

https://dspace.org/

The identification and characterisation of a novel Apoptotic Gene,Snama, in drosophila melanogaster

Mather, Arshad Saleh

2006-11-15T13:01:51Z
http://hdl.handle.net/10539/1745

Downloaded from a DSpace repository. https://dspace.org/

THE IDENTIFICATION AND CHARACTERISATION OF A NOVEL APOPTOTIC GENE, *SNAMA*, IN *Drosophila melanogaster*

Arshad Saleh Mather

A thesis submitted to the Faculty of Science, University of the Witwatersrand, Johannesburg, in fulfilment of the requirements for the degree of Doctor of Philosophy.

Johannesburg, 2005

DECLARATION

I declare that this thesis is my own, unaided work. It is being submitted for the degree of Doctor of Philosophy in the University of the Witwatersrand, Johannesburg. It has not been submitted before for any degree or examination at any other University.

_____

ARSHAD SALEH MATHER

_____ Day of _____200_____

ABSTRACT

SNAMA is the *Drosophila melanogaster* homologue of a group of proteins that are known to bind p53 and the retinoblastoma protein (Rb). This multi domain protein consists of a conserved N-terminal domain called Domain With No Name (DWNN), a zinc finger, a cysteine rich RING finger-like domain, a probable p53 binding region, and a glutamic acid-rich and lysine-rich region. These associated domains indicate that SNAMA plays an important regulatory role in the cell and may function in RNA processing and in apoptosis.

The DWNN domain was first identified in Cytotoxic T-cell resistant Chinese hamster ovary (CHO) cells using promoter trap mutagenesis to screen for genes involved in apoptosis. Subsequently, this domain was identified in other eukaryotic organisms including animals and plants.

The *SNAMA* transcriptional unit consists of 9 exons and 8 introns that code for a 1231 amino acid protein with the 76 residue N-terminal DWNN domain. The DWNN domain has a 23.5% sequence identity to the ubiquitin protein and a predicted folded structure similar to ubiquitin. Western blots identified multiple bands indicative of ubiquitin tagged proteins. Taken together this suggests a role in the ubiquitin pathway either as an ubiquitin domain protein or the DWNN domain of SNAMA tagging other proteins. The cysteine rich RING finger-like domain has a histidine to serine substitution at the fourth position of the putative RING finger and represents a

distinct class of RING finger-like proteins that could have ubiquitin ligase activity. Northern blot analysis identified a single 4.6 kbp transcript expressed abundantly throughout development early in embryogenesis but reduced in older embryos and in adult male and females. SNAMA probably interacts with Dmp53 as a suppressor of apoptosis or a negative regulator of an activator of apoptosis. It is a vital gene required for development, as the mutant P-element insertion line in which the P-element is inserted in the first intron of SNAMA is lethal when homozygous. Acridine orange staining of these mutant flies showed a direct correlation between the presence of SNAMA and apoptosis. An increase in the levels of apoptosis occurred in embryos with relatively low levels of *SNAMA* expression. The mode of this action is either direct, or via other proteins that are involved in the apoptotic pathway.

This thesis is dedicated to my daughter

Maseeha Mather.

"Every day you may make progress. Every step may be fruitful. Yet there will stretch out before you an ever-lengthening, ever-ascending, ever-improving path. You know you will never get to the end of the journey. But this, so far from discouraging, only adds to the joy and glory of the climb." **Sir Winston Churchill**

ACKNOWLEDGEMENTS

My supervisor, Dr Monde Ntwasa for all his guidance, encouragement, and enthusiasm throughout my studies.

My wife, Safiyya, for her constant encouragement and understanding over the many years. She has been patient and supportive of all my decisions. I would also like to thank my parents for their support and encouragement.

Dr Yasien Sayed, Christopher Arnot, Rodney Hull, Mpho Rakgotho and Shune Oliver for all their assistance.

The South African National Research Foundation for financial support.

TABLE OF CONTENTS

DECLARATION……..… ii

ABSTRACT……. iii

DEDICATION……..… v

ACKNOWLEDGEMENTS……...…… vi

ABBREVIATIONS…….. xii

LIST OF FIGURES……..… xiii

LIST OF TABLES…… xvi

CHAPTER 1: INTRODUCTION…… 1

1.1 Genes involved in programmed cell death…… 3

1.2 *Drosophila melanogaster* cell death genes…… 4

1.3 The role of caspases in apoptosis…… 5

1.4 Regulators of programmed cell death…… 7

1.5 Signalling pathways in programmed cell death…… 10

1.5.1 p53 signalling and regulation by Mdm2……...…… 11

1.5.2 The domain organisation of p53 and/or pRB interacting proteins…… 16

1.6 The ubiquitin system…… 23

1.7 Ubiquitin enzymes…… 27

1.8 The RING finger E3s …….. 29

1.9 The process of protein degradation…… 33

1.10 Deubiquitinating enzymes (DUBs)……...…… 35

1.11 Ubiquitin-like proteins…… 37

1.11.1 Ubiquitin-like proteins (UBLs)……. 37

1.11.2 Ubiquitin domain proteins (UDPs)…… 42

1.12 Aims…… 45

CHAPTER 2: MATERIALS AND METHODS..... 47

2.1 Media..... 48
2.2 Buffers and solutions..... 49
2.3 Chemicals and kits..... 50
2.4 Bacterial strains..... 51
2.5 Oligonucleotides..... 52
2.6 Genetic techniques..... 53
2.6.1 Fly stocks and genetic manipulation..... 53
2.6.2 Generation of the P-element excision line..... 54
2.7 General molecular biology methods..... 54
2.7.1 Small scale preparation of plasmid DNA..... 54
2.7.2 Large scale preparation of plasmid DNA..... 55
2.7.3 Digestion of DNA fragments with restriction enzymes..... 56
2.7.4 Purification of restriction enzyme-digested DNA fragments..... 56
2.7.5 Preparation of competent *E. coli* cells..... 56
2.7.6 Transformation of *E. coli* cells with plasmid DNA..... 57
2.7.7 Isolation of total RNA from *D. melanogaster* embryos..... 57
2.7.8 Preparation of whole embryo extracts..... 59
2.7.9 Preparation of genomic DNA..... 59
2.8 Biochemical methods..... 59
2.8.1 *In situ* hybridization to *D. melanogaster* embryos..... 59
2.8.2 Northern hybridization..... 65
2.8.3 The polymerase chain reaction (PCR)..... 67
2.8.4 Reverse transcription PCR (RT-PCR)..... 67
2.8.5 Western blotting..... 68
2.8.6 Acridine orange staining..... 70
2.8.7 Cuticle preparation..... 70
2.9 Bioinformatics..... 71

CHAPTER 3: IDENTIFICATION AND SEQUENCE ANALYSIS OF *SNAMA*, THE DWNN-CONTAINING HOMOLOGUE, IN *Drosophila melanogaster* 73

3.1 Introduction 74
3.2 Sequence analysis of the *SNAMA* gene 79
3.2.1 The 5′ untranslated region 86
3.2.2 The 3′ untranslated region 87
3.2.3 Splice junctions 88
3.3 The domain organisation of SNAMA 91
3.3.1 The DWNN domain of the SNAMA protein 96
3.3.2 The cysteine rich RING finger-like domain 99
3.3.3 The ubiquitin-like fold of the DWNN domain of SNAMA 102

CHAPTER 4: TEMPORAL AND SPACIAL DISTRIBUTION OF SNAMA IN *Drosophila melanogaster* 109

4.1 Introduction 110
4.2 *SNAMA* is a developmentally regulated gene 110
4.3 *SNAMA* cDNA sequence 115
4.4 SNAMA is expressed throughout development 118
4.5 SNAMA may be localised in the nucleus of *D. melanogaster* embryos 123
4.6 SNAMA interacts with Dmp53 127

CHAPTER 5: THE BIOLOGICAL FUNCTION OF SNAMA IN *Drosophila melanogaster* 132

5.1 Introduction 133
5.2 *SNAMA* is an essential gene required for development 133
5.2.1 P-element mutations in *Drosophila melanogaster* 133
5.2.2 Identification of the position of the P-element in *l(2)rQ313* flies 135
5.2.3 Mobilisation of the P-element from *l(2)rQ313* flies 140
5.2.4 Cuticle analysis of *l(2)rQ313* embryos 142
5.3 *SNAMA* is characteristic of genes involved in apoptosis 145

CHAPTER 6: GENERAL DISCUSSION AND CONCLUSION 148

6.1 The domain organisation of SNAMA protein and its functional implications 149
6.2 The sequence analysis of *SNAMA* 156
6.3 The cytological position of *SNAMA* 162
6.3.1 Genes occupying position 60B 163
6.3.1.1 Muscle LIM protein at 60A (Mlp60A) 163
6.3.1.2 Genghis Khan (Gek) 164
6.3.1.3 Enoki mushroom (Enok) 165
6.3.1.4 Tamo 165
6.3.1.5. Zinc finger RP-8 (Zfrp8) 166
6.4 The biological function of SNAMA 167
6.4.1 The DWNN domain of SNAMA has a ubiquitin-like fold 167
6.4.2 SNAMA is involved in apoptosis 172
6.5 Conclusions 175

