallel. The 2nd elution fraction from each respective purification was isolated and evaluated via western blot. Lanes 1-2, increasing concentration of F2/ empty strain (YW778), lanes 3-6, increasing concentration of F2/ YW778 strain expressing HIS-CeAPN-1.

3.4.3. New NLS-Ceapn-1-His Construct

The inability to purify the HIS-CeAPN-1 recombinant protein had forced us to reconsider strategies beyond general purification schemes. In essence, simple and multi-step purification approaches failed to yield a pure fraction. The problem may stem from the Talon affinity beads themselves and a possible failure of the proteins to securely bind to the matrix. It may be suggested that the manner in which the recombinant protein was folded may have interfered with binding. Hence, it should be remembered that the His tag is found at the N-terminal end between the nuclear localization signal

and the gene. Thus, the NLS motif or other structural features of the protein may be impeding access to the His tag and subsequently preventing any strong association with the affinity beads. Previous constructs utilized to express and purify recombinant forms of yApn1 consisted of a tag on the opposite end of the bipartite nuclear localization signal. Interestingly, the purification of the yApn1 recombinant protein was not a large hurdle as it was observed with the corresponding CeAPN-1 polypeptide. Consequently, a second construct was designed in a fashion that the His tag was moved to the C-terminal of the gene while maintaining the NLS at the N-terminal.

3.4.4. Single Step Purification Approach

Protein extracts expressing CeAPN-1-HIS were loaded onto a His affinity Talon column in a single step manner without prior manipulation (i.e. no fractionation and anion exchange column). Washing and elution steps of the column were performed by standard conditions as suggested by the resin manufacturer. The resulting western blot (figure 3-11A) demonstrates that the antibody is fairly specific and the protein of study is well detected. A coomassie stained SDS-PAGE gel (figure 3-11B) showing the true purity of the elution F-2, however, does indicate the presence of multiple bands. Evidently, the coomassie reveals that the fraction in question is far from being homogeneous. Despite the isolation of an impure fraction, an AP assay (3-11C) indicates that the partially purified elution is, indeed, capable in processing AP sites. In fact, the AP site carrying 42mer (1.5 ng) is shown to be well converted to the desired 20mer in the presence of F2.
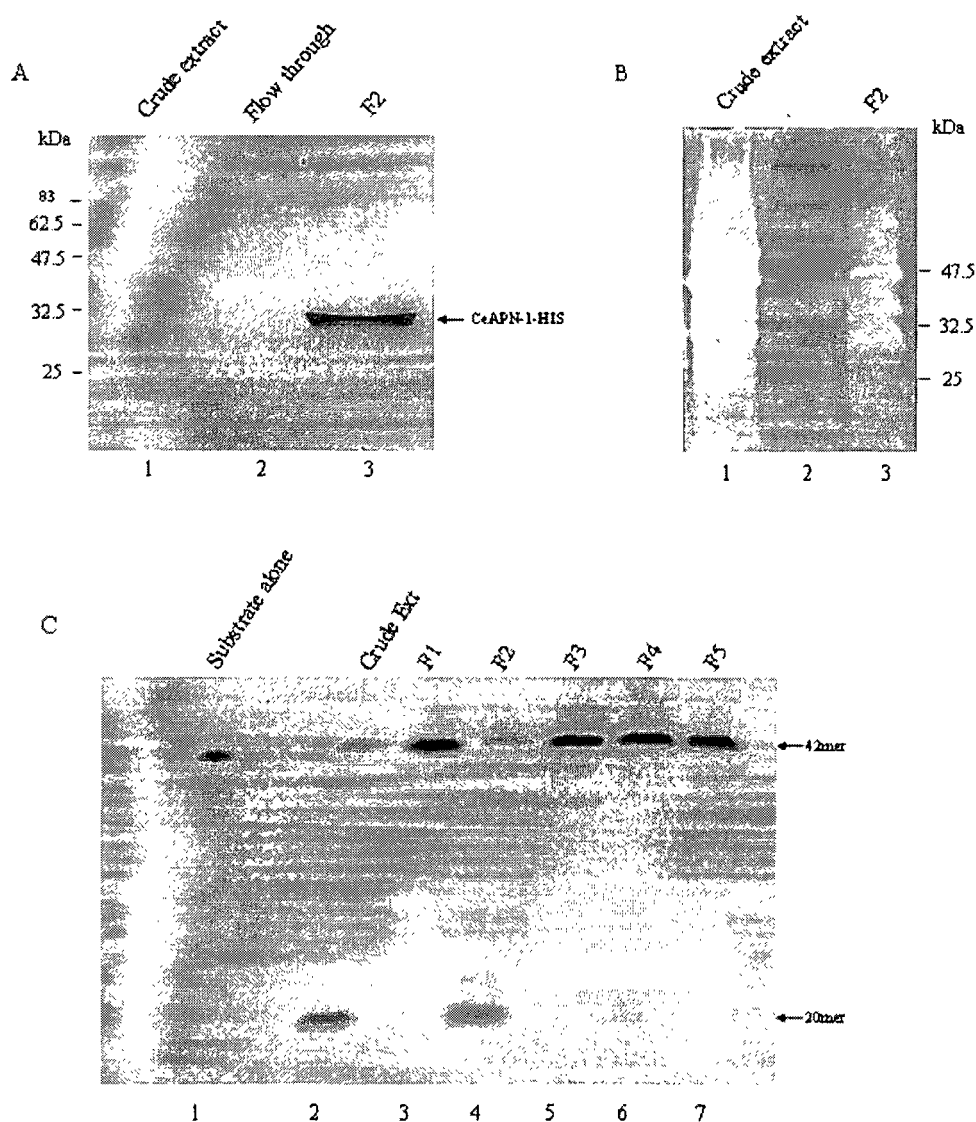
Fig 3-11: Talon Column Purification of CeAPN-1-HIS Protein

A one step purification scheme (utilizing Talon resin) was performed, in order, to obtain a homogeneous fraction composed of CeAPN-1-HIS. Thus, a western blot (A) and coomasie stain (B) were completed to monitor its effectiveness. A) lane 1, 10 µg protein extracts used for subsequent purification; lane 2, flow through obtained following binding with resin; lane 3, 500 ng fraction 2 (F2) collected during final elution from beads. B) lane 1, 10 µg crude extracts used for subsequent purification; lane 3, 500 ng fraction 2 (F2). An AP endonuclease activity

assay (C) was performed (as shown in figure 3-9A) to confirm the presence of CeAPN-1 in the collected fraction and assess its enzymatic capacity. The upper band represents the 42bp oligonucleotide substrate (carrying an AP site), while the presence of a bottom band is representative of processed substrate (20mer product). (C) Lane 1, 1.5 ng substrate alone; lane 2, 5 µg crude extract before loading onto beads with substrate (1.5 ng); lane 3, 2 µl of collected elution fractions (200 µl each, F1-F5) with 1.5 ng substrate.

3.5. Characterization of Partially Purified Extracts Containing CeAPN-1-HIS

Although a homogeneous fraction consisting solely of CeAPN-1-HIS had yet to be obtained, partially purified extracts were utilized for preliminary kinetics analysis and functional studies. Thereupon, an active fraction (F2) obtained following its isolation from the Talon column was utilized to perform this initial investigation (as seen in figures 3-12A-B).

3.5.1. Protein Concentration Dependence of AP Site Processing

Figures 3-12A & 3-12B represent the AP assay and its corresponding graphical depiction. The figures demonstrate that substrate conversion occurs fairly quickly while increasing enzyme concentration which is, in turn, followed by a plateau in product formation. It appears that a significant amount of the substrate (almost 80%) remains unprocessed despite increasing the concentration of AP endonuclease. Hence, it is possible that the unconverted substrate may be a specific form or species that is unable to be processed by the enzyme in question. In fact, one can suggest that a significant concentration of the substrate was not optimally processed by the UDG enzyme, which is normally utilized to produce the AP site and prime the substrate for AP endonuclease.

Also, another likely possibility to explain the failure to achieve complete substrate turnover may be linked to the fact that a particular component, yet unknown, may be limiting.
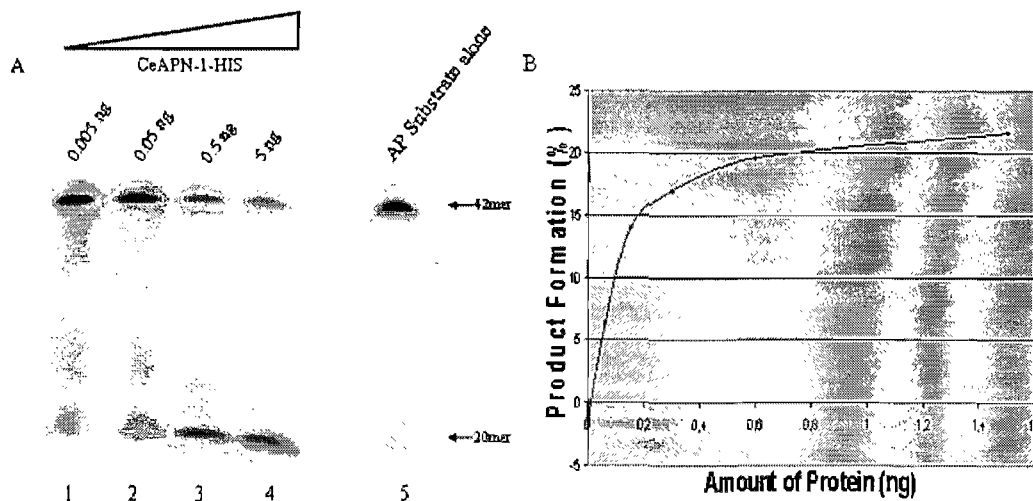
Figure 3-12: Protein Concentration Dependence of AP Site Processing

AP endonuclease assay (A) was conducted to evaluate the enzyme concentration dependence of AP site processing. A) Lane 1, 1.5 ng substrate alone; lanes 2-5, substrate (1.5 ng) with an increasing concentration of CeAPN-1-HIS (partially purified). B) Product formation obtained as a function of protein amount is depicted using the Imagequant program.

3.5.2. Time Dependence of AP Site Processing

Figure 3-13A shows the AP assay performed to analyze the effect of time on AP site processing. A graphical representation of the corresponding assay is depicted in figure 3-13B. It is evident that with increasing time (while maintaining the concentration of 42mer and enzyme at a fixed level), the substrate turnover increases

steadily in reaction and plateaus following 10 minutes. A similar phenomenon is observed, as in the concentration dependence experiments (figure 3-12), whereby full conversion of the substrate is not achieved. Hence, this may stem from the quality of the substrate accessible to the enzyme in question or the possibility that another factor may be limiting.
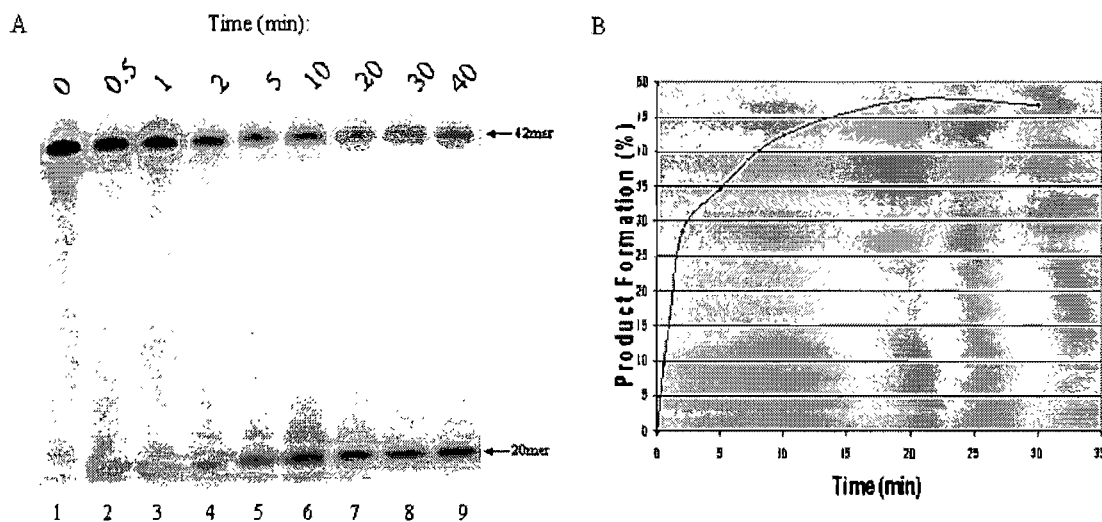
Figure 3-13: Time Dependence of AP Site Processing

AP endonuclease assay (A) was conducted to evaluate the dependence of time on AP site processing. A) Lanes 1-9, 1.5 ng of substrate with fixed amount (5 ng) of CeAPN-1-HIS (partially purified) while increasing time of reaction (0 min to 40min). B) Product formation obtained as a function of time is depicted using the Imagequant software.