CHAPTER 7: REFERENCES..... 176
APPENDIX..... 199

ABBREVIATIONS

APC	Anaphase promoting complex
bp	base pair
CHO	Chinese hamster ovary
CTL	cytotoxic T lymphocytes
Cys	Cysteine
DWNN	Domain with no name
ER	Endoplasmic reticulum
EST	expressed sequence tag
Gly	Glycine
His	Histidine
hr	hours
kbp	kilo base pair
kDa	kilo-dalton
LB	Luria-Bertani mediam
Lys	Lysine
MDa	mega-dalton
MHC	major histocompatibility complex
OD	optical density
ORF	open reading frame
PAGE	polyacrylamide gel electrophoresis
PCD	programmed cell death
RT-PCR	Reverse transcriptase polymerase chain reaction
UPEs	upstream promoter elements

LIST OF FIGURES

Figure 1: Schematic representation of p53 regulation by Mdm2 ubiquitin ligase activity…… 13

Figure 2: Schematic diagram of Mdm2 indicating domains that are important for the regulation of p53 function…… 15

Figure 3: Schematic diagram showing the domain organization of SNAMA orthologous proteins…….…. 19

Figure 4: The three-dimensional structure of ubiquitin…… 26

Figure 5: Schematic representation of the ligation scheme employed by the RING finger showing the unique 'cross brace'…….…. 31

Figure 6: A sequence comparison of ubiquitin to ubiquitin-like proteins……..…… 39

Figure 7: Amino acid sequence and alignment of the conserved DWNN domains with homologous DWNN domains from a variety of eukaryotic organisms……. 76

Figure 8: The sequence of the 4.595 kbp *SNAMA* gene…….…. 85

Figure 9: The domain arrangement of SNAMA homologous proteins……. 93

Figure 10: The amino acid sequence of *Drosophila melanogaster* SNAMA protein showing conserved domains and motifs…….. 94

Figure 11: Schematic representation of the "cross-brace" structure adopted by the unique RING finger of SNAMA…….…… 101

Figure 12: The alignment of SNAMA with ubiquitin and ubiquitin-like sequences, Nedd8 and Rub1, generated using FUGUE…… 104

Figure 13: Secondary structure prediction of the DWNN domain of SNAMA using PSIPRED…….. 105

Figure 14: Secondary structure prediction of ubiquitin and SNAMA using the Network protein sequence analysis tool…… 106

Figure 15: The three-dimensional structure of ubiquitin compared to the DWNN domain…….. 108

Figure 16: Reverse transcriptase PCR on various developmental stages of *D. melanogaster embryonic* RNA……. 112

Figure 17: Sequence alignment of the two RT-PCR products SNAMA-1 and SNAMA-2 with genomic DNA sequence……..…… 114

Figure 18: pOT2 vector with the *SNAMA* insert…….…… 117

Figure 19: Developmental Northern blot of *D. melanogaster* RNA from various embryonic developmental stages and adult flies……. 119

Figure 20: RT-PCR on various developmental stages of *D. melanogaster* RNA…….. 122

Figure 21: *In situ* staining of developmental stages of *D. melanogaster* embryos to detect *SNAMA* transcription…….. 126

Figure 22: Western blot analysis of proteins from embryonic, larval and adult stages of wild-type *D. melanogaster* development……. 130

Figure 23: Western blot analyses of proteins from *D. melanogaster* larvae that have been fed with camptothecin…… 131

Figure 24: Schematic representation of the primer positions on the genomic DNA from *l(2)rQ3131* flies…… 136

Figure 25: PCR amplification to detect P-element insertion in *l(2)rQ313* and *l(2)rQ313/GFP*……. 139

Figure 26: PCR of genomic DNA from excision cross flies to test for excisions of the P-element…….. 141

Figure 27: Cuticle pattern of *l(2)rQ313* embryos…… 144

Figure 28: Acridine orange staining of embryos viewed using confocal microscopy…… 147

Figure 29: Schematic representation of the domain arrangement of SNAMA protein identified with the SMART programme..... 151

Figure 30: A schematic representation of the genomic organisation of the *SNAMA* gene..... 157

Figure 31: Phylogenetic analysis of the ubiquitin-like domain of the DWNN family of proteins and comparison of the ubiquitin sequence to DWNN..... 168

LIST OF TABLES

Table 1: Media used in this study…….…. 48

Table 2: General molecular biology buffers and solutions…… 49

Table 3: Bacterial strains used in this study…….…. 51

Table 4: Oligonucleotides used for PCR and RT-PCR……..… 52

Table 5: Solutions used for *in situ* hybridisation study……. 60

Table 6: Web addresses for the various bioinformatics tools used…… 72

Table 7: A species comparison of the DWNN sequences……. 78

Table 8: *SNAMA* intron and exon splice junctions…… 90

Table 9: Introns from the *SNAMA* gene…… 90

Table 10: Sequence relationships among ubiquitin-like proteins…… 98

Table 11: Presence of the di-glycine residues at the C-terminal end of the DWNN domain in SNAMA homologous proteins…… 170

CHAPTER 1

INTRODUCTION

1.1 Genes involved in programmed cell death

1.2 *Drosophila melanogaster* cell death genes

1.3 The role of caspases in apoptosis

1.4 Regulators of programmed cell death

1.5 Signalling pathways in programmed cell death

1.6 The ubiquitin system

1.7 Ubiquitin enzymes

1.8 The RING finger E3's

1.9 The process of protein degradation

1.10 Deubiquitinating enzymes (DUBs)

1.11 Ubiquitin-like proteins

1.12 Aims

1. **INTRODUCTION**

Throughout the last century, *Drosophila melanogaster* has been the workhorse for genetic studies in eukaryotes. Due to its elaborate genetics and the availability of many techniques that facilitate genetic manipulations, the fruit fly is one of the preferred models for studying developmental and molecular biology. Two processes that are important for the development and immune protection of organisms are apoptosis and the ubiquitin-proteasome pathways. These processes play major roles in a variety of basic pathways during cell life and death as well as in health and disease. Defects in these processes have been implicated in the pathogenesis of many diseases, cancer, allergy and autoimmunity. Characterisation of novel genes in the apoptosis and ubiquitin protein degradation pathways identifies potential drug targets to develop therapeutic agents for the treatment of diseased conditions.

The focus of this thesis is to investigate the functional role of a previously unidentified *D. melanogaster* protein named SNAMA (Xhosa word meaning "sticky"). SNAMA is the *D. melanogaster* homologue of a group of proteins with a highly conserved N-terminal domain named DWNN (domain with no name). The novel DWNN domain has a three-dimensional structure similar to that of ubiquitin and it is likely that proteins bearing this domain play a role similar to ubiquitin. The DWNN proteins were identified in a mutational study on Chinese hamster ovary cells (CHO). Genes were disrupted in CHO cells and screened for resistance to apoptosis

causing agents. Based on the initial screening experiments that identified DWNN-domain containing proteins and the ubiquitin-like fold of the DWNN domain it is likely that this group of proteins plays a role in apoptosis. This thesis investigates the role of SNAMA in apoptosis, and whether SNAMA acts in an ubiquitin-like manner. This chapter presents a review of the components of the apoptotic pathway and the ubiquitin proteolytic protein degradation pathway. Attention will be focused on the apoptotic-signalling pathway involving p53 and the ubiquitin-like proteins and their roles.

1.1 GENES INVOLVED IN PROGRAMMED CELL DEATH

Programmed cell death (PCD) is a normal physiological process of cellular suicide essential for maintaining homeostasis. The process is essential for the removal of unwanted cells during tissue modelling in development and to maintain a stable balance between proliferation of normal cells and non-inflammatory death of damaged cells. Two types of programmed cell death have been reported: apoptosis and autophagy. The morphological features of apoptosis are cell shrinkage, accompanied by membrane blebbing resulting in separation of the cell into a cluster of membrane – bound bodies, nuclear and cytoplasmic condensation and DNA fragmentation. Autophagy involves the destruction of entire tissues and the presence of autophagic vacuoles.

The genes involved in apoptosis were first identified in *Caenorhabditis elegans* (Fraser, 1999). These genes were Ced-3 (a cysteine protease/caspase), Ced-4 (involved in caspase activation), Ced-9 (an inhibitor of apoptosis) and Egl-1 (a BH3 domain-only protein which induced apoptosis by binding Ced-9) (Meier and Evan, 1998). Structural or functional homologues of these genes have been found in mammals and *D. melanogaster* (Bangs and White, 2000) indicating that the core components of PCD are highly conserved. Higher organisms show a greater degree of complexity in the regulatory mechanisms of programmed cell death. *D. melanogaster* provides a system of intermediate complexity between *C. elegans* and mammals in which to study PCD (Meier *et al.*, 2000).

1.2 *Drosophila melanogaster* CELL DEATH GENES

The first PCD genes identified in *D. melanogaster* were the death activator genes of the intrinsic pathway *reaper* (White *et al.*, 1994), *hid* (head involution defect/Wrinkled) (Grether *et al.*, 1995) and *grim* which cooperate to mediate PCD in the embryo (Chen *et al.*, 1996b). These three genes are found at position 75C on chromosome three (Haining *et al.*, 1999; Igaki *et al.*, 2000) and all share a similar 14-amino-acid sequence at their N-terminus called the RHG motif (Wing *et al.*, 1998). When the combined activity of Rpr, Hid and Grim in a cell exceeds a set threshold, caspases are activated and the cell undergoes apoptosis (Nordstrom and Abrams, 2000). The human functional homologue of Rpr, Hid or Grim is the apoptosis inducer Smac/Diablo (DeLaurenzi and Melino, 2000; Du *et al.*, 2000).

Reaper (Rpr) encodes a 65 residue cytoplasmic protein with no known homologues. *Rpr* promotes Bcl2-suppressible apoptotic activity in *Xenopus* egg extracts (Evans *et al.*, 1997). The *Rpr*-induced activity is mediated by Scythe, the *Xenopus* protein containing a ubiquitin-like domain near the N-terminus (Thress *et al.*, 1998). This suggests that a *Scythe*-like *Drosophila* orthologue may be an important transducer of the *Rpr* signal (Abrams, 1999). A *D. melanogaster Scythe* homologue has not yet been identified. Two pathways have been identified that activate *Rpr* expression; the steroid hormone ecdysone (Jiang *et al.*, 2000) and the *D. melanogaster* homologue of p53, Dmp53, which acts through a p53-response element within the upstream regulatory region of *Rpr* (Brodsky *et al.*, 2000; Ollmann *et al.*, 2000).

The predicted *Hid* gene product is a 410 amino acid protein rich in serine and proline. It is found both in cells undergoing apoptosis and in those that do not ultimately die. Expression of *D. melanogaster Hid* in mammalian cells induces apoptosis (Haining *et al.*, 1999). The activity of *Hid* is controlled by growth factors acting via the Ras/Raf/MAPK pathway. Activation of these pathways promotes the down regulation of *Hid* expression resulting in suppression of Hid-induced killing (Kurada and White, 1998; Bergmann *et al.*, 1998).

The third gene identified, *Grim*, maps between *Rpr* and *Hid*. The deduced open reading frame of *Grim* encodes a protein of 138 amino acids. *Grim* expression is developmentally controlled and regulated by mechanisms similar to *Rpr* expression

(Chen *et al.*, 1996b). *Grim* initiates apoptosis, even when expression of *Rpr* is not able to, indicating distinct regulators and/or effectors of this gene.

1.3 THE ROLE OF CASPASES IN APOPTOSIS

The caspases exist as large inactive precursors (called procaspases) of between 20 to 30 kDa that contain three domains: an NH2–terminal domain, a large subunit (~20 kDa), and a small subunit (~10 kDa) (Thornberry and Lazebnik, 1998). Activation involves cleavage at aspartic acid residues by either another caspase, or by autocatalysis (Nicholson, 1999), followed by the association of the large and small subunits to form a heterodimer. The binding of cofactors or the removal of inhibitors from the procaspase triggers this cleavage. In apoptosis, caspases are activated in an amplifying proteolytic cascade, cleaving one another in sequence resulting in an exponential rate of activation that ensures the protease can be functional quickly (Raff, 1998). There are seven caspases in *D. melanogaster* named Dcp-1, Dredd/Dcp-2, Drice, Dronc, Decay, Strica/Dream and Damm/Daydream. These have been reviewed in (Kumar and Doumanis, 2000).

Caspases have been divided into two classes. The class I caspases (initiator caspases) contain long prodomains comprised of two subunits and the class II caspases (effector caspases) either lack this prodomain or have very short prodomains (Thornberry and Lazebnik, 1998; Nicholson, 1999; Kumar, 1999; Kumar and Colussi, 1999). The long prodomain of class I caspases contains two types of domains: the caspase recruitment

domain (CARD) or the death effector domain (DED) (Aravind *et al.*, 1999). In mammals, the CARD of caspase-9 is required for the recruitment of procaspase-9 by Apaf-1 via a N-terminal CARD domain on Apaf-1 (Li *et al.*, 1997; Zou *et al.*, 1999). One of the two DEDs of caspase-8 interacts with the DED in the adaptor protein FADD (Fas-associated death domain) (Aravind *et al.*, 1999). The recruitment of the procaspases with their adaptors is essential for promoting autocatalytic activation.

D. melanogaster Dredd, Dronc and Strica are class I caspases with long prodomains. Dredd contains two DEDs in the prodomain region and Dronc has a CARD (Chen *et al.*, 1998; Inohara *et al.*, 1997; Dorstyn *et al.*, 1999). Strica has a unique Serine/Threonine-rich domain at its N-terminus and lacks any DED or CARD. The function of the Serine/Threonine domain is not yet known (Doumanis *et al.*, 2001). Dcp-1, Drice, Decay and Damm are similar to the mammalian class II effector caspases lacking the long prodomains (Kumar, 1999). The caspases are required for Rpr, Grim and Hid cell killing.

1.4 REGULATORS OF PROGRAMMED CELL DEATH

In mammals, apoptosis is regulated by the adaptor protein Apaf-1 which is required for caspase-9 activation (Zou *et al.*, 1999). The activation of caspase-9 requires the release of cytochrome *c* from the mitochondria (Green and Reed, 1998). Cytochrome *c* binds to a specific region at the C-terminal end of Apaf-1 called the WD-40 repeats. The WD-40 repeats inhibit Apaf-1 function and cytochrome *c* binding thus initiates

the oligomerisation of Apaf-1 molecules (Li *et al.*, 1997; Zou *et al.*, 1999). Apaf-1 then recruits procaspase-9 resulting in the proximity-induced autocatalytic activation of the caspase. The *D. melanogaster* homologue of Apaf-1 is Dark/Dapaf-1/Hac-1 (Kanuka *et al.*, 1999; Rodriguez *et al.*, 1999; Zhou *et al.*, 1999a). This protein is referred to as Apaf-1-related killer (Ark) in the *D. melanogaster* online database (http://flybase.bio.indiana.edu/genes/). Dark binds to the initiator caspases, Dredd and Dronc, through their N-terminal domains. These caspases are cleaved to the active form to activate apoptosis. There are conflicting reports about the binding ability of Dark to the executioner caspase Drice (Kanuka *et al.*, 1999; Rodriguez *et al.*, 1999). Dark has a WD-40 domain region at the C-terminal end which binds cytochrome *c*, but it is not known whether Dark requires cytochrome *c* for its oligomerisation. The specific role of Dark in Rpr, Hid and Grim mediated PCD is not clear, but there is evidence that *Dark* expression affects Rpr, Hid and Grim mediated PCD.

Another class of regulators of PCD in mammals is the Bcl-2 family of proteins (Raff, 1998). Fifteen Bcl-2 family members have been identified. Members of this family include Bcl-2, Bcl-xL and Bcl-w that suppress apoptosis, while Bax, Bad, Bid and Bok promote cell death (Adams and Cory, 1998). A BH4 domain is found exclusively on the anti-apoptotic Bcl-2 family members. The Bcl-2 gene product regulates apoptosis in several ways (Merry and Korsmeyer, 1997). They either bind to the Ced-4-like adaptor protein blocking their ability to activate procaspase-9, or they bind to the apoptosis promoters inhibiting their function. Many proteins of the Bcl-2 family have a hydrophobic tail that enables them to bind to the outer membrane

of the mitochondria, the endoplasmic reticulum and the nucleus, and are thought to function mainly on these organelles (Raff, 1998). Debcl/dBorg-1/dRob-1 and Buffy/dBorg-2 are *D. melanogaster* homologues of Bcl-2 (Igaki *et al.*, 2000; Chen and Abrams, 2000). Both these proteins have the BH1, BH2, BH3 and C-terminal transmembrane domains of the Bcl-2 family of proteins, but lack the BH4 domain present on anti-apoptotic Bcl-2 proteins. Debcl and Buffy are related to the pro-apoptotic Bok (Richardson and Kumar, 2002). Deblc expression promoted increased apoptosis indicating a pro-apoptotic function in different embryonic developmental stages (Colussi *et al.*, 2000). In contrast, despite the absence of the BH4 domain, Buffy was found to function as an anti-apoptotic factor (Quinn *et al.*, 2003).