3.5.3. Concentration Dependence in the Processing of 3'α,β Unsaturated Aldehyde Substrate (3'-Blocked Ends):

AP endonucleases have classically been shown to be bifunctional; thus, in addition to its AP endonuclease component, they have a well established 3'-diesterase

activity, which enables the processing of 3' blocked ends. Figure 3-14 reveals the 3'

diesterase assay that was performed to analyze the bifunctionality, if any, of CeAPN-1-

HIS. The 3'α,β unsaturated aldehyde substrate (3' blocked end) is obtained by

incubating the labeled 42mer, already treated with the enzyme UDG, with Endo III (AP

lyase) which forms the desired 3' blocked end.

The assay confirms through *in-vitro* means that CeAPN-1, akin to its

counterparts of the Endo IV family, is composed of a 3' diesterase component. The

3'diesterase activity is revealed by the subtle shift caused by the turnover of 3'α,β

unsaturated aldehyde group to the 3'OH group.
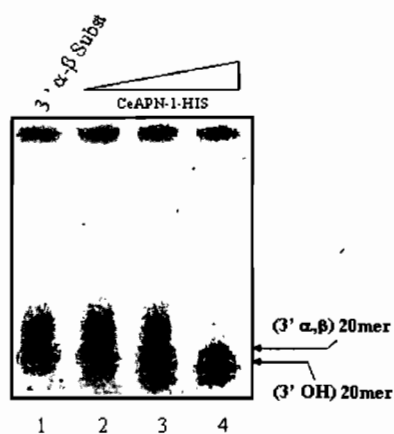
Figure 3-14: Concentration Dependence of 3'α,β Unsaturated Aldehyde

3'-diesterase assay was conducted to evaluate the dependence of enzyme

concentration upon the processing 3'α,β unsaturated aldehyde. Lanes 1,

1.5 ng substrate bearing the 3'α,β unsaturated aldehyde group following

Endo III treatment, lane 2-4, substrate with an increasing concentration

of the partially purified fraction containing CeAPN-1-HIS (0.05, 0.5, 5.0

ng).

3.6. Complete Purification of the CeAPN-1 Recombinant Protein

3.6.1. GST-His-Ceapn-1 Construct

Successive experiments have showed that the complete purification of the His tagged protein requires a complex approach. Two constant hurdles have been encountered: the inability of the protein to be strongly expressed and secondly, the inability of the recombinant protein to tightly bind to the affinity matrix. In the hopes to simplify the purification scheme and optimize full protein isolation, a second tag was fused to the Ceapn-1-His construct. A similar GST tag system was fused to y*APN1* to perform subsequent characterization [79, 88]. Nonetheless, the utilization of a GST tag was initially avoided for this particular project because of previous experiments demonstrating that a GST tagged CeAPN-1 recombinant protein was not active. It was believed that GST, being a much larger tag, may interfere with the integrity of the enzyme active site. However, it was only later recognized that the presence of the NLS (not included previously) may have the ability to preserve enzyme activity even with a bulkier tag. It should be noted that an additional modification was also performed in order to alleviate the weak expression of CeAPN-1; thus, the ADH constitutive promoter was substituted with a P_{cup} copper inducible promoter. This latter promoter had been successfully used in the expression of CeEXO-3 in yeast [85]. Although the ADH promoter is known to be fairly strong, it failed to strongly express CeAPN-1 in yeast.

3.6.2. Purification of GST Tagged Protein

The GST-tagged recombinant protein was purified via a glutathione sepharose resin column. In essence, crude extracts were applied to the matrix in question, washed

with the appropriate conditions and thereafter, the elution fractions were collected following the addition of reduced glutathione.

The resulting 5 elution fractions are represented on a sypro stained gel (figure 3-15A). It is evident that contrary to previous attempts at purification, only a single band is observed in lanes 1 and 2 (figure 3-15A) and, moreover, one that coincides with the expected size of the recombinant protein (61 kDa). A silver stained gel was also performed (fig 3-15B) where an increasing amount (2, 10, 35 ng) of the isolated fraction is represented. The silver stain, being more sensitive than the sypro-stain, may reveal, if any, the presence of non-specific components that were not revealed in the latter. However, in this particular trial, the isolated fraction appears to be very pure. The presence of the protein in the corresponding fraction is confirmed via western blot (figure 3-15C) where an increasing concentration of crude extracts (5, 80, 150, 200, 250 μg) and pure protein (1, 5, 20, 35 ng) are individually represented and probed with anti-His (1:1000). The antibody does not appear to detect any fragmented nor non-specific proteins in the eluted fraction.
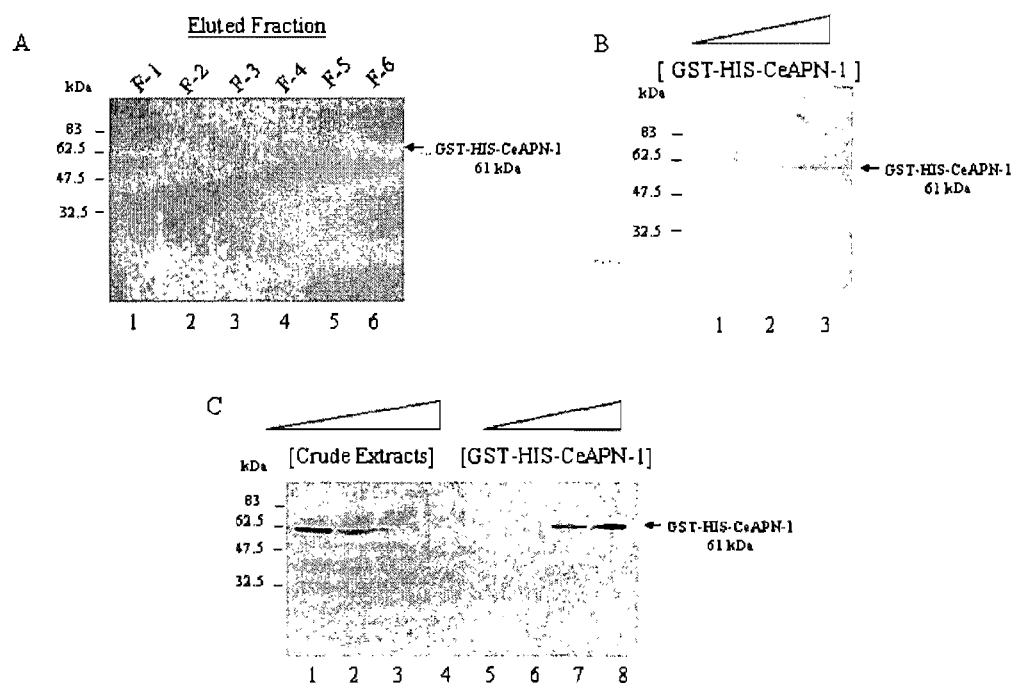
Figure 3-15: Analysis of Eluted Fractions from Glutathione Sepharose Resin

Extracts were loaded onto a glutathione sepharose column. Subsequent elution fractions were collected following the application of 10 mM glutathione. Panel A demonstrates the sypro stained gel conducted to evaluate equal volume (35 µl) of the individual fractions, F1-F6 (lanes 1-6) isolated from the resin. Panel B reveals a silver stained gel loaded with an increasing amount (10, 25, 50 ng) of the most concentrated fraction (F1) obtained from the elution profile (lanes 1-3). Panel C shows a western blot analysis comparing the increasing concentration of crude extracts (50, 100, 150, 200 µg) (lanes1-5) and the most concentrated eluted fraction, F1 (10, 20, 80, 120 ng) (lanes 6-8).

3.7. AP Endonuclease Activity Assay of GST Recombinant Protein

Although purified, the corresponding AP assay failed to show that the purified protein sustained significant activity. The purified *C. elegans* recombinant protein appears to differ in this particular aspect to its characterized homologues, Apn1 and

endo IV, in *S. cerevisiae* and *E. coli* [55]. Figure 3-16 shows the collected data from the exposed polyacrylamide-urea gel where the activity reactions were loaded. Thus, it is demonstrated that the purified protein is indeed unable to process the substrate. Although, a faint band coinciding with the desired 20mer (demonstrating cleaved 42mer) is slightly visible, its intensity remains weak in comparison to what is observed in the presence of crude extract (prior to purification) and purified endo IV. The inactivity of the purified protein is not believed to stem from a construct defect. Thus, the crude extracts, used for subsequent purification, demonstrate to have strong activity. One can suggest that the activity was lost during the actual purification process.
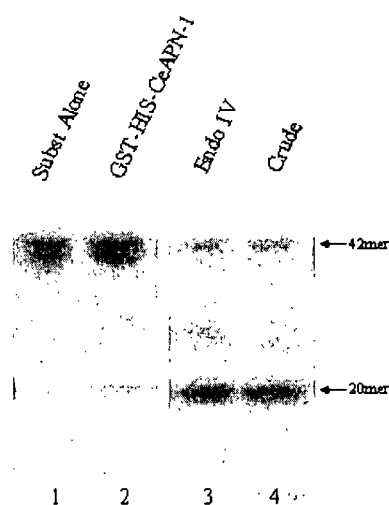
Figure 3-16: AP Activity Analysis of Purified Protein

Activity of purified fraction (F1) was evaluated against 1.5 ng of the $U^{21} \cdot G$ substrate (42 bp oligonucleotide containing AP site). Lane 1, substrate alone; lane 2, substrate incubated with 1 ng of purified fraction (F1); lane 3, substrate incubated with 1 ng of bacterial endo IV; Lane 4, substrate with 2 µg crude extracts used for subsequent column purification.

3.8. Addition of YW778 Protein Extracts Stimulate CeAPN-1 Activity

Due to the fact that the crude extracts have shown to retain activity in contrast to the fully purified fraction, it was hypothesized that another factor may be involved in the stimulation of GST-HIS-CeAPN-1. Thereupon, extracts from the YW778 strain were prepared and added in reaction with GST-HIS-CeAPN-1 and substrate. Hence, it is observed (figure 3-17, lanes 1-4) that the addition of extracts does, indeed, appear to dramatically rescue protein activity. This particular phenomenon appears to be dependent on the concentration of extract added in reaction; thus, lanes 1-4 show a steady increase in substrate turnover and a subsequent plateau.
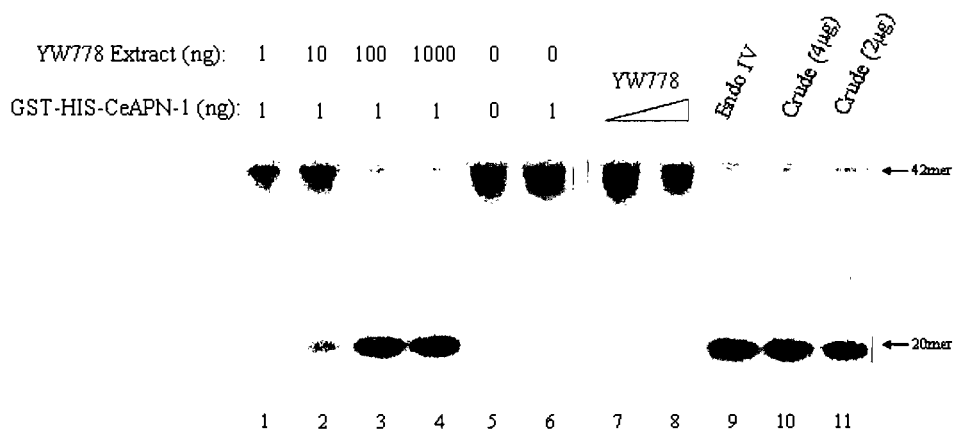
Figure 3-17: Addition of YW778 (*apn1Δ*, *apn2Δ*) Protein Extracts Stimulate CeAPN-1 Activity

Detection of AP activity of purified GST-HIS-CeAPN-1 fraction (1 ng) was monitored in the presence of an increasing concentration (1-1000 ng) of YYW778 crude extract (lanes 1-4). Lane 5, 1.5 ng substrate alone, lane 6, 1.5 ng of substrate with purified protein, lane 7-8, substrate (1.5 ng) with increasing concentration (10 ng, 1000 ng) of the YW778 AP endonuclease deficient extract. 1 ng of purified endo IV (with substrate) was included in lane 9 as a positive control.

Methodology

It has been tentatively shown that the addition of extracts appears to stimulate

the activity of the CeAPN-1 recombinant protein. In order to better understand the

underlying component causing this phenomenon, the extracts were pre-treated as shown

in the figure below:

YW 778 (*apn1Δ apn2Δ*) Cell
culture

↓

1.   Protein
extracts

↓

2.   Heated // Boil

↓

Spin at 10 000 rpm

↓

3.   Collecting supernatant
(soluble fraction)

↓

4.   DIALYSIS
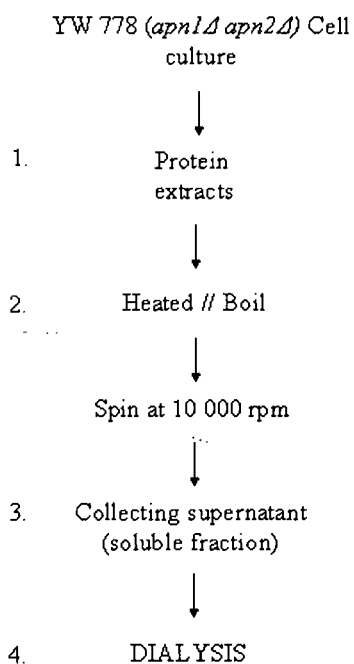
Figure 3-18: Treatment of Protein Extracts

Protein extracts from the AP endonuclease deficient yeast strain
(YW778) were treated in a series of steps. Extracts collected following
each of the numbered steps were added individually into reaction with
purified GST-HIS-CeAPN-1 as well as substrate and subsequently
evaluated for enzyme activity (AP endonuclease).

The heat denaturation step was included to inactivate the proteins found in the

extracts. The absence of activity observed following the addition of heated extracts

would suggest an involvement of proteins in the stimulation of the Endo IV member. The dialysis step, on the other hand, was conducted in the event that the previous heat treatment would not affect activity stimulation. Hence, this particular step would be effective in removing all metal factors in the extracts (components under 5000 Da). Thereupon, the absence of activity following the addition of dialyzed extracts would be indicative that a metal ion may be involved in the function of GST-HIS-CeAPN-1.

3.8.1. Addition of Heated Extracts

In order to perform a preliminary analysis to 'narrow down' the identity of the component, if any, causing the stimulation of activity, protein extracts were heated to 95° C for 5 minutes and added to reaction with purified protein and substrate. The heat treatment is exploited to denature proteins in extracts and render them inactive; consequently, it would allow one to deduce whether the component responsible for the stimulation is a metal factor or a protein.

Figure 3-19 shows the corresponding data obtained following the AP assay. It is observed that despite strongly heating the extracts, substrate turnover (illustrated by the formation of 20mer product) is preserved. Substrate turnover is not observed when heated extracts are assayed in the absence of CeAPN-1 (Lanes 2-3). This particular reaction serves as a negative control demonstrating the extracts alone do not contain AP endonuclease activity. One should further note, as shown earlier, that the CeAPN-1 alone produces very little, if not at all, of the processed 20mer.

Figure 3- 19: Addition of Heated Protein Extracts

AP endonuclease deficient (YW778) extracts were pre-heated for 5 minutes at 95° C and thereafter, added in an increasing concentration in reaction with purified protein and substrate (maintained concentration). Lane 1, substrate alone (1.5 ng); lanes 2-5, 1.5 ng substrate with increasing concentration of both heated and non heated AP endonuclease deficient extracts respectively (10 ng, 1000 ng); lane 6, 1.5 ng substrate with 1 ng of purified endo IV; lane 7, substrate with 4 μg crude extracts utilized in the purification of GST-HIS-CeAPN-1; lanes 8-11, constant level of substrate (1.5 ng) and purified GST-HIS-CeAPN-1 incubated with an increasing concentration of heated YW778 extracts.

3.8.2. Addition of Specific Metal Factors

Due to the fact that heated extracts were able to stimulate CeAPN-1 activity, it was assumed that a protein would likely not be involved. In fact, a metal factor was thought to be the more likely candidate since it remains unaffected by the heat treatment. Thereupon, several metal factors were individually added to reaction with purified recombinant protein in order to assess for activity stimulation. The metal factors evaluated included Zn^{2+}, Co^{2+} and Mn^{2+} as shown in figure 3-20. Metal ions (transition metals) often play an important role in stabilizing protein structure [98].

However, none of the factors suggest to confer any activity stimulation. In essence, the 42mer substrate remains unchanged despite the addition of metal factors. Additional factors, including Mg^{2+}, Ni^{2+} and Ca^{2+} have been tested in the past with partially purified extracts but similarly did not reveal any activity stimulation (data not shown).
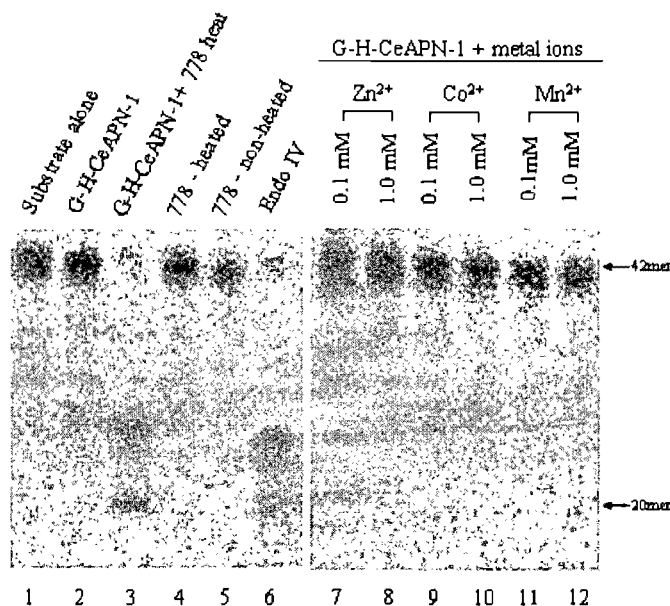
Figure 3-20: Evaluation of Metal Ion Dependence of G-H-CeAPN-1

Activity of G-H-CeAPN-1 (GST-HIS-CeAPN-1) was evaluated in the presence of various divalent metal ions, which included Zn^{2+}, Co^{2+}, and Mn^{2+}. Thus, individual metal ions were added as 2 specific concentrations (0.1 mM and 1.0 mM) while maintaining the level of purified protein and substrate. Lane 1, 1.5 ng substrate alone (constant in all other reactions involving substrate); lane 2, 1 ng of purified protein with substrate, lane 3, substrate incubated with 1 ng purified protein and 100 ng of heated AP endonuclease deficient (YW778) extracts; lanes 4-5, substrate with 100 ng heated and non-heated YW778 extracts respectively; lane 6, substrate with 1 ng of purified endo IV; lanes 6-11, substrate incubated with 1 ng of purified protein and an increasing concentration (0.1 mM and 1 mM) of the respective individual metal ions (Zn^{2+}, Co^{2+}, and Mn^{2+}).

3.8.3. Addition of Dialyzed Extracts

In order to further assess the nature of the observed phenomenon, extracts were thoroughly dialyzed (tubing pore size consisted of a molecular weight cut off of 5000 Da) following heat treatment and subsequently, applied in reaction with pure protein and substrate. The dialysis was performed to allow metals and other small components (under 5000 Da) to be removed from the extracts. As a result, despite heating and dialysis steps, fig 3-21 reveals that the extracts still stimulate the activity of GST-HIS-CeAPN-1.
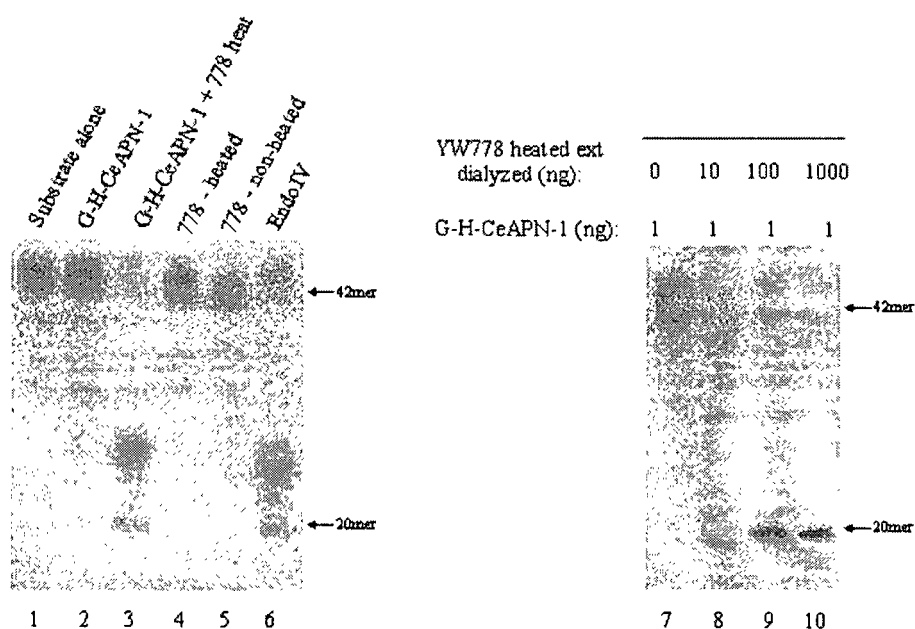
Figure 3-21: Addition of Dialyzed Extracts

YW778 extracts were dialyzed (following heat treatment) in order to fully remove any metal components that may be found in the sample. These treated extracts were thereafter, incubated in an increasing concentration with the purified recombinant protein G-H-CeAPN-1 (GST-HIS-CeAPN-1) and 1.5 ng of U^{21}·G substrate (42 bp oligonucleotide). Lane 1, substrate alone in reaction; lane 2, substrate incubated with 1 ng GST-HIS-CeAPN-1; lane 3, substrate with purified protein (1ng) and heated 100 ng of YW778 extracts, lane 4-5, substrate with 100 ng of heated and non-heated YW778 extracts respectively; lane 6, substrate with 1 ng of endo IV protein; lanes 7-10, substrate with purified recombinant protein and an increasing concentration of dialyzed YW778 extracts.

3.8.4. Summary of Observed Results:

YW 778 *(apn1Δ apn2Δ)* Cell culture   Effect on CeAPN-1 Activity

1. Protein extracts   ⟹   +

2. Heated // Boil   ⟹   +

Spin at 10 000 rpm

3. Collecting supernatant (soluble fraction)   ⟹   +

4. DIALYSIS   ⟹   +
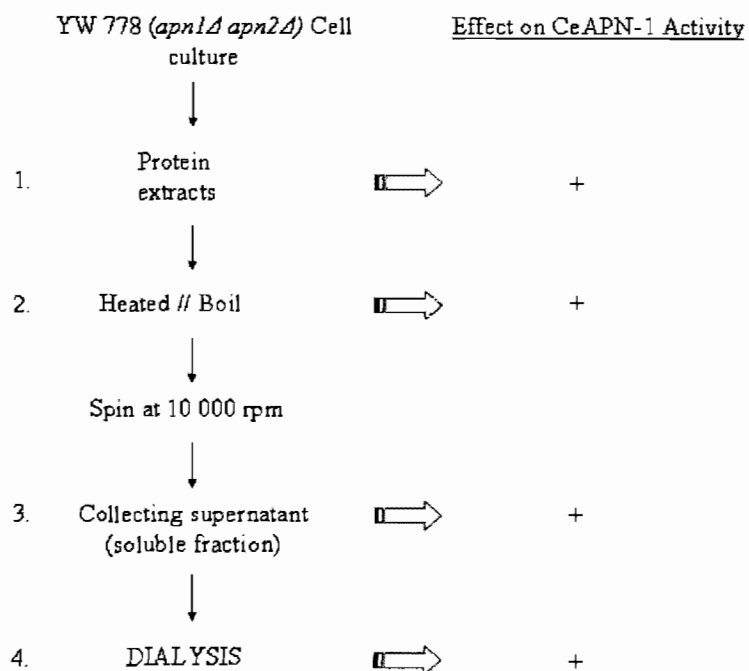
Figure 3-22: Summary of Observed Results Following the Addition of YW778 Protein Extracts

AP endonuclease activity associated with GST-HIS-CeAPN-1 was manifested following the addition of AP endonuclease deficient extracts. This activity was maintained despite the various treatments issued to the extracts (i.e. heating and dialysis)

3.9. Domain Function Analysis

It is known that 5 conserved domains exist within the Endo IV family. Four of the five domains have been correlated to govern protein activity; thus, key residues in these segments have been implicated in DNA binding and the coordination of zinc atoms within the active site.