The inhibitors of apoptosis proteins (IAPs) regulate apoptosis by binding caspases thereby inhibiting their function (Goyal, 2001). All IAPs have one or more BIR motifs (baculovirus inhibitor of apoptosis type repeats) essential for their anti-apoptotic activity, while other IAPs have been found to have in addition a C-terminal RING finger domain (Clem and Duckett, 1997). Two *D. melanogaster* IAP homologues have been reported, Diap1/thread and Diap2 (Hay, 2000). Diap1 functions to inhibit the caspases Dcp-1, Drice and Dronc by blocking the activation of these proteins. Rpr, Hid and Grim promote apoptosis by disrupting the Diap1-caspase interaction (Goyal *et al.*, 2000; Wang *et al.*, 1999). Binding of Diaps to Rpr, Hid and Grim occurs via a homologous N-terminal region. This region of homology is found in mammalian Smac/Diablo which functions similar to Rpr, Hid or Grim. However, it has been shown that the IAP-binding domain of Smac/Diablo is not required for pro-

apoptotic activity (Roberts *et al.*, 2001). This data suggest that Smac/Diablo-induced cell death is independent of IAPs. It has also been suggested that Grim-induced cell death is independent of Diap and caspases (Wing *et al.*, 2001). Functional mapping studies suggest that IAPs function by simultaneously preventing binding of Rpr, Hid and Grim to their downstream effectors while also blocking caspase function (Vucic *et al.*, 1998; Vucic *et al.*, 1997; Chen *et al.*, 1996a).

1.5 SIGNALLING PATHWAYS IN PROGRAMMED CELL DEATH

One of the PCD signalling pathways in mammals involves a protein called Fadd. The *D. melanogaster* homologue of Fadd has been named *Drosophila* Fadd (dFadd). dFadd contains a death domain similar to the mammalian Fadd death domain. It also has the death-inducing domain found in the caspase Dredd. *In vitro* studies have indicated the binding of dFadd to Dredd through the death-inducing domain, which activates the death-inducing activity of Dredd by proteolytic processing of Dredd (Hu and Yang, 2000). The role of dFadd may be to activate Dredd (Richardson and Kumar, 2002), and since Dredd plays a role in the innate immune response, dFadd may also function in the innate immune response pathway. This assumption, however, needs to be confirmed experimentally. No death receptors (tumour necrosis factor receptor [TNFR] family) have been identified in *D. melanogaster* and neither has a genomic sequence been identified that has both a death domain and a death effector domain characteristic of mammalian death receptors (Vernooy *et al.*, 2000; Aravind *et al.*, 2001).

Grim and *Rpr* are expressed specifically in cells that undergo PCD while *Hid* has a general pattern of expression (White *et al.*, 1994; Chen *et al.*, 1996b; Grether *et al.*, 1995). *Rpr* expression is regulated by the steroid hormone ecdysone and by developmental signals (Jiang *et al.*, 2000). *Hid* is negatively regulated by the epidermal growth factor receptor (EGFR)-Ras-Raf-MAPK signalling pathway (Kurada and White, 1998; Bergmann *et al.*, 1998). The MAPK phosphorylation sites in Hid are essential for cell survival promoted by the EGFR pathway and *Hid* mRNA expression is regulated by the Pointed and Yan transcription factors which act downstream of the EGFR-Ras signalling pathway.

1.5.1 p53 signalling and regulation by Mdm2

The mammalian tumour suppressor gene, p53, encodes a 53 kDa protein that mediates PCD or cell cycle arrest in response to stress signals such as DNA damage (Levine, 1997). The p53 protein is a DNA-binding transcription factor that binds to enhancer/promoter elements of downstream target genes and thereby regulating transcription. Activation of p53 in response to cellular stress or DNA damage involves phosphorylation of the p53 protein. Once activated, p53 moves into the nucleus, and binds to specific enhancers thus regulating transcription of adjacent genes. Under normal cellular conditions, p53 is maintained at very low levels and these levels are only increased in response to stress signals. The increased levels of p53 either induce cell cycle arrest in the G1 phase or G2/M phase, or initiate apoptosis. The apoptotic pathway is initiated by activation of pro-apoptotic genes

Bax, *Fas*, and *IGF-BP*. *Bax* activates Apaf-1 and caspase 9.

Dmp53 (*Drosophila melanogaster* p53) is the homologue of p53 (Ollmann *et al.*, 2000). Dmp53 induces PCD in response to DNA damage but does not play a role in cell cycle arrest (Lee *et al.*, 2003; Nordstrom and Abrams, 2000). The 42 kDa Dmp53 protein functions by inducing expression of *Rpr* through the radiation-inducible element (*p53RE*) located ~5 kbp upstream of the *Rpr* translational start codon (Brodsky *et al.*, 2000). Dmp53 also activates other pro-apoptotic genes such as components of the Bcl-2/Apaf-1 pathway (Steller, 2000).

Mdm2, $p14^{ARF}$ and E2F-1 maintain the regulation of low cellular p53 levels (Jin and Levine, 2001; Momand *et al.*, 2000). Mdm2 and p53 are involved in a negative feedback loop with each other. p53 positively regulates Mdm2 by activating its transcription and Mdm2 negatively regulates p53 by promoting its ubiquitination and degradation of (Figure 1). The N-terminal region of Mdm2 binds p53 at the N-terminal end and acts as an ubiquitin ligase by catalyzing the transfer of ubiquitin to p53 via its RING finger (see Figure 2A) (Honda *et al.*, 1997). The amino acid residues on p53 that make contact with Mdm2 are: Phe19, Leu22, Trp23 and Pro27 (Figure 2B) (Chen *et al.*, 1993).

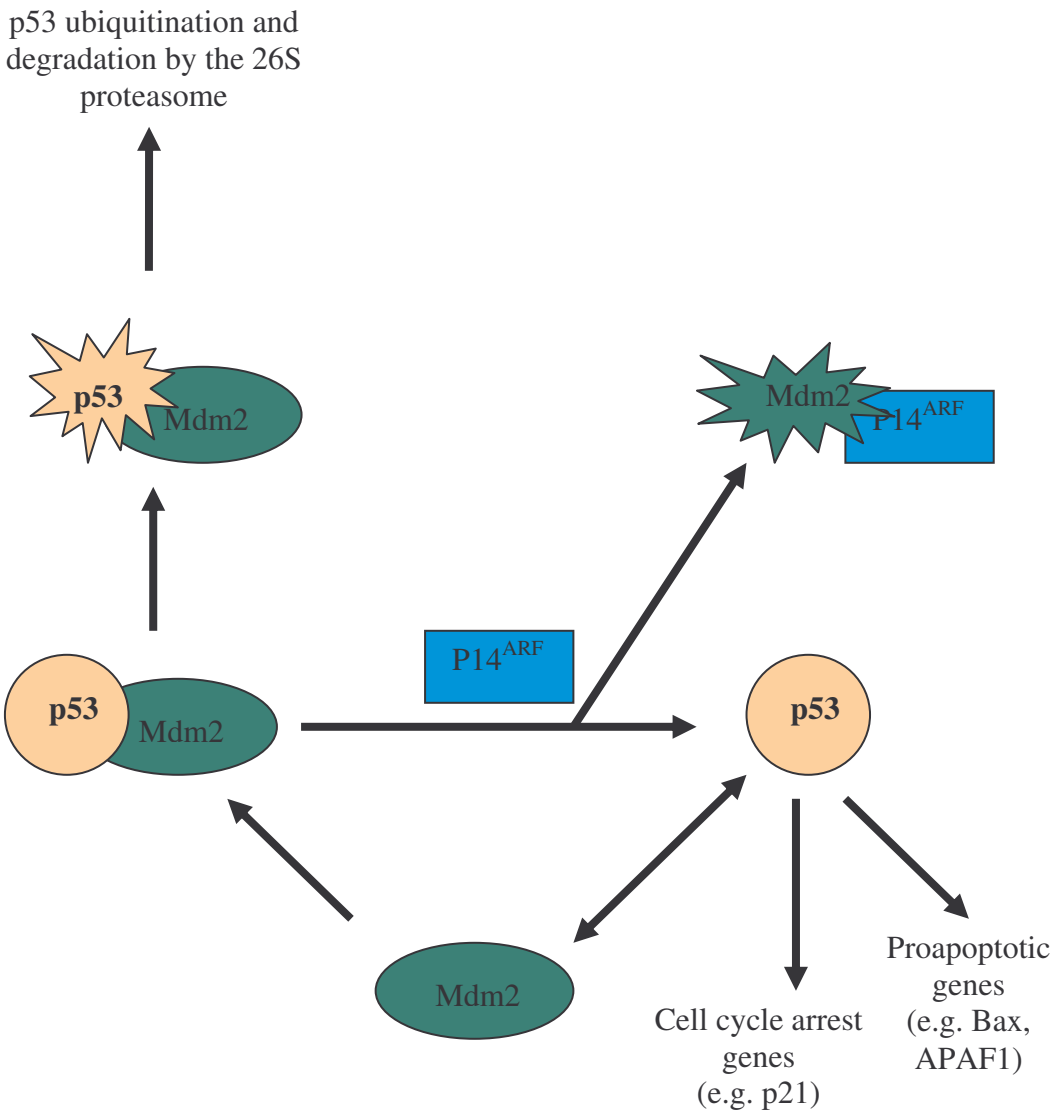
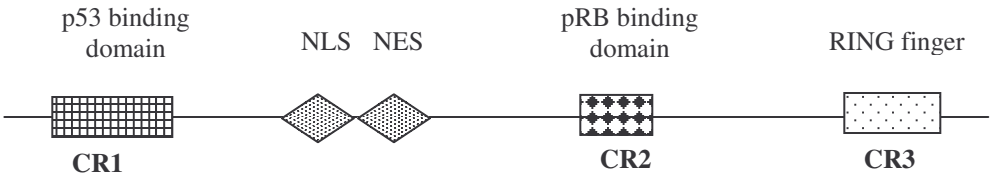


Figure 1: Schematic representation of p53 regulation by Mdm2 ubiquitin ligase activity. p53 functions to initiate apoptosis or cell cycle arrest. Binding of p53 to Mdm2 inhibits p53 activity and ubiquitinates p53. The p53 is targeted for degradation by the 26S proteasome. Mdm2 may also bind $p14^{ARF}$, which prevents Mdm2 binding to p53.

The ubiquitin tagged p53 is then degraded by the 26S proteasome. A conserved Cys464 on Mdm2 has been implicated in this mechanism (Honda *et al.*, 1997). Mdm2 has three regions of high identity called CR1, CR2 and CR3 (Figure 2A). CR1 (residues 42-94) binds proteins that are involved in cell growth such as p53, p73, E2F-1 and DP1. CR2 (residue 301-329) is a putative zinc-binding domain and overlaps a region required for binding the Retinoblastoma (pRB) tumour suppressor protein. pRB and p53 act as cell cycle checkpoint components. Unphosphorylated pRB suppresses cellular proliferation in G_1 (Chen *et al.*, 1989) and is found in an inactivated form in various types of human cancers. CR3 (residues 444-483) encodes the RING finger domain and binds two zinc atoms. It contains the Cys464 residue essential for ubiquitination of p53. Also present is a nuclear localization sequence (NLS) and a nuclear export sequence (NES) (Roth *et al.*, 1998). These sequences allow Mdm2 to be shuttled between the nucleus and the cytoplasm and are essential for p53 degradation (Freedman and Levine, 1998). The regulation of p53-Mdm2 binding is controlled by the phosphorylation of residues within or near the interaction domains on either p53 or Mdm2. Phosphorylation of Ser15 on p53 (Shieh *et al.*, 1997) or Ser20 (Unger *et al.*, 1999) (Figure 2B), and Ser17 on Mdm2 (Mayo *et al.*, 1997) has been shown to prevent the binding of p53 to Mdm2 and therefore prevents p53 degradation which would lead to increased cellular levels of p53. An Mdm2 homologue for *D. melanogaster* has not yet been identified. The regulatory mechanism of Dmp53 is not clearly understood although it has been suggested that a PEST region at the N-terminus is responsible for Dmp53 regulation (Jin *et al.*, 2000).

A



B



Figure 2: (A) Schematic diagram of Mdm2 indicating domains that are important for the regulation of p53 function. CR1 is the p53-binding domain, CR2 the zinc-binding motif that binds the Retinoblastoma tumour suppressor protein (pRB), and CR3 the RING finger motif with ubiquitin ligase activity. The nuclear localization sequence (NLS) and the nuclear export sequence (NES) are also indicated. **(B)** The p53 sequence that binds Mdm2. Arrows indicate the residues that make contact with Mdm2 (Chen *et al.*, 1993); underlined residues represent conserved residues predicted to be required for binding to Mdm2 and boxed residues are phosphorylated to inhibit p53-Mdm2 binding.

PEST sequences refer to proline (P), glutamic acid (E), serine (S), and threonine (T)-rich regions in proteins. The function of PEST containing proteins is generally controlled by proteolysis, mostly via the ubiquitin pathway (Rechsteiner and Rogers, 1996).

The E2F-1 protein activates the *ARF* gene. The transcript of this gene, called $p14^{ARF}$, is involved in the degradation of E2F-1 and it interacts with Mdm2. $p14^{ARF}$ binding to Mdm2 stabilises p53 by preventing ubiquitination and degradation of p53 by Mdm2. p53 also functions in suppressing transcription of the *ARF* gene (Stott *et al.*, 1998).

1.5.2 The domain organisation of p53 and/or pRB interacting proteins

Not only does p53 interact with Mdm2, but it also forms interactions with other cellular proteins. I have identified a group of proteins that interact with p53 as well as the tumour suppressor retinoblastoma RB protein pRB. The link between the pRB and the p53 pathways in human cells is that p53 induces expression of the cyclin-dependent kinase inhibitor $p21^{WAF/CIP1}$. This prevents phosphorylation of pRB at the G_1/S transition, which inhibits E2F transcriptional activity and cell cycle progression (Simons *et al.*, 1997). The proteins that interact with p53 and pRB share a similar domain organisation and have the ability to influence the nuclear matrix. All the proteins have a RING finger or RING finger-like domain (Mather *et al.,*2005). These proteins include human RBBP6 (Sakai *et al.*, 1995), PACT (Simons *et al.*, 1997),

P2P-R (Gao *et al.*, 2002), p53BP3 (Zhou *et al.*, 1999b) and Topors (Haluska *et al.*, 1999). A member of this group has also been identified in yeast. The yeast protein, Mpe1, forms part of the cleavage and polyadenylation factor (CPF) (Vo *et al.*, 2001). RBBP6, PACT, P2P-R and Mpe1 are SNAMA orthologous proteins. All these proteins are characterised by a conserved cysteine-rich motif resembling a RING finger domain (Figure 3). Many proteins with a RING finger domain have been shown to have ubiquitin ligase activity (Deshaies, 1999) and this domain will be discussed in detail in section 1.11.

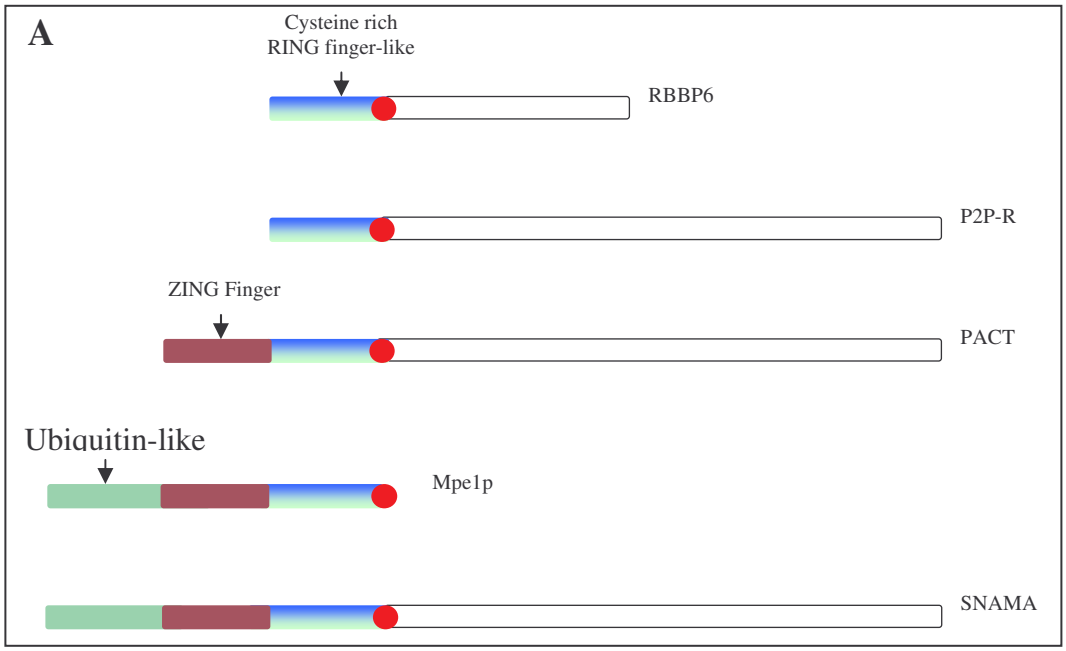
RBBP6

The human RBBP6 protein was identified from a H69c (human small lung carcinoma) library by virtue of its ability to bind the tumour suppressor retinoblastoma RB protein (pRB) (Sakai *et al.*, 1995). pRB suppresses cellular proliferation and is inactivated in various human cancers (DeCaprio *et al.*, 1989; Weinberg, 1992). Underphosphorylated pRB protein is a G_1 cell cycle checkpoint component (Weinberg, 1995). pRB undergoes functional inactivation by being phosphorylated, which allows the cell cycle to proceed into late G_1. RBBP6 is a 948 amino acid polypeptide with a predicted molecular weight of 140 kDa and contains multiple short repeat sequences. These sequences include SRS (12 times), YRE (4 times), VPPP (4 times), and SYSRS (3 times). The SRS sequence is found in DNA binding proteins. SR domain proteins are characteristic of a family of proteins involved in pre-mRNA splicing and are important auxiliary factors involved in

multiple steps of spliceosome assembly (Fu, 1995). RBBP6 binds to underphosphorylated pRB via a short sequence near the C-terminus of the protein.