Active site consisting of Zn²⁺ cluster
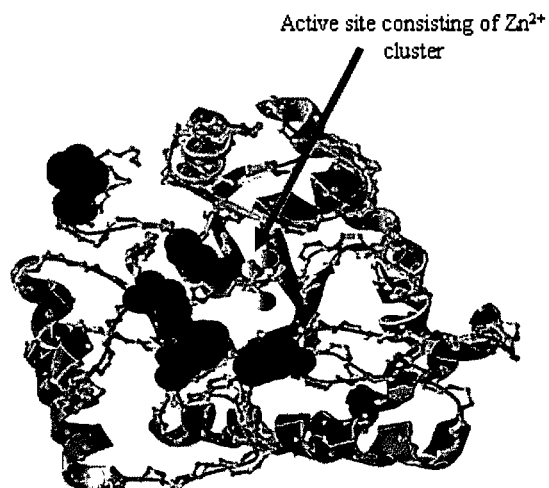
Figure 3-23: Model of CeAPN-1 Based on Known endo IV Structure

The Swiss Model Server program (v3.5) was utilized to develop a model of the CeAPN-1 protein structure. The model is inspired by the known and well characterized structure of the bacterial homologue, endo IV.

Using the obtained CeAPN-1 model structure (refer to Method and Materials 2.16), one can visualize and extrapolate the predicament awaiting the CeAPN-1 protein by changing key amino acid residues. Thus, 2 sites have been initially chosen to perform site directed mutagenesis (shown in fig 3-24):

(1) E97G - found in the first domain and makes up a segment of acidic residues that may be important in contributing to protein stability.

(2) E143G - found in the second domain and believed to be key in coordinating Zn^{2+}.
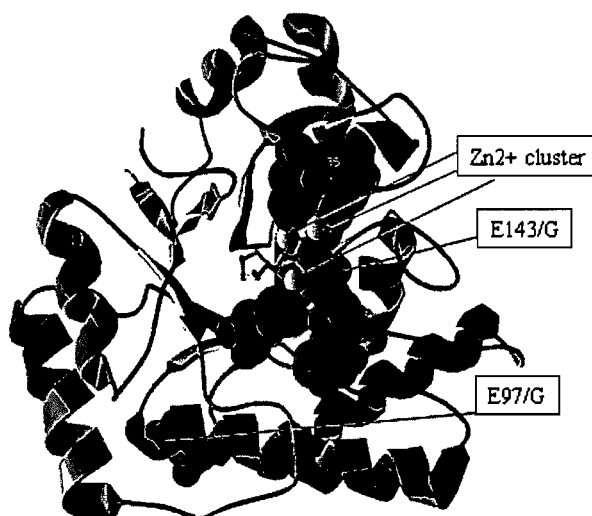
Figure 3-24: Chosen Sites for Mutagenesis

Through the utilization of an alignment of Endo IV family members, two sites have been chosen for site directed mutagenesis to enable subsequent domain function analysis. Hence, these include, E143G and E97G.

It is observed that E143 is embedded deep inside the protein and likely involved in playing a significant role in catalytic activity (Fig 3-24 and 3-25). E97, in contrast, is found on the outer shell of the protein structure and its substitution with glycine is believed to affect protein stability. In fact, Fig 3-25 shows that E97 is found within a rich cluster of acidic residues which is represented by the red colour tone. In consequence, the glycine substitution appears to disperse the rich acidic cluster and subsequently, believed to negatively affect the function of the enzyme (fig 3-25).

A

CeAPN-1 (wt)   CeAPN-1-97   CeAPN-1-143

90 °C

B

CeAPN-1 (wt)   CeAPN-197   CeAPN-1-143
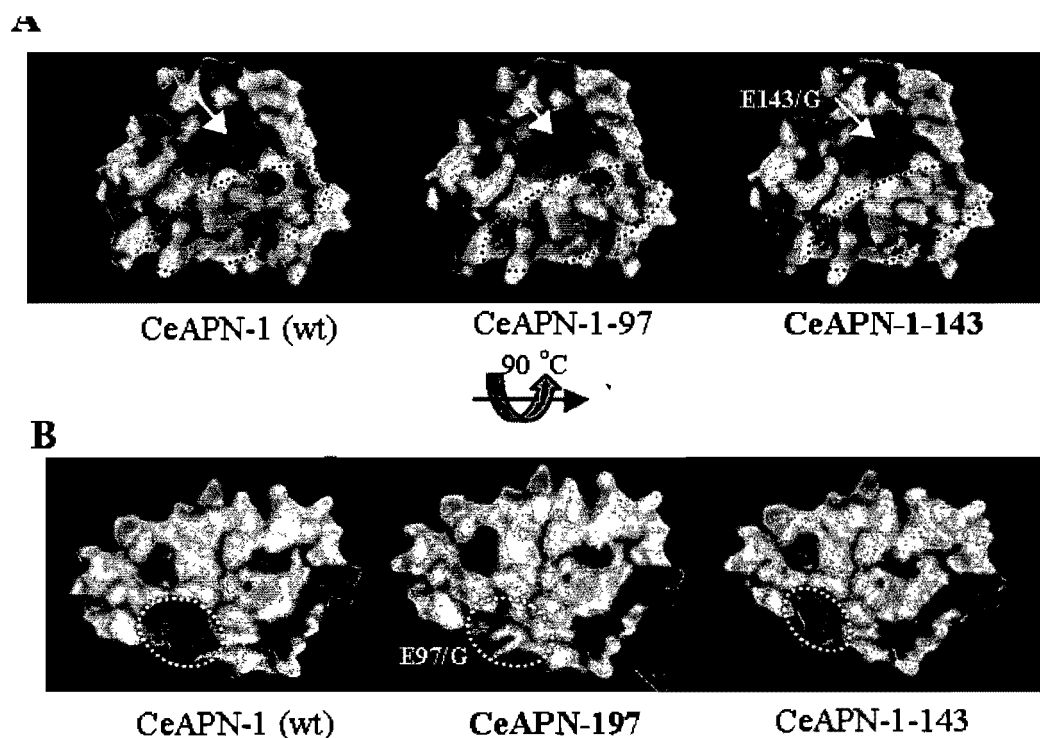
Figure 3-25: Visualization of Variant Proteins

Utilization of the Swiss Model server enables visualization of the variant proteins. E97 is shown to be on the outer shell of the protein. Interestingly, the E97G mutation is shown to cause a dispersal of the acidic cluster of amino acids (represented in red). E143, in contrast, is found in an interior groove of the protein associated with the active site. Consequently, mutation of this particular site (E143G) was believed to have negative consequences on the activity of the protein.

3.9.1. Expression of Variant Proteins Following Copper Induction in the YW778 Background

Yeast strains expressing the variant proteins were induced with copper to allow for optimal expression. Thereafter, the cultures were pelleted, thoroughly washed and finally evaluated via western blot (Fig 3-26). Figure 3-26 demonstrates that the E97G fails to be stably expressed (lane 8). However, there is an accumulating band signal at

50 kDa which may be a sign of protein degradation. In contrast, E143G does, in fact, demonstrate to be stably expressed without an excessive accumulation of degradation.
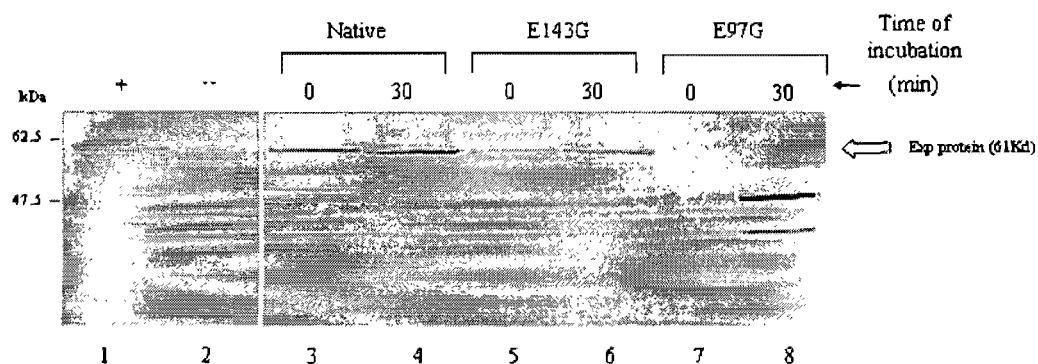
Figure 3-26: Protein Expression Following Copper Induction

Comparison of expression levels following the introduction of the mutant constructs in yeast. Genes in question are under the control of a copper inducible promoter. To optimize expression, cells were incubated with a specific concentration of copper (1 mM) for 30 minutes. Following incubation, reactions were halted and the cells were pelleted and subsequently lysed to form protein extracts. Subsequent 35 µg samples were probed with anti-HIS to detect expression. Lane 1, positive control (extracts expressing previously used GST-HIS-CeAPN-1); lane 2, negative control (extracts stemming from parent strain), lanes 3-8, extracts collected following the introduction of the native construct and the respective site directed mutants (E143G and E97G) with copper induction (0 min and 30 min).

3.9.2. Gradient Plate Analysis of Variant and Native Proteins:

The corresponding strains expressing the variant proteins were thereafter tested, via cross species complementation analysis, in order to establish whether the residue substitutions have an impact on the ability to process damaged DNA. This was achieved by allowing the strains to grow along increasing gradients composed of MMS, H_2O_2 and Bleomycin respectively. Thus, as highlighted earlier, MMS was used due to its alkylating properties and subsequent indirect role in forming AP sites. H_2O_2 was included because it causes strand breaks with 3'-blocked ends. Lastly, Bleomycin was also exploited due to its strong oxidant capacities to produce single strand breaks forming blocked ends (i.e. 3'-phosphate and 3' phosphoglycolate termini), double strand breaks and AP sites [71, 72]. Hence, Fig 3-27 shows the wild-type strain YW465, containing normal AP endonuclease activity, is able to grow the full length of the MMS gradient. The mutant YW778 strain lacking AP endonuclease activity (both Endo IV and Exo III components), in contrast, is unable to process the created AP sites and therefore, fails to grow a significant length on the plate. Strains transformed with both site directed mutants, Ceapn-1(E143G) and Ceapn-1(E97G), are equally unable to circumvent the increasing MMS gradient and fail to restore parental resistance. Thus, the gradient test confirms that the variant E143G, although expressed (unlike E97G), carries a significant defect which renders it unable to complement for yeast Apn1 and is functionally inactive. However, as shown in Fig 3-5, the native form of the protein demonstrates that it is able to rescue the mutant strain (lacking yApn1 and yApn2). This general phenomenon is further evident on both oxidant gradients, H_2O_2 and Bleomycin, as shown in Fig 3-28. As expected, the strain expressing the native

recombinant CeAPN-1 is able to restore parental resistance to the DNA damaging agents in question. The mutant strains, however, sustain a level of growth which is inferior to that of the parental strain.

Host Strain	Introduced Plasmid
YW465	--
YW778	--
YW778	pTW340
YW778	pGST-His-Ceapn-1
YW778	pE97G
YW778	pE97G
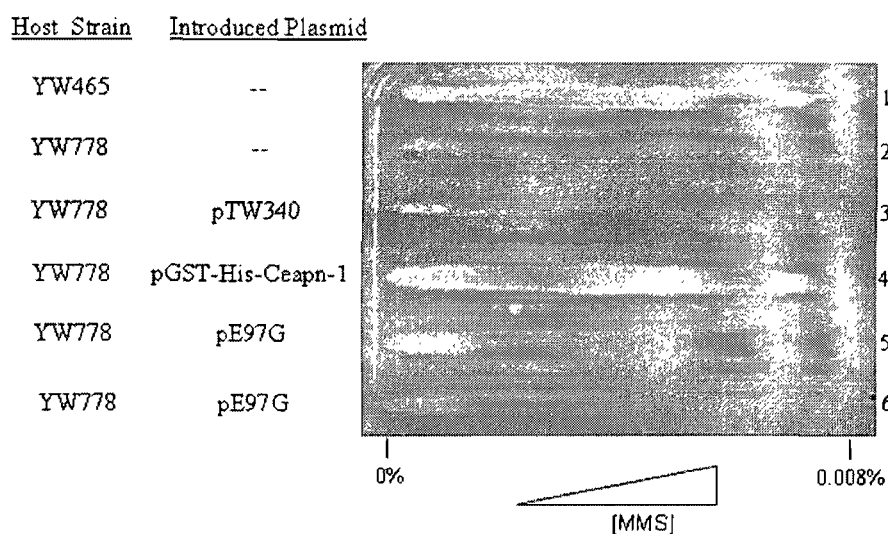
Figure 3-27: Cross Species Complementation Analysis Using MMS
An increasing gradient of MMS (with fixed 0.5 mM copper) was utilized to assess strain sensitivity following introduction of the native and site directed mutant constructs. Lane 1, wild type strain, containing AP endonuclease activities is able to grow the full length of the gradient. However, the AP endonuclease deficient strain (YW778) grows a very minimal distance (lane 2). Lane 3, similarly demonstrates minimal growth when the 778 strain is introduced with pTW340 which bears *OGG1* (same backbone as pCeapn-1 constructs). Lane 4 reveals parallel growth to that of the wild type as Ceapn1 is introduced into the mutant strain. Lanes 5-6, the respective introduction of the mutants is unable to demonstrate cross species complementation.