PACT

The mouse PACT protein was identified by virtue of its ability to bind p53 (Simons *et al.*, 1997). RBBP6 is the truncated human homologue of PACT and occurs as a result of a mutational event in the small lung carcinoma H69c cell line. PACT and RBBP6 share 97% homology from the N-terminus to the end of the SR region suggesting that these two proteins are highly conserved and that the SR region may be of functional significance. PACT was shown to interact with pRB at the N-terminus and p53 at the C-terminus and is implicated in pre-mRNA splicing. PACT binds p53 at the p53 DNA binding region and competes for p53 specific DNA binding. The 250 kDa PACT protein is detected on Western blots only after acetylation and is localized in nuclear speckles, where splicing factors are found (Simons *et al.*, 1997). PACT contains a SR rich region and a basic lysine rich region at the C-terminus of PACT giving this protein a positive charge. The lysine rich region accounts for the unusual behaviour of PACT on SDS-PAGE.



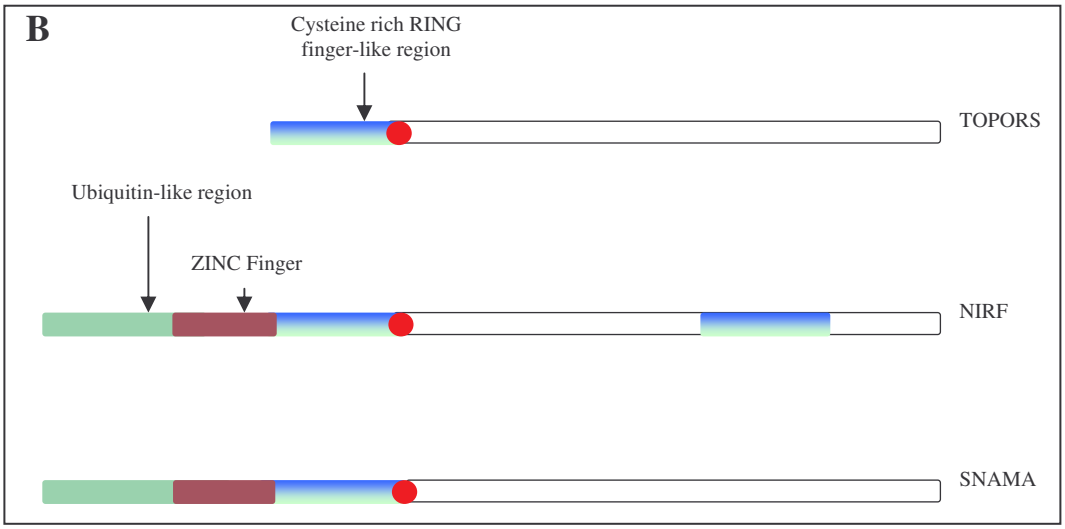


Figure 3: (A) Schematic diagram showing the domain organization of SNAMA orthologous proteins. **(B)** Proteins that share a similar domain configuration with SNAMA. Proteins belonging to this group have a similar domain arrangement. All the proteins shown above contain a cysteine rich RING finger-like domain. TOPORS and NIRF bind p53/pRB downstream the first RING finger-like region.

P2P-R

P2P-R has been identified as being the alternately spliced product of the PACT gene (Scott *et al.*, 2003). P2P-R lacks a single exon containing 102 nucleotides that encodes 34 amino acids. The murine protein was identified in 1997 as proliferation potential proteins (P2P) related (P2P-R) (Witte and Scott, 1997). The 250 kDa P2P-R protein is a basic protein containing multiple domains. There is a cysteine rich domain near the N-terminus that resembles the consensus sequence of the RING finger domain, a proline rich domain, an RS region, and a C-terminal lysine rich region. The N-terminal domain is implicated in cell growth control. The protein has multiple potential nuclear localization signals. The P2P-R protein binds p53 as well as pRB and single-stranded DNA. Amino acids 735 to 908 of P2P-R bind to the pocket domain of pRB, and amino acids 1204 to1314 bind to the oligomerisation and C-terminal regulatory domain of p53 (Witte and Scott, 1997). Expression and localization of P2P-R is regulated by differentiation and cell cycle events. P2P-R localizes to the nucleoli and nuclear speckles in interphase cells, and associates with the periphery of chromosomes during mitosis (Gao *et al.*, 2002). This suggests that P2P-R shares characteristics with other nuclear proteins such as nucleolin and fibrillarin, which associates with the periphery of chromosomes during mitosis. Overexpression of the protein induces mitotic arrest in prometaphase and apoptosis (Gao and Scott, 2002). Apoptosis is induced when P2P-R binds to regions on p53 used to transactivate expression of apoptotic genes. The region on P2P-R responsible for the camptothecin-induced apoptosis is a 158 amino acid sequence at the C-terminal end of the protein (Gao and Scott, 2003). Camptothecin is a topoisomerase I

inhibitor that has been shown to induce apoptosis in MCF-7 cells by increasing the levels of both p53 and p21 waf1/cip1 (Liu and Zhang, 1998). P2P-R has also been shown to interact with scaffold attachment factor B (SAF-B), a MAR (matrix attachment region) binding factor, and the MAR's binding factor nucleolin (Witte and Scott, unpublished).

MpeI

Mpe1 is an evolutionarily conserved *S. cerevisae* protein with a homologue identified in the human genome (Vo *et al.*, 2001). The human homologue has been identified as the DWNN protein and in *D. melanogaster* as SNAMA. Mpe1 has a RNA-binding zinc finger and an uncharacterised N-terminal region. It lacks the lysine rich regions and the nuclear bipartite nuclear targeting region found in the other members in this group of proteins. It is an essential yeast gene necessary for in vitro 3′ end processing of mRNA and is an integral subunit of the CPF complex but not directly involved in the stability of the complex. Mpe1 is essential for cell viability.

Topors

The yeast two-hybrid system was used to screen a HeLa cDNA library and identified a novel gene with p53 binding ability (p53BP3) (Zhou *et al.*, 1999b). p53BP3 is an 815 amino acid protein that has two bipartite nuclear localization signals and is a nuclear protein. The protein is rich in serine and arginine residues characteristic of the SR proteins involved in pre-mRNA splicing (Fu, 1995). The p53 binding region is located at amino acids 226 to 501. The protein is expressed ubiquitously in a wide

variety of tissues. The cDNA sequence of p53BP3 matched the cDNA sequence of *Topors* almost exactly (Haluska *et al.*, 1999). Topors was identified by its ability to bind human topoisomerase I. The protein has a serine/arginine rich region and a RING finger domain. Topors localises in the nucleus of cells with a pattern similar to the nuclear localisation patterns of SR-rich splicing proteins and RING finger proteins. The RING of Topors is similar in sequence to RING finger proteins implicated in transcriptional control (Haluska *et al.*, 1999). The *D. melanogaster* homologue of Topors (dTopors) interacts with *Drosophila* topoisomerase I and Dmp53 (Secombe and Parkhurst, 2004). dTopors has similar structural motifs as Topors. These include the RING finger, SR repeats and bipartite nuclear localization sequences. The SR repeats are reduced in dTopors and lack the consensus RNA binding domain of SR domain splicing factors. dTopors binds specifically to the basic region of Hairy without affecting the ability of Hairy to bind DNA (Secombe and Parkhurst, 2004). Hairy is a primary pair-rule gene encoding a transcriptional repressor that is required for proper segmentation. The process of transcriptional repression plays an important role in many aspects of development. In addition, dTopors has E3 ubiquitin ligase activity and is able to ubiquitinate Hairy and target it for degradation.

The p53 and/or Rb binding proteins have zinc binding domain and a cysteine rich RING finger-like region (Figure 3). In addition they may also have the uncharacterised N-terminal ubiquitin-like region and SR rich regions. SR domain

proteins are important auxiliary factors involved in multiple steps of spliceosome assembly and also play roles in the regulation of alternative splicing (Fu, 1995). Another gene that belongs to this new class of proteins is NIRF (Mori *et al.*, 2002). The nuclear protein NIRF has an ubiquitin-like domain, a zinc finger, two RING fingers and a SR rich region. There is a potential pRb-binding region within the RING finger indicating possible pRB binding. The domain configuration, and its homology to Np95/ICBP90 suggest that NIRF is involved in functions related to cell proliferation. NIRF interacts with the PEST-containing nuclear protein (PCNP) (Mori *et al.*, 2002).

Taken together, the unique domain organisation of these proteins suggest that they are involved in biological processes such as transcriptional regulation, cell cycle control and apoptosis and links splicing to these processes. During evolution these proteins may have maintained both constitutive and regulated splicing functions while others have become transcriptional regulators involved in cell cycle control. In higher organism these proteins appears to function in the ubiquitin pathway as well as in apoptosis.