Host Strain	Introduced Plasmid		
YW465	--		1
YW778	--		2
YW778	pTW340		3
YW778	pGST-His-Ceapn-1		4
YW778	pE143G		5
YW778	pE97G		6
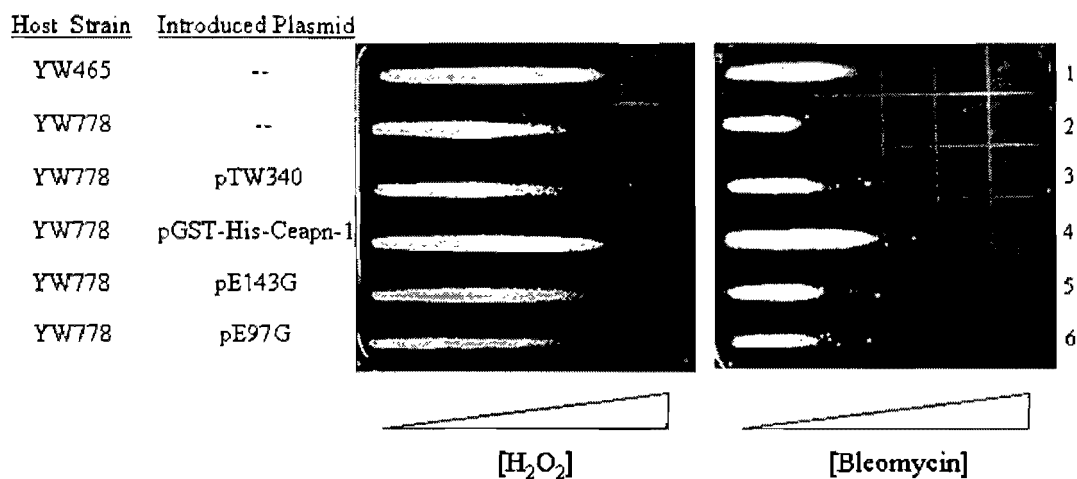
[H₂O₂]   [Bleomycin]

Figure 3-28: Gradient Plate Analysis Using H₂O₂ and Bleomycin

Sensitivity was monitored similarly to as shown in fig 3-27 but on increasing gradients composed of H_2O_2 (4.0 mM) and Bleomycin (200 µg). The AP endonuclease deficient strain (YW778, lane 2) upholds an inferior level of survival to that of the parent strain (YW465, lane1) with both DNA damaging agents. This effect is maintained following the introduction of pTW340, bearing *OGG1* (same backbone as pCeapn-1 constructs) as seen lane 3. Lane 4 reveals that the expression of CeAPN-1 restores parental resistance towards the repair of the respective strand breaks caused by the agents in question. Lanes 5-6 show that the introduction of the constructs bearing Ceapn-1 mutations cause a similar level of growth to that observed with YW778.

4.0. DISCUSSION

The discovery of CeAPN-1 in *C. elegans* corresponds to the first member of the Endo IV AP endonuclease family to be found in a higher eukaryote, representative of an animal model [84]. This particular organism was, in fact, shown to be composed of both Exo III and Endo IV families; a unique composition that has previously never been observed in a multicellular system [84]. Despite the fact that a mammalian member of the Endo IV family has yet to be found, it is believed that the study of an enzyme from a metazoan model would provide significant clues on the existence of a potential homologue in mammals. Such homology would likely arise from a functional rather than a strictly structural identity.

Although the long term goals in the characterization of CeAPN-1 rest on *in-vivo* studies, a classical biochemical approach remains primordial in the understanding of this enzyme and its true biological role. The approach taken in this study is fundamental in the determination of the types of lesions the protein can directly process, the necessity of additional factors for activity and a comparison in the degree of similarity it shares with other Endo IV family homologues. As a result, in order to perform a classical *in-vitro* biochemical characterization of CeAPN-1, the first hurdle remains the ability to obtain a pure fraction of the protein in question.

4.1. Utilization of His Affinity Tag

CeAPN-1 was initially decided to be tagged with a His tag due to the fact that it is much smaller than other affinity tags and therefore, less likely to interfere with enzyme activity. As mentioned in the introduction, the utilization of a larger tag (GST) was believed to negate enzyme function [86]. Evidently, a smaller tag of 6 amino acid

repeats (His) would be less susceptible to directly affect proper protein folding. It should further be recognized that the affinity tag was fused to the N-terminal end of the gene of interest; thus, previous studies with corresponding homologue proteins showed no inhibition of activity when fusing tags to the N-terminal [79].

Despite the creation of the fusion protein, many difficulties were encountered in order to achieve full purity. Routinely, multiple proteins had been observed in collected fractions following elution steps. It should be noted that these fractions in question, although not homogeneous (multiple signals observed in sypro stained gels), did demonstrate to be active in repairing damaged DNA (AP sites and 3'-blocked ends). The identity of the contaminating bands, observed via sypro staining, may have been suggested to be due to particular components binding very tightly to the protein of interest. Another hypothesis to explain the excessive banding pattern following column elution could stem from the possibility that CeAPN-1 may be undergoing certain post translational modifications, which may include its degradation. However, these situations were not encountered while studying other proteins of the same family, notably, *E. coli* Endo IV and *S. cerevisiae* Apn1. As a result, it was strongly believed that the contaminating proteins arose from non-specific interactions with the His affinity (Talon) resin. Experiments performing a parallel purification using the transformed yeast strain expressing CeAPN-1 (YW778 / pCeapn-1) and the untransformed AP endonuclease deficient background (YW778) have later proven that the vast majority of the contaminating bands can be accounted as being non-specific.

Numerous approaches and measures had been attempted to achieve full purity for subsequent characterization. Thus, extracts were progressively fractionated, pH was fine tuned, multiple column resins were exploited and the configuration of the construct

was modified (implementing His tag to C-terminal end). Nonetheless, the final result commonly yielded very little protein if any and sometimes produced inactive fractions. It should be noted that two constant hurdles had been encountered throughout the various purification attempts: firstly, the challenge in purifying a protein that was not highly expressed in the system used, and secondly, the inability of the recombinant protein to strongly bind to the affinity beads. As a result, new approaches had been envisioned to overcome these hurdles by possibly changing expression systems or adding another tag to the single tag construct. As a result, very little difficulty was encountered using the bacterial system to express and purify the yeast Apn1 homologue. It should be noted that a similar approach (using *E. coli*) had been initially adopted for the characterization of CeAPN-1. Nonetheless, the expressed product (20 kDa) obtained from bacteria did not correlate with the expected size (30 kDa) and prevented thorough characterization from being performed [55]. It was, in fact, believed that the *C. elegans* protein was truncated through a proteolytic process in the bacterial expression system [55]. Consequently, the utilization of insect cells via a baculovirus system was thought to be a more viable option to express the recombinant worm protein. This type of system has become very prevalent in recent years for the expression and purification of specific proteins. In essence, the benefits of insect cells arise from the fact that they represent a more advanced eukaryotic system which is able to conduct post translational modifications [99]. Moreover, insect cells provide an environment that allows proper protein folding. It is also very important to note that insect cells are known to be able to express proteins at high levels [99]. A major disadvantage, however, with the utilization of insect cells lies in the absence of a specific test or assay that permits the evaluation of

whether functional cross species complementation can occur with the introduction of a foreign gene [61].

4.1.1. Utilization of a GST Affinity Tag:

A second option resided in the preservation of a yeast expression system while changing the promoter to optimize the expression level. Moreover, in doing so, it was also deemed that the subsequent purification steps may be rendered more efficient by fusing an additional GST tag to the His-Ceapn-1 or Ceapn-1-His genetic constructs. GST tagged proteins are known to be easily purified via a glutathione sepharose matrix. Thus, glutathione-S- transferase has a strong affinity for glutathione and allows the recombinant protein to bind tightly to the resin while non-specific components are washed away. The protein of interest is eluted under mild conditions using reduced glutathione.

Interestingly, a GST system was initially avoided since it was equally utilized in the past to purify a recombinant form of the CeAPN-1 protein. As a result, the introduction of the pGST-Ceapn-1 in an AP endonuclease deficient yeast strain (YW778) failed to demonstrate complementation and activity. This inability was originally believed to be linked to the possibility that the tag may have interfered with proper protein folding. However, it was learned only thereafter that the utilized construct lacked a key N-terminal domain which resembled very closely a nuclear localization signal (NLS). Hence, it was later assumed that the inability of the CeAPN-1 recombinant to substitute for yApn1 was linked more largely on the absence of the NLS-like domain rather than the fused GST tag.

A P_{cup} copper inducible promoter was exploited to drive the expression of the resulting GST-tagged sequence. It should be noted that the His-Ceapn-1 fragment was recombined into the pTW340 backbone (containing the P_{cup} promoter and GST tag). Evidently, the final expressed product results in a dual tagged protein composed of HIS and GST affinity tags. The presence of two affinity tags coupled with the insertion of an NLS was thought to concurrently alleviate the purification process and preserve the protein activity respectively. Indeed, the dual tagged protein was purified with relative ease, in comparison to previous attempts, using a single step approach via the glutathione resin. The resulting SDS stained gel revealed the presence of minimal contaminating species being co-eluted with the protein of interest.

4.2. Activity Assays of GST Tagged Protein (GST-HIS-CeAPN-1)

Despite the purity of the collected fraction, *in-vitro* assays revealed that the enzyme lacked activity. Initially, assay conditions were modified by re-evaluating the ionic strength and various reaction conditions; however, these efforts did not produce substrate turnover (data not shown). It should be noted that the crude extracts did demonstrate activity, which suggested that the activity of the purified protein may be, in fact, lost during the purification process. Accordingly, it was assumed that a possibility existed that CeAPN-1 may require a particular component to stimulate its activity. Thus, extracts from the YW778 strain (AP endonuclease deficient) were introduced to the reaction with the purified GST recombinant protein.

Interestingly, the addition of the AP endonuclease deficient extracts did appear to dramatically stimulate the activity capacity of the enzyme. Thus, enzyme activity can be observed with the addition of as little as 10 ng of total extract. Undoubtedly, the

presence of the extracts alone (void of AP endonuclease) is unable to process the damaged DNA being utilized in this assay.

Further experimentation was performed to assess the observed phenomenon causing activity stimulation in the presence of YW778 extracts. Thus, to evaluate the possibility that a protein may be involved in the stimulation of CeAPN-1, extracts were rigorously heated (95°C, 5 minutes); thus, inactivating (denaturing) most proteins in the extracts. The pre-treated extracts were added in reaction with purified GST-HIS-CeAPN-1 and assayed for AP endonuclease activity. It was believed that the absence of activity would promote the notion that CeAPN-1 requires a particular polypeptide to ensure AP endonuclease activity.

However, despite the addition of heat treated extracts to the purified CeAPN-1 fusion protein, the stimulation of enzyme activity was continued to be observed. Thus, collected results revealed a distinct band characteristic of the converted 20mer, following the addition of heat treated extracts, which paralleled very closely to what was observed in the presence of non-heated extracts. The substrate turnover continued to increase with the concentration of extracts (up to 1000 ng) while substrate and CeAPN-1 levels were maintained.

The continued stimulation caused by the addition of heat treated extracts allows one to conclude that a protein would not likely be the necessary agent behind the phenomenon. Larger proteins would be more easy to rule out since they are composed of multiple folding 'signatures' and consequently, it would be highly improbable that they would have the opportunity to re-fold correctly. Thus, a vastly smaller protein, consisting of minimal 'folding signatures' may be more capable of promptly re-folding into its native state following denaturing conditions.

An initial possible candidate that could be causing this activity stimulation may be linked to a metal factor. Thus, Exo III family members have showed dependence towards Mg^{2+} ions in order to render these proteins active. It could be assumed that CeAPN-1 requires a specific metal factor for activity which may, in fact, differentiate it from other studied Endo IV members in yeast and bacteria. To test this hypothesis, YW778 extracts were dialyzed against a buffer free of metal ions and subsequently added to reaction with purified GST-HIS-CeAPN-1. Furthermore, a second experiment aimed at observing whether any activity stimulation would exist following the addition of a fixed concentration of metal ions (increasing concentration) in reaction with the purified recombinant protein.