1.6 THE UBIQUITIN SYSTEM

Intracellular protein degradation plays an important role in the control of the levels of specific proteins and the elimination of damaged proteins. This proteolytic process is a complex process involving a cascade of enzymes that are selective for protein

substrates to be degraded. The process is a major system for selective protein degradation in eukaryotic cells and maintains protein turnover. The steps involved in this process have been reviewed (Hershko and Ciechanover, 1998). Briefly, protein degradation by the ubiquitin process involves two steps: (1) proteins that are targeted for degradation are covalently ligated to ubiquitin molecules in an ATP dependent reaction; and (2) the tagged proteins are then degraded by the 26S proteasome complex that acts only on ubiquitinated proteins. The ubiquitination process is a posttranslational process that requires three enzymatic steps. Initially, free ubiquitin is activated in its C-terminal glycine by a ubiquitin-activating enzyme, E1, forming a thiolester bond to a cysteine residue on E1. Following activation a ubiquitin carrier protein, E2, forms another thiolester bond at the active site cysteine residue on E2 with the activated ubiquitin. Finally the ubiquitin is transferred to the target protein by an E3 ubiquitin protein ligase that is specifically bound to the substrate protein. This results in a covalent isopeptide bond formed between a conserved glycine of the ubiquitin and a lysine of the target protein. In successive reactions additional ubiquitin moieties are linked to one another via a Gly76-Lys48 isopeptide bond to synthesize the polyubiquitin chain. This chain then serves as a recognition signal for the 26S proteasome complex. Polyubiquitin chains are also assembled through conjugation to lysine residues of ubiquitin other than Lys48, and the resulting chains appear to function in other biological processes such as a signal for recruitment of the chain-elongation factor E4, DNA repair, cellular response to stress, inheritance of mitochondrial DNA, endocytosis of specific plasma membrane proteins and

ribosomal function (Spence *et al.*, 1995; Arnason and Ellison, 1994; Fisk and Yaffe, 1999; Soetens *et al.*, 2001; Spence *et al.*, 2000).

Ubiquitin is a highly conserved 76 amino acid protein (Ciechanover *et al.*, 1978) found in eukaryotic organisms (Figure 4). The protein folds into 2 α-helices and 5 β-strands (Cornilescu *et al.*, 1998). It is a thermostable protein found throughout the cell that can either exist as monomers, or conjugated to other proteins. Ubiquitin is conjugated with other proteins through a covalent bond between the conserved glycine at the C-terminal end of ubiquitin and a lysine residue on a target protein. Single molecules of ubiquitin may be attached to the lysine of targeted proteins, but more commonly, ubiquitin chains are added. Ubiquitin is involved in many cellular processes such as cell cycle regulation, innate immunity, signal transduction, development, and apoptosis. The ubiquitin protein is not expressed as a monomer, but instead encoded by a family of genes whose translational products are fusion proteins (Ozkaynak *et al.*, 1984). The ubiquitin gene can be fused to a ribosomal protein gene resulting in a translational product that is a ubiquitin-ribosomal fusion protein, or ubiquitin can exist as a linear repeat resulting in a polyubiquitin molecule. These fusion proteins are then cleaved by the deubiquitinating enzyme (DUB). DUB is a C-terminal hydrolase responsible for the release of ubiquitin monomers.

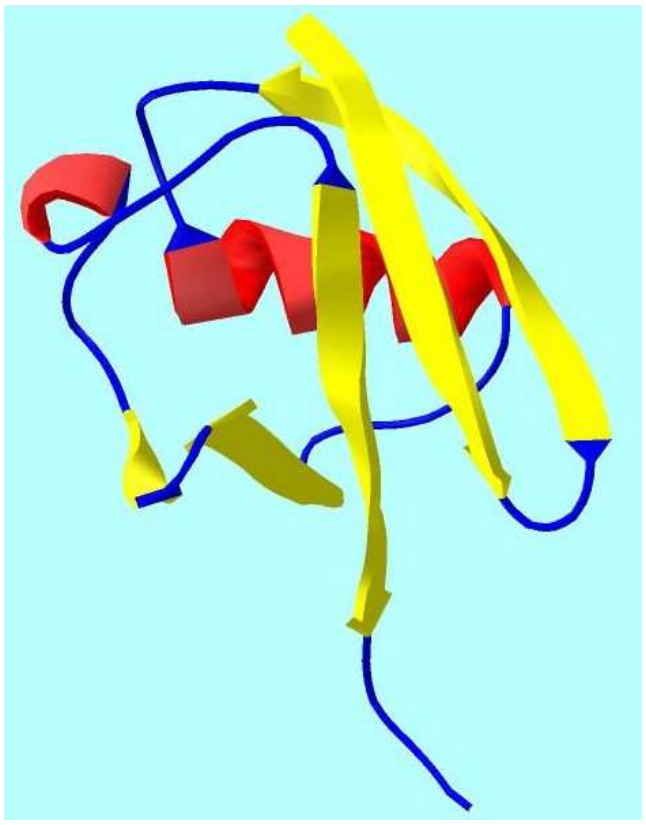


Figure 4: The three-dimensional structure of ubiquitin. The 2-α helices (red) and the 5-β strands (yellow) are clearly visible. This image was generated using the DeepView/Swiss-PdbViewer Version 3.7 (entry #1D3Z; Cornilescu *et al.,* 1998).

1.7 UBIQUITIN ENZYMES

The first enzyme involved in the ubiquitin pathway is the ubiquitin-activating enzyme E1. E1 is involved in the activation of ubiquitin by forming an E1-ubiquitin thioester via a ubiquitin-adenylate intermediate. The activated ubiquitin is then transferred to E2. E1 appears to be a homodimer with an apparent molecular mass of 210 kDa that is composed of two identical 105 kDa subunits (Ciechanover *et al.*, 1982). E1 is highly conserved in evolution and is found both in the cytosol and in the nucleus suggesting that the ubiquitin-ligase system exist in both these compartments. The roles of E1 in the nucleus are not yet known but it is suggested that E1 may be involved in ubiquitination of specific nuclear proteins such as histones (Finley and Chau, 1991) or in the degradation of other nuclear proteins.

The ubiquitin-carrier proteins (E2) are a family of enzymes with distinct substrate specificities and function. These enzymes mediate the transfer of ubiquitin from E1 to the targeted protein. E2s may participate in the recognition of the protein substrate, either directly or in combination with an E3 enzyme (Hochstrasser, 1996; Jentsch, 1992). There is no substantial experimental evidence supporting the idea of the direct binding of E2s to protein substrates but rather E2s associate with E3s that in turn bind specific protein substrates. All known E2s have a conserved domain of ~16 kDa that contains the cysteine residue required for the formation of the ubiquitin-E2 thiol ester intermediate. Certain E2 enzymes are involved in E3 independent mono-ubiquitination that does not target proteins for degradation. Some E2s have

overlapping functions while others are very specific. In *D. melanogaster* the E2 enzyme UbcD1 is involved in the detachment of telomeres in mitosis and meiosis (Cenci *et al.*, 1997). Another example is the *D. melanogaster bendless* gene required for the establishment of synaptic connectivity in embryo development (Muralidhar and Thomas, 1993). E2s in yeast have also been identified conjugating ubiquitin-like proteins Smt3 and Rub1 and not ubiquitin.

The E3 enzymes play an important role in the selectivity of ubiquitin-mediated protein degradation. E3s lack sequence homologies between each other but show sequence similarities between members of the same E3 family. There are a variety of mechanisms by which different types of E3s promote ubiquitin-protein ligation. E3 has been defined as an enzyme that binds, directly or indirectly, specific protein substrates and promotes the transfer of ubiquitin, directly or indirectly, from a thiol ester intermediate to amide linkages with proteins or polyubiquitin chains (Hershko and Ciechanover, 1998). The ubiquitin protein ligases are divided into two large groups, the HECT (homologous to E6-APC-terminus) domain and the RING finger motif (**R**eal **I**nteresting **N**ew **G**ene) containing proteins. The HECT domain E3s function to transfer the activated ubiquitin protein from E2 to an internal cysteine residue on E3 followed by transfer of the ubiquitin to an NH_2 group on the target protein. The RING finger-containing enzyme serves as a scaffold that brings the E2 and the substrate close to each other to allow for the transfer of the ubiquitin from E2 to the substrate.

The *hect* domain family of proteins harbour a 350 amino acid residue sequence homologous to the C-terminal domain of the E6-associated protein (Huibregtse *et al.*, 1995; Huibregtse *et al.*, 1993). This domain contains a conserved active site cysteine residue near the C-terminus to which the activated ubiquitin molecule is transferred from E2 (Scheffner *et al.*, 1995). The highly variable N-terminal region of different *hect* proteins is responsible for the specificity to protein substrates (Huibregtse *et al.*, 1995). The first member of the *hect* family identified was the 100 kDa E6-AP (E6-associated protein). This protein together with human papilloma virus (HPV) oncoprotein E6 is required for the degradation of p53 (Scheffner *et al.*, 1993). Some examples of characterised *hect* proteins include NEDD4 (Staub *et al.*, 1996), which ubiquitinates epithelial sodium channel subunits and the S. *cerevisiae* homologue Rsp5p (Huibregtse *et al.*, 1997).