As seen with heat treated extracts, the addition of dialyzed extracts maintained an ability to stimulate the catalytic function of the recombinant protein. Moreover, the addition of individual metal ions, notably, Zn^{2+}, Co^{2+}, Mn^{2+}, did not demonstrate any ability whatsoever to positively affect enzyme activity. Therefore, it could be suggested that divalent metal ions are probably not responsible for the stimulation of CeAPN-1 activity.

4.3. Conclusions Stemming from Activity Assays:

Preliminary data suggests that CeAPN-1 requires a particular component to increase its activity level. This phenomenon was revealed by the addition of YW778 extracts to purified protein during activity assays. The stimulus was originally thought to be linked to three possible candidates: a metal factor, protein or DNA/RNA (not investigated in this study). No indication has thus far been observed to indicate that a metal ion is causing the activity stimulation. As highlighted above, experiments

utilizing dialyzed extracts and the addition of individual factors (Zn^{2+}, Co^{2+}, Mn^{2+}) with a purified fraction of CeAPN-1 failed to convey that a metal ion is implicated. A probable candidate could be a very small protein comprised of heat stable properties. Although, the addition of boiled extract to purified protein was not shown to reduce the extent of activity stimulation, this does not entirely refute the possibility that a protein may be at play. Thus, a protein candidate would have a low molecular weight and the ability to regain its native form immediately following significant thermal treatment. As a result, the exact role or mechanism governing this overall process is unknown. If a protein is involved, it must have some role in allowing GST-CeAPN-1 to access the lesions in question with greater ease. It is unlikely that pH or buffer composition is the cause of this phenomenon; however, additional experimentation is required to fully rule out these possibilities.

Interestingly, it is widely documented in the literature that certain proteins require small molecular weight polypeptides or nucleic acids for their activity. In fact, Ruzindana-Umunyana *et al.* 2000 has reported that adenain, a cysteine protease responsible for uncoating and virion maturation in adenoviruses, requires an 11 amino acid peptide cofactor in order to stabilize its optimal conformation for enzyme function [100]. In another report, it has been similarly demonstrated by Abdel-Ghany *et al.* 1984 that a polypeptide is needed for the activity of a protein kinase present in the plasma membrane of Ehrlich ascites tumor cells. The activating component was shown to be heat stable and believed to be identical to histone 1 [101].

An interesting experiment to shed light on the activity stimulation would be via the utilization of a potent protease (i.e. trypsin) to cleave all proteins present in the YW778 extracts. The treated extracts could thereafter be boiled, added to purified

recombinant forms of CeAPN-1 and finally, assayed for activity. As a result, the absence of activity would promote the notion that a protein is indeed involved. Measures may be devised to isolate the protein in question by running the extracts in a series of purification columns (i.e. anion exchange, gel filtration) and evaluating the presence of AP activity following each step. Ideally, once fully purified, the protein may be sent for mass spectrometry analysis in order to identify the component in question. It would also be intriguing to use heat treated extracts stemming from different models (i.e. mammalian cells and bacteria) to observe if stimulation is maintained following the addition of purified protein. Thus, an important evaluation to perform would lie in the demonstration of whether worm extracts deficient in AP endonuclease activity are able to confer activity stimulation. The confirmation of protein involvement would subsequently make the utilization of siRNA an interesting approach. In essence, this technique would allow one to observe the extent at which worms can counter DNA damage in the absence of this particular protein that may be involved.

Nonetheless, the continued presence of activity following the addition of protease/heat treated extracts may indicate that the stimulation is the result of other possibilities. DNA and RNA may become potential candidates triggering the dual activity stimulation of the recombinant protein. One can suggest that the nucleic acids may bind to the protein and contribute to changing the conformation of the polypeptide. Nucleic acids have been observed to be an important component for the activity of particular proteins. Such is the case demonstrated by DNA protein kinase (DNA-PK), a nuclear serine/threonine kinase, which requires double stranded DNA for its activity [102]. The possibility of nucleic acids being required by CeAPN-1 could be analyzed

directly by separately incubating the YW778 extracts with DNase and RNase and thereafter, evaluating whether the addition of the treated sample with purified CeAPN-1 stimulates AP endonuclease activity.

As discussed in the introduction Endo IV members have been shown to be multifunctional. Although not completed during this specific experimental trial, it would undoubtedly be an interesting goal to assess whether CeAPN-1 is also equally endowed with such activities (i.e. NIR and 3'→5' exonuclease). Moreover, having showed the possibility of a thermally stable component stimulating AP endonuclease function, this would naturally raise the question of whether this phenomenon is consistent with the other biological activities associated with Endo IV members.

4.4. Domain Function Analysis:

Close inspection of the E. coli endo IV crystal structure has allowed one to recognize that the N-terminal and C-terminal portions of the protein are fairly close to each other [68]. The active site is composed of positively charged amino acids (arginine and lysine residues) and a cluster of Zn^{2+} ions which forms the positive core (within the groove) that grants entry to the negatively charged DNA strand. The most conserved region of amino acids shared between the various Endo IV family members is known to be located between residues Asp 179 and His182 [68]. The structural data appears to suggest that this group of amino acids provides stabilization to the beta barrel and, in turn, to the tri-nuclear zinc cluster [68].

Site directed mutagenesis studies have been performed with the yeast Apn1 and the bacterial endo IV proteins [88, 93]. Although fairly conserved, two particular mutations have shown that the yeast protein differs from its bacterial counterpart.

Hence, a variant of the yApn1 protein, specifically consisting of an amino acid substitution of the Glu158 residue with Gly, yielded a 150 fold decrease in activity [88, 93]. Nonetheless, additional experiments have showed that the E158G variant does retain an ability to bind to DNA [88]. Interestingly, the corresponding variant form of endo IV, E145G, also showed a 150 fold reduction in enzymatic activity but revealed no capacity for DNA binding [93]. Consequently, it is believed that the E145 residue in the endo IV protein has a dual role involved in both catalysis and substrate binding [93]. The equivalent residue, E158, in yeast Apn1 may, in contrast, be solely involved in enzymatic activity via its coordination of the specific zinc ion [88].

As a result, as mentioned earlier, one of the principle goals of this study wished to investigate the structural and domain features CeAPN-1 shares or differentiates it from its counterparts. Undoubtedly, having 44.9 % amino acid identity with the *E. coli* endo IV, the structure of CeAPN-1 is thought to resemble very closely that of its bacterial homologue [84]. As a result, the Swiss Model application enables an interesting rough visualization of the *C. elegans* protein based on its homology to endo IV. The program confirms a strong likelihood that CeAPN-1 is also composed of similar features which include a structure of 8 alpha helices and 8 beta strands believed to be essential in substrate binding and a fairly large pocket for catalysis [68, 94, 96, 97].

In this analysis, two variant proteins were produced and compared to the *C. elegans* native protein and the other homologue variants (yeast and bacteria). Thus, the first amino acid substitution was conducted at the Asp143 (E143) position to produce Glycine143 (E143G). The E143 residue was chosen due to the fact that it was thought to be a key player in the coordination of the Zinc cluster found in the active site [68]. As

described earlier, substitution of the corresponding residue in the bacterial endo IV (E145G) and yeast Apn1 (E179G) has produced interesting data to demonstrate a divergence of the respective homologue proteins and the individual roles the residues play in catalysis and DNA binding [93] [88]. Thus, it would be intriguing to demonstrate whether the worm homologue resembles more the bacterial endo IV or the yeast Apn1 counterpart.

In addition, a second amino acid substitution was conducted at the E97 position. Very little is actually known about this particular residue despite the fact that is found in the first conserved domain amongst the Endo IV member sequences. Hence, unlike E143, corresponding mutations have not been conducted while studying bacterial endo IV and yeast Apn1. Moreover, a compelling aspect justifying the investigation of E97 arises from its distinct features. Along with R95, it is the only charged residue which is found in homology region I that is conserved in all four models (*E. coli*, *S. pombe*, *S. cerevisiae* and *C. elegans*) (referred to the alignment given in [55]). Two separate amino acid substitutions found in the first conserved region, S87P and C98R (via chemical mutagenesis), have been conducted using *E. coli* endo IV [93]. Interestingly, the introduction of the former variant (S87P) in an AP endonuclease deficient strain has shown to cause partial complementation in the presence of specific DNA damaging agents. On the other hand, the latter alteration, C98R, fails to cause any complementation under the same conditions [93]. However, the Ceapn-1 sequence, unlike what is observed in *S. cerevisiae* and *E. coli*, does not contain a corresponding Cysteine but an Alanine residue (refer to alignment given in [55]).

Preliminary results collected following the introduction of the E143G mutation to the *C. elegans* gene, via site directed mutagenesis, is in accordance with what was

reported with the yeast and bacterial variants [88, 93]. A gradient plate assay, in fact, indicated that a strain with the introduced E143G mutation was unable to complement for yApn1 when exposed to an increasing concentration of MMS. It should be noted that a Western Blot analysis revealed that the E143G variant was being stably expressed. Unfortunately, further work was not conducted due to the inability to purify to homogeneity a sufficient amount for subsequent characterization studies.

In contrast, the E97G variant was not detected at its expected size via Western blot analysis by the anti-GST antibody (presence of lower molecular weight bands reacting with antibody may be indicative of degradation). The failure to identify the presence of E97G (at predicted molecular weight) could be due to the nature of the substitution which may be causing a dispersal of the acidic cluster surrounding the residue; consequently, producing a structural instability leading to the protein's rapid degradation.

4.5. Conclusion and Future Perspectives:

Deficiencies in DNA repair processes are known to trigger the onset of disease and aging in humans. BER is one of the most important repair pathways since it takes part in the processing of the most common insults affecting DNA. Despite the fact that repair mechanisms overlap in terms of their functions, defects in specific components of BER have been correlated to impairing survival. In humans, defects in the glycosylase, MYH, has been linked to familial adenomatous polyposis and multiple colorectal adenomas (with or without carcinoma) [103, 104]. Studies have as well indicated a susceptibility to cancer in subjects having abnormal variants of the Polβ protein or loss of the OGG1 genetic locus [105-107]. Alternatively, much has yet to be learned

concerning the polymorphic variants of the human AP endonuclease, APE1 (Exo III family). Although, it has been reported that such variants have been correlated to impairing normal functioning, its precise association to cancer is still not clearly displayed [108, 109]. However, it should be noted that APE1 haploinsufficient mice have been documented to have high spontaneous mutagenesis, increased tendency towards cancer and, consequently, decreased survival. The APE1 double knockout mouse, in contrast, is embryonically lethal. Undoubtedly, the continued study of BER and its various components is believed to enhance our understanding of various diseases such as cancer [110, 111].

4.5.1 The Study of a C. elegans Protein

The finding of APN-1 in C. elegans represented the first member of the Endo IV family to be found in a multicellular animal model. Although, a human structural homologue of the Endo IV family has not been discovered yet, it is believed that a functional counterpart does indeed exist [55]. Thus, the characterization of an APN-1 member stemming from a metazoan model was believed to serve as a first step to potentially provide clues concerning a possible functional human homologue. Interestingly, such a study involving the C. elegans protein could provide additional insight on the corresponding Endo IV protein in the malaria parasite Plasmodium falciparum. Better understanding of the AP endonuclease enzyme could enable the elucidation of a pathway in the parasite that may serve as a potential therapeutic target against malaria [112].

Furthermore, the characterization of the C. elegans protein was also thought to shed some light on DNA repair mechanisms associated with the multicellular organism.

Contrary to bacteria and yeast, the worm model has, in fact, not been as widely utilized in the study of DNA repair. Increased awareness of *C. elegans* may potentially promote its value as a model/marker for environmental toxicity studies [113].

4.5.2 Future Work

The work conducted in this study serves as a first step in the understanding of CeAPN-1. Thus, this particular investigation sought to characterize CeAPN-1 *in-vitro* through classical biochemical means. It was found that the functional activity of the polypeptide may be stimulated via an extra component. The identity of this molecule, however, remains to be fully determined. As highlighted earlier, via the means of purification and mass spectrometry, one could potentially identify or at least better understand the process causing this particular enhancement in activity.

The next step, beyond this specific study, lies in the *in-vivo* characterization of the protein. Hence, a major objective in this type of investigation consists in demonstrating how mutations may accumulate in worms spontaneously and when exposed to DNA damaging agents when CeAPN-1 and CeEXO-3 expression is individually repressed. This worm response could be compared to a situation of dual repression involving both AP endonucleases.