1.8 THE RING FINGER E3S

The RING finger (**R**eal **I**nteresting **N**ew **G**ene) is a cysteine rich motif found in many proteins. The function of these proteins is to mediate cellular processes such as oncogenesis, apoptosis, development and viral infection (Borden, 2000). The RING finger motif appears to be involved in protein-protein interactions resulting in multi-protein complexes and has potential zinc binding ability. It has been identified as a ubiquitin ligase that facilitates the interaction between E2 proteins and their substrates allowing E2-dependent ubiquitination (Deshaies, 1999) by transferring ubiquitin to target substrates as well as to RING proteins (Tyers and Willems, 1999; Lorick *et al.*,

1999; Joazeiro and Weissman, 2000; Jackson *et al.*, 2000). A mutagenesis experiment has implicated Trp408 in E2 binding (Joazeiro *et al.*, 1999). This aromatic residue is exposed to the outside of the domain and could directly contribute to E2-ubiquitin binding via hydrophobic or aromatic stacking interactions. The RING domain is essential for E3 activity and is defined by a pattern of cysteine and histidine residues arranged as Cys-x_2-Cys-$x_{(9-39)}$-Cys-$x_{(1-3)}$-His-$x_{(2-3)}$-Cys/His-x_2-Cys-$x_{(4-48)}$-Cys-x_2-Cys, where x can be any amino acid (Figure 5) (Saurin *et al.*, 1996). It binds two zinc atoms. The zinc atom is coordinated to four cysteines or three cysteines and a histidine in a tetrahedral arrangement (Borden and Freemont, 1996). This binding is also possible with other zinc ligands resulting in a RING finger family of proteins. These are structurally grouped as either the RING-H2 family that has a histidine residue in place of a Cys4 and the RING-HC that contains only one histidine at position 4. RING fingers with a histidine substitution have also been found. The histidine to asparagine substitution occurring at position 4 in the retinoblastoma-binding protein RBQ-1 is one such example (Saurin *et al.*, 1996). Other RING finger variants include the mouse double minute protein Mdm2 where Cys3 is replaced by threonine (Boddy *et al.*, 1994) without affecting the ubiquitin ligase activity, and CART1 where Cys7 is substituted for an aspartate residue (Regnier *et al.*, 1995). All these substitutions maintain the zinc binding ability and the overall three-dimensional structure of the RING domain.

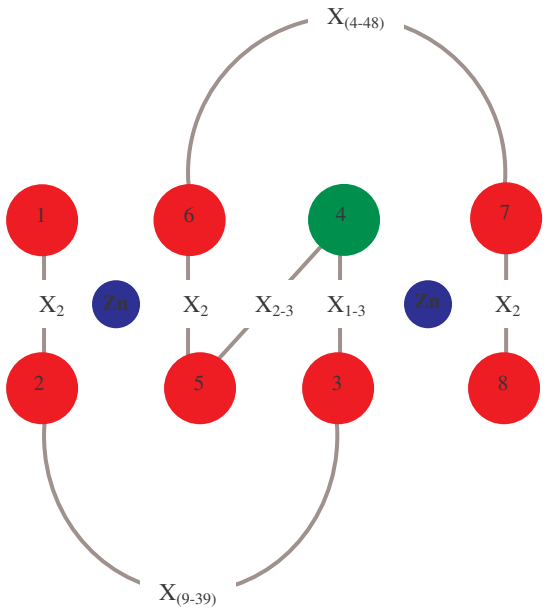


Figure 5: Schematic representation of the ligation scheme employed by the RING finger showing the unique 'cross brace'. Red spheres represent cysteine residues, green spheres represent histidine residues and the blue spheres correspond to zinc atoms. The RING finger coordinates two zinc ions. A His4 to Asn4 substitution is found in the RBBP6 protein.

The RING finger domain-containing E3 family is composed of two distinct groups: single and multisubunit E3 proteins. The N-end rule E3 and the E3α are examples of monomers or homodimers containing both the RING finger domain as well as the substrate-binding/recognition domain. The N-end rule relates the *in vivo* half-life of a protein to the identity of the proteins N-terminal residues (Varshavsky, 1992). E3s are ~200 kDa proteins that bind N-end rule protein substrates that have basic (Type I) or bulky-hydrophobic (Type II) N-terminal amino acid residues with specific binding site for these residues. N-end rule E3 also has a putative "body" site (Type III) to which substrates that do not have a N-end rule terminal amino acid residue bind. E3α has a binding site for a specific E2 (Reiss *et al.*, 1989). This facilitates the transfer of activated ubiquitin from E2 to substrate proteins.

An example of the multisubunit E3 enzymes is the cyclosome (Sudakin *et al.*, 1995) or the anaphase-promoting complex (APC) (King *et al.*, 1995). These have ubiquitin ligase activity specific for cell-cycle regulatory proteins with a nine amino acid "destruction box" motif. The anaphase-promoting complex APC contains the RING finger domain protein Apc11, which has been shown to bind the substrate and catalyse its ubiquitination (Leverson *et al.*, 2000). Another group of multisubunit E3 enzymes contains a RING finger domain involved in E2 recruitment and assembly of other components of the complex, but not in substrate recognition. These E3s are involved in the degradation of cell cycle regulators such as the Sic1p Cdk inhibitor or the G1 cyclin Cln2p. They have been designated as the phosphoprotein-ubiquitin

ligase complex (PULCS) (Hershko and Ciechanover, 1998), since phosphorylation of the substrate is required to convert it to a form susceptible to the action of the ubiquitin ligase complex.

The ubiquitin chain elongation factor, E4, is an additional subset of the E3 group of enzymes. They function in the elongation of a polyubiquitin chain in a specific manner (Koegl *et al.*, 1999). E4 has an essential functional domain called the U-box domain that acts as a scaffold for the transfer of ubiquitin from E2 to a conjugated ubiquitin molecule on the target protein or to elongate short polyubiquitin chains. This process elongates the ubiquitin chains. The U-box is a modified RING finger that lacks the metal chelating residues of RING finger domains (Aravind and Koonin, 2000). Instead the U-box fold is stabilized by salt-bridges and hydrogen bonds. The Cys residues of the RING finger are not conserved in the U-box. The U-box proteins can function as a ubiquitin ligase independent of an E3 (Hatakeyama *et al.*, 2001), or in some instances requiring an E3. E4 is identical to the yeast protein Ufd2 (ubiquitin fusion degradation). The Ufd2 protein is essential for the assembly of the polyubiquitin chain on artificial substrates that are targeted for degradation.

1.9 THE PROCESS OF PROTEIN DEGRADATION

The 26S proteasome is a large catalytic protease (~2.5 MDa) that degrades polyubiquitin proteins into small peptides. The protein is made up of two subunits: a 20S core particle (CP) responsible for the catalytic activity and a regulatory 19S

regulatory particle (RP). The 20S CP is a barrel-shaped molecule composed of four stacked rings: two identical outer α-rings and two identical β-rings. The α- and β-rings are composed of seven subunits with a general structure of $\alpha_{1-7}\beta_{1-7}\beta_{1-7}\alpha_{1-7}$. The catalytic sites are found on the β-subunits (Groll *et al.*, 1997; Lowe *et al.*, 1995). The 20S barrel protein can be capped at either end by the 19S RP. The 19S RP functions to recognise ubiquitinated proteins and substrates of the proteasome. The 19S RP also functions as a lid that opens the outer α-rings to allow entry of the substrate through a gated channel into the proteolytic centre. The 19S regulatory protein is thought to assist in the unfolding of substrates so that they may be inserted into the proteolytic core (Groll *et al.*, 1997; Horwich *et al.*, 1999; Kohler *et al.*, 2001; Strickland *et al.*, 2000).

An important function of the proteasome is the generation of antigenic peptides to be presented by the major histocompatibility complex (MHC) class I molecules to initiate an immune response (Kloetzel, 2001; Rechsteiner *et al.*, 2000; Yewdell and Bennink, 2001). The proteasome is responsible for the degradation of viral proteins following an invasion on the cell. The viral proteins are hydrolysed into short peptides to be transported to the ER and presented to the immune system by MHC class I.

Signals for proteins to be ubiquitinated include phosphorylation of the protein at specific sites and/or the recruitment of ancillary substrate binding factors. One of

these phosphorylation sites is a region found on rapidly degraded proteins known as PEST elements. These regions are rich in proline, glutamic acid, serine and threonine amino acids (Rogers *et al.*, 1986; Rechsteiner and Rogers, 1996). The phosphorylation site for Cdks and other protein kinases have PEST elements rich in S/TP sequences, which are the minimum consensus phosphorylation sites for these proteins (Yaglom *et al.*, 1995). Another