Furthermore, as explained in the introduction, it should be stressed that Ceapn-1 is composed of a large N-terminal domain (118 residues) which distinguishes itself from the other members of the Endo IV family in two aspects. Thus, the N terminus region is thought to act as a nuclear localization signal which is indicated by the highly concentrated presence of basic residues (characteristic of NLS) [87]. Unlike yeast Apn1, having a bipartite NLS at the C-terminus, the *C. elegans* protein may be composed of an

NLS found at the N-terminus [78]. Moreover, the presumed NLS region within the Ceapn-1 gene appears to differentiate from its counterpart in its sheer size of approximately 118 amino acids. The large size of this domain suggests that it may play a broader role than first thought (acting solely as a nuclear localization signal). Hence, the domain might play a role in recognizing or accessing lesions as well as possibly being implicated in targeting the protein to different organelles (i.e. the mitochondrion) [79]. The investigation of the N-terminal domain could begin by exploiting GFP tagged constructs and analyzing the localization of the expressed product in worms through fluorescence microscopy [79]. One could also create mutations within the N-terminal domain to characterize the key residues and confirm the fundamental underlying importance of the region [78].

Being a DNA repair protein, it would also be interesting to determine if CeAPN-1 is able to associate to other proteins, notably transcription factors or other components of BER. Such interactions would not be out of the realm of possibility; in fact, the human AP endonuclease, Ape1, has been in fact correlated to interact with, amongst others, PCNA and specific glycosylases (MYH) [34]. Moreover, previous studies have shown yApn1 interacting with Pir1 for transport into the mitochondrion [79]. As a result, an analysis in search of similar phenomena may be designed using a series of IP experiments with *C. elegans* extracts.

5. REFERENCES

1.   Rass, U., I. Ahel, and S.C. West, *Defective DNA repair and neurodegenerative disease.* Cell, 2007. **130**(6): p. 991-1004.

2.   Friedberg, E.C. and E.C. Friedberg, *DNA repair and mutagenesis.* 2nd ed. 2006, Washington, D.C.: ASM Press. xxix, 1118 , [12] de pl.

3.   Singer, B. and B. Hang, *What structural features determine repair enzyme specificity and mechanism in chemically modified DNA?* Chem Res Toxicol, 1997. **10**(7): p. 713-32.

4.   Verly, W.G., *Repair of AP sites in DNA.* Biochimie, 1982. **64**(8-9): p. 603-5.

5.   Johnson, R.E., et al., *Identification of APN2, the Saccharomyces cerevisiae homolog of the major human AP endonuclease HAP1, and its role in the repair of abasic sites.* Genes Dev, 1998. **12**(19): p. 3137-43.

6.   Karihtala, P. and Y. Soini, *Reactive oxygen species and antioxidant mechanisms in human tissues and their relation to malignancies.* Apmis, 2007. **115**(2): p. 81-103.

7.   Wiseman, H. and B. Halliwell, *Damage to DNA by reactive oxygen and nitrogen species: role in inflammatory disease and progression to cancer.* Biochem J, 1996. **313 (Pt 1)**: p. 17-29.

8.   Vance, J.R. and T.E. Wilson, *Repair of DNA strand breaks by the overlapping functions of lesion-specific and non-lesion-specific DNA 3' phosphatases.* Mol Cell Biol, 2001. **21**(21): p. 7191-8.

9.   Povirk, L.F. and C.W. Houlgrave, *Effect of apurinic/apyrimidinic endonucleases and polyamines on DNA treated with bleomycin and neocarzinostatin: specific formation and cleavage of closely opposed lesions in complementary strands.* Biochemistry, 1988. **27**(10): p. 3850-7.

10. Boiteux, S. and M. Guillet, *Abasic sites in DNA: repair and biological consequences in Saccharomyces cerevisiae.* DNA Repair (Amst), 2004. **3**(1): p. 1-12.

11. Peltomaki, P., *DNA mismatch repair and cancer.* Mutat Res, 2001. **488**(1): p. 77-85.

12. Yu, Z., et al., *Human DNA repair systems: an overview.* Environ Mol Mutagen, 1999. **33**(1): p. 3-20.

13. Felton, K.E., D.M. Gilchrist, and S.E. Andrew, *Constitutive deficiency in DNA mismatch repair.* Clin Genet, 2007. **71**(6): p. 483-98.

14. Limp-Foster, M. and M.R. Kelley, *DNA repair and gene therapy: implications for translational uses.* Environ Mol Mutagen, 2000. **35**(2): p. 71-81.

15. Weterings, E. and D.C. van Gent, *The mechanism of non-homologous end-joining: a synopsis of synapsis.* DNA Repair (Amst), 2004. **3**(11): p. 1425-35.

16. Couedel, C., et al., *Collaboration of homologous recombination and nonhomologous end-joining factors for the survival and integrity of mice and cells.* Genes Dev, 2004. **18**(11): p. 1293-304.

17. Ferreira, M.G. and J.P. Cooper, *Two modes of DNA double-strand break repair are reciprocally regulated through the fission yeast cell cycle.* Genes Dev, 2004. **18**(18): p. 2249-54.

18. Shrivastav, M., L.P. De Haro, and J.A. Nickoloff, *Regulation of DNA double-strand break repair pathway choice.* Cell Res, 2008. **18**(1): p. 134-47.

19. Wyman, C., D. Ristic, and R. Kanaar, *Homologous recombination-mediated double-strand break repair.* DNA Repair (Amst), 2004. **3**(8-9): p. 827-33.

20. Phillips, E.R. and P.J. McKinnon, *DNA double-strand break repair and development.* Oncogene, 2007. **26**(56): p. 7799-808.

21. Wu, X. and Z. Wang, *Relationships between yeast Rad27 and Apn1 in response to apurinic/apyrimidinic (AP) sites in DNA.* Nucleic Acids Res, 1999. **27**(4): p. 956-62.

22. Wijnhoven, S.W., et al., *Tissue specific mutagenic and carcinogenic responses in NER defective mouse models.* Mutat Res, 2007. **614**(1-2): p. 77-94.

23. Sarasin, A. and A. Stary, *New insights for understanding the transcription-coupled repair pathway.* DNA Repair (Amst), 2007. **6**(2): p. 265-9.

24. Wyatt, M.D., et al., *3-methyladenine DNA glycosylases: structure, function, and biological importance.* Bioessays, 1999. **21**(8): p. 668-76.

25. Hitomi, K., S. Iwai, and J.A. Tainer, *The intricate structural chemistry of base excision repair machinery: implications for DNA damage recognition, removal, and repair.* DNA Repair (Amst), 2007. **6**(4): p. 410-28.

26. Krokan, H.E., R. Standal, and G. Slupphaug, *DNA glycosylases in the base excision repair of DNA.* Biochem J, 1997. **325 (Pt 1)**: p. 1-16.

27. Memisoglu, A. and L. Samson, *Base excision repair in yeast and mammals.* Mutat Res, 2000. **451**(1-2): p. 39-51.

28. Fortini, P., et al., *Different DNA polymerases are involved in the short- and long-patch base excision repair in mammalian cells.* Biochemistry, 1998. **37**(11): p. 3575-80.

29. Almeida, K.H. and R.W. Sobol, *A unified view of base excision repair: lesion-dependent protein complexes regulated by post-translational modification.* DNA Repair (Amst), 2007. **6**(6): p. 695-711.

30. Liu, Y., H.I. Kao, and R.A. Bambara, *Flap endonuclease 1: a central component of DNA metabolism.* Annu Rev Biochem, 2004. **73**: p. 589-615.

31. Caldecott, K.W., et al., *An interaction between the mammalian DNA repair protein XRCC1 and DNA ligase III.* Mol Cell Biol, 1994. **14**(1): p. 68-76.

32. Kubota, Y., et al., *Reconstitution of DNA base excision-repair with purified human proteins: interaction between DNA polymerase beta and the XRCC1 protein.* Embo J, 1996. **15**(23): p. 6662-70.

33. Gellon, L., et al., *Intrinsic 5'-deoxyribose-5-phosphate lyase activity in Saccharomyces cerevisiae Trf4 protein with a possible role in base excision DNA repair.* DNA Repair (Amst), 2008. **7**(2): p. 187-98.

34. Hang, B. and B. Singer, *Protein-protein interactions involving DNA glycosylases.* Chem Res Toxicol, 2003. **16**(10): p. 1181-95.

35. Flaherty, D.M., M.M. Monick, and G.W. Hunninghake, *AP endonucleases and the many functions of Ref-1.* Am J Respir Cell Mol Biol, 2001. **25**(6): p. 664-7.

36. Unk, I., et al., *Stimulation of 3'-->5' exonuclease and 3'-phosphodiesterase activities of yeast apn2 by proliferating cell nuclear antigen.* Mol Cell Biol, 2002. **22**(18): p. 6480-6.

37. Marti, T.M. and O. Fleck, *DNA repair nucleases.* Cell Mol Life Sci, 2004. **61**(3): p. 336-54.

38. Ramotar, D., et al., *Schizosaccharomyces pombe apn1 encodes a homologue of the Escherichia coli endonuclease IV family of DNA repair proteins.* Biochim Biophys Acta, 1998. **1396**(1): p. 15-20.

39. Ribar, B., T. Izumi, and S. Mitra, *The major role of human AP-endonuclease homolog Apn2 in repair of abasic sites in Schizosaccharomyces pombe.* Nucleic Acids Res, 2004. **32**(1): p. 115-26.

40. Roldan-Arjona, T., C. Anselmino, and T. Lindahl, *Molecular cloning and functional analysis of a Schizosaccharomyces pombe homologue of Escherichia coli endonuclease III.* Nucleic Acids Res, 1996. **24**(17): p. 3307-12.

41. Tanihigashi, H., et al., *The role of Schizosaccharomyces pombe DNA repair enzymes Apn1p and Uve1p in the base excision repair of apurinic/apyrimidinic sites.* Biochem Biophys Res Commun, 2006. **347**(4): p. 889-94.

42. Fraser, J.L., E. Neill, and S. Davey, *Fission yeast Uve1 and Apn2 function in distinct oxidative damage repair pathways in vivo.* DNA Repair (Amst), 2003. **2**(11): p. 1253-67.

43. Duncan, B.K. and B. Weiss, *Specific mutator effects of ung (uracil-DNA glycosylase) mutations in Escherichia coli.* J Bacteriol, 1982. **151**(2): p. 750-5.

44. Kavli, B., et al., *Excision of cytosine and thymine from DNA by mutants of human uracil-DNA glycosylase.* Embo J, 1996. **15**(13): p. 3442-7.

45. Sakumi, K., et al., *Purification and structure of 3-methyladenine-DNA glycosylase I of Escherichia coli.* J Biol Chem, 1986. **261**(33): p. 15761-6.

46. O'Brien, P.J. and T. Ellenberger, *The Escherichia coli 3-methyladenine DNA glycosylase AlkA has a remarkably versatile active site.* J Biol Chem, 2004. **279**(26): p. 26876-84.

47. Izumi, T., et al., *Molecular cloning and characterization of the promoter of the human N-methylpurine-DNA glycosylase (MPG) gene.* Carcinogenesis, 1997. **18**(9): p. 1837-9.

48. Chen, J., B. Derfler, and L. Samson, *Saccharomyces cerevisiae 3-methyladenine DNA glycosylase has homology to the AlkA glycosylase of E. coli and is induced in response to DNA alkylation damage.* Embo J, 1990. **9**(13): p. 4569-75.

49. O'Brien, P.J. and T. Ellenberger, *Human alkyladenine DNA glycosylase uses acid-base catalysis for selective excision of damaged purines.* Biochemistry, 2003. **42**(42): p. 12418-29.

50. Dizdaroglu, M., *Base-excision repair of oxidative DNA damage by DNA glycosylases.* Mutat Res, 2005. **591**(1-2): p. 45-59.

51. Cadet, J., et al., *Oxidative base damage to DNA: specificity of base excision repair enzymes.* Mutat Res, 2000. **462**(2-3): p. 121-8.

52. O'Connor, T.R., et al., *Fpg protein of Escherichia coli is a zinc finger protein whose cysteine residues have a structural and/or functional role.* J Biol Chem, 1993. **268**(12): p. 9063-70.

53. Bailly, V. and W.G. Verly, *Possible roles of beta-elimination and delta-elimination reactions in the repair of DNA containing AP (apurinic/apyrimidinic) sites in mammalian cells.* Biochem J, 1988. **253**(2): p. 553-9.

54. Sakumi, K. and M. Sekiguchi, *Structures and functions of DNA glycosylases.* Mutat Res, 1990. **236**(2-3): p. 161-72.

55. Ramotar, D., *The apurinic-apyrimidinic endonuclease IV family of DNA repair enzymes.* Biochem Cell Biol, 1997. **75**(4): p. 327-36.

56. Otterlei, M., et al., *Repair of chromosomal abasic sites in vivo involves at least three different repair pathways.* Embo J, 2000. **19**(20): p. 5542-51.

57. Demple, B., A. Johnson, and D. Fung, *Exonuclease III and endonuclease IV remove 3' blocks from DNA synthesis primers in H2O2-damaged Escherichia coli.* Proc Natl Acad Sci U S A, 1986. **83**(20): p. 7731-5.

58. Sung, J.S. and B. Demple, *Roles of base excision repair subpathways in correcting oxidized abasic sites in DNA.* Febs J, 2006. **273**(8): p. 1620-9.

59. Bailly, V. and W.G. Verly, *AP endonucleases and AP lyases.* Nucleic Acids Res, 1989. **17**(9): p. 3617-8.

60. Levin, J.D., A.W. Johnson, and B. Demple, *Homogeneous Escherichia coli endonuclease IV. Characterization of an enzyme that recognizes oxidative damage in DNA.* J Biol Chem, 1988. **263**(17): p. 8066-71.

61. Cunningham, R.P., et al., *Endonuclease IV (nfo) mutant of Escherichia coli.* J Bacteriol, 1986. **168**(3): p. 1120-7.

62. Ljungquist, S., *A new endonuclease from Escherichia coli acting at apurinic sites in DNA.* J Biol Chem, 1977. **252**(9): p. 2808-14.

63. Doetsch, P.W. and R.P. Cunningham, *The enzymology of apurinic/apyrimidinic endonucleases.* Mutat Res, 1990. **236**(2-3): p. 173-201.

64. Chan, E. and B. Weiss, *Endonuclease IV of Escherichia coli is induced by paraquat.* Proc Natl Acad Sci U S A, 1987. **84**(10): p. 3189-93.

65. Li, Z. and B. Demple, *SoxS, an activator of superoxide stress genes in Escherichia coli. Purification and interaction with DNA.* J Biol Chem, 1994. **269**(28): p. 18371-7.

66. Kerins, S.M., R. Collins, and T.V. McCarthy, *Characterization of an endonuclease IV 3'-5' exonuclease activity.* J Biol Chem, 2003. **278**(5): p. 3048-54.

67. Ishchenko, A.A., et al., *The 3'->5' exonuclease of Apn1 provides an alternative pathway to repair 7,8-dihydro-8-oxodeoxyguanosine in Saccharomyces cerevisiae.* Mol Cell Biol, 2005. **25**(15): p. 6380-90.

68. Hosfield, D.J., et al., *Structure of the DNA repair enzyme endonuclease IV and its DNA complex: double-nucleotide flipping at abasic sites and three-metal-ion catalysis.* Cell, 1999. **98**(3): p. 397-408.

69. Ivanov, I., J.A. Tainer, and J.A. McCammon, *Unraveling the three-metal-ion catalytic mechanism of the DNA repair enzyme endonuclease IV.* Proc Natl Acad Sci U S A, 2007. **104**(5): p. 1465-70.

70. Paspaleva, K., et al., *Crystal structure of the DNA repair enzyme ultraviolet damage endonuclease.* Structure, 2007. **15**(10): p. 1316-24.

71. Burger, R.M., *Cleavage of Nucleic Acids by Bleomycin.* Chem Rev, 1998. **98**(3): p. 1153-1170.

72. Hagensee, M.E. and R.E. Moses, *Bleomycin-treated DNA is specifically cleaved only by endonuclease IV in E. coli.* Biochim Biophys Acta, 1990. **1048**(1): p. 19-23.

73. Siwek, B., et al., *The relative importance of Escherichia coli exonuclease III and endonuclease IV for the hydrolysis of 3'-phosphoglycolate ends in polydeoxynucleotides.* Nucleic Acids Res, 1988. **16**(11): p. 5031-8.

74. Loeb, L.A., et al., *Apurinic sites as common intermediates in mutagenesis.* Basic Life Sci, 1986. **38**: p. 341-7.

75. Ramotar, D., S.C. Popoff, and B. Demple, *Complementation of DNA repair-deficient Escherichia coli by the yeast Apn1 apurinic/apyrimidinic endonuclease gene.* Mol Microbiol, 1991. **5**(1): p. 149-55.

76. Souza, L.L., et al., *Endonuclease IV and exonuclease III are involved in the repair and mutagenesis of DNA lesions induced by UVB in Escherichia coli.* Mutagenesis, 2006. **21**(2): p. 125-30.

77. Levin, J.D., R. Shapiro, and B. Demple, *Metalloenzymes in DNA repair. Escherichia coli endonuclease IV and Saccharomyces cerevisiae Apn1.* J Biol Chem, 1991. **266**(34): p. 22893-8.

78. Ramotar, D., et al., *Intracellular localization of the Apn1 DNA repair enzyme of Saccharomyces cerevisiae. Nuclear transport signals and biological role.* J Biol Chem, 1993. **268**(27): p. 20533-9.

79. Vongsamphanh, R., P.K. Fortier, and D. Ramotar, *Pir1p mediates translocation of the yeast Apn1p endonuclease into the mitochondria to maintain genomic stability.* Mol Cell Biol, 2001. **21**(5): p. 1647-55.

80. Ramotar, D., et al., *Cellular role of yeast Apn1 apurinic endonuclease/3'-diesterase: repair of oxidative and alkylation DNA damage and control of spontaneous mutation.* Mol Cell Biol, 1991. **11**(9): p. 4537-44.

81. Kunz, B.A., et al., *Specificity of the mutator caused by deletion of the yeast structural gene (APN1) for the major apurinic endonuclease.* Proc Natl Acad Sci U S A, 1994. **91**(17): p. 8165-9.

82. Bennett, R.A., *The Saccharomyces cerevisiae ETH1 gene, an inducible homolog of exonuclease III that provides resistance to DNA-damaging agents and limits spontaneous mutagenesis.* Mol Cell Biol, 1999. **19**(3): p. 1800-9.

83. Ho, R., et al., *Yeast apurinic/apyrimidinic endonuclease Apn1 protects mammalian neuronal cell line from oxidative stress.* J Neurochem, 2007. **102**(1): p. 13-24.

84. Masson, J.Y., S. Tremblay, and D. Ramotar, *The Caenorhabditis elegans gene CeAPN1 encodes a homolog of Escherichia coli and yeast apurinic/apyrimidinic endonuclease.* Gene, 1996. **179**(2): p. 291-3.

85. Shatilla, A., et al., *Characterization of Caenorhabditis elegans exonuclease-3 and evidence that a Mg2+-dependent variant exhibits a distinct mode of action on damaged DNA.* Biochemistry, 2005. **44**(38): p. 12835-48.

86. Shatilla, A., et al., *Identification of two apurinic/apyrimidinic endonucleases from Caenorhabditis elegans by cross-species complementation.* DNA Repair (Amst), 2005. **4**(6): p. 655-70.

87. Cansizoglu, A.E., et al., *Structure-based design of a pathway-specific nuclear import inhibitor.* Nat Struct Mol Biol, 2007. **14**(5): p. 452-4.

88. Jilani, A., et al., *Characterization of two independent amino acid substitutions that disrupt the DNA repair functions of the yeast Apn1.* Biochemistry, 2003. **42**(21): p. 6436-45.

89. Sambrook, J., E.F. Fritsch, and T. Maniatis, *Molecular cloning : a laboratory manual.* 2nd ed. 1989, Cold Spring Harbor, N.Y.: Cold Spring Harbor Laboratory. 3 v.

90. Rosenberg, I.M., *Protein analysis and purification : benchtop techniques.* 2nd ed. 2004, Boston: Birkhäuser. xxvi, 520.

91. Masson, J.Y. and D. Ramotar, *Normal processing of AP sites in Apn1-deficient Saccharomyces cerevisiae is restored by Escherichia coli genes expressing either exonuclease III or endonuclease III.* Mol Microbiol, 1997. **24**(4): p. 711-21.

92. Shatilla, A. and D. Ramotar, *Embryonic extracts derived from the nematode Caenorhabditis elegans remove uracil from DNA by the sequential action of uracil-DNA glycosylase and AP (apurinic/apyrimidinic) endonuclease.* Biochem J, 2002. **365**(Pt 2): p. 547-53.

93. Yang, X., et al., *Characterization of amino acid substitutions that severely alter the DNA repair functions of Escherichia coli endonuclease IV.* Biochemistry, 1999. **38**(12): p. 3615-23.

94. Kopp, J. and T. Schwede, *The SWISS-MODEL Repository of annotated three-dimensional protein structure homology models.* Nucleic Acids Res, 2004. **32**(Database issue): p. D230-4.

95. Schwede, T., et al., *SWISS-MODEL: An automated protein homology-modeling server.* Nucleic Acids Res, 2003. **31**(13): p. 3381-5.

96. Guex, N., A. Diemand, and M.C. Peitsch, *Protein modelling for all.* Trends Biochem Sci, 1999. **24**(9): p. 364-7.

97. Kinoshita, K. and H. Nakamura, *eF-site and PDBjViewer: database and viewer for protein functional sites.* Bioinformatics, 2004. **20**(8): p. 1329-30.

98. Johnson, A.W. and B. Demple, *Yeast DNA diesterase for 3'-fragments of deoxyribose: purification and physical properties of a repair enzyme for oxidative DNA damage.* J Biol Chem, 1988. **263**(34): p. 18009-16.

99. Wang, N.P., et al., *Expression of the human retinoblastoma gene product pp110RB in insect cells using the baculovirus system.* Cell Growth Differ, 1990. **1**(9): p. 429-37.

100. Ruzindana-Umunyana, A., S. Sircar, and J.M. Weber, *The effect of mutant peptide cofactors on adenovirus protease activity and virus infection.* Virology, 2000. **270**(1): p. 173-9.

101. Abdel-Ghany, M., C. Riegler, and E. Racker, *A placental polypeptide activator of a membranous protein kinase and its relation to histone 1.* Proc Natl Acad Sci U S A, 1984. **81**(23): p. 7388-91.

102. Finnie, N.J., et al., *DNA-dependent protein kinase activity is absent in xrs-6 cells: implications for site-specific recombination and DNA double-strand break repair.* Proc Natl Acad Sci U S A, 1995. **92**(1): p. 320-4.

103. Al-Tassan, N., et al., *Inherited variants of MYH associated with somatic G:C-->T:A mutations in colorectal tumors.* Nat Genet, 2002. **30**(2): p. 227-32.

104. Chmiel, N.H., A.L. Livingston, and S.S. David, *Insight into the functional consequences of inherited variants of the hMYH adenine glycosylase associated with colorectal cancer: complementation assays with hMYH variants and pre-steady-state kinetics of the corresponding mutated E.coli enzymes.* J Mol Biol, 2003. **327**(2): p. 431-43.

105. Shinmura, K. and J. Yokota, *The OGG1 gene encodes a repair enzyme for oxidatively damaged DNA and is involved in human carcinogenesis.* Antioxid Redox Signal, 2001. **3**(4): p. 597-609.

106. Boiteux, S. and J.P. Radicella, *The human OGG1 gene: structure, functions, and its implication in the process of carcinogenesis.* Arch Biochem Biophys, 2000. **377**(1): p. 1-8.

107. Starcevic, D., S. Dalal, and J.B. Sweasy, *Is there a link between DNA polymerase beta and cancer?* Cell Cycle, 2004. **3**(8): p. 998-1001.

108. Hadi, M.Z., et al., *Functional characterization of Ape1 variants identified in the human population.* Nucleic Acids Res, 2000. **28**(20): p. 3871-9.

109. Wilson, D.M., 3rd and V.A. Bohr, *The mechanics of base excision repair, and its relationship to aging and disease.* DNA Repair (Amst), 2007. **6**(4): p. 544-59.

110. Huamani, J., et al., *Spontaneous mutagenesis is enhanced in Apex heterozygous mice.* Mol Cell Biol, 2004. **24**(18): p. 8145-53.

111. Meira, L.B., et al., *Heterozygosity for the mouse Apex gene results in phenotypes associated with oxidative stress.* Cancer Res, 2001. **61**(14): p. 5552-7.

112. Haltiwanger, B.M., N.O. Karpinich, and T.F. Taraschi, *Characterization of class II apurinic/apyrimidinic endonuclease activities in the human malaria parasite, Plasmodium falciparum.* Biochem J, 2000. **345 Pt 1**: p. 85-9.

113. Guven, K., et al., *The toxicity of dithiocarbamate fungicides to soil nematodes, assessed using a stress-inducible transgenic strain of Caenorhabditis elegans.* J Biochem Mol Toxicol, 1999. **13**(6): p. 324-33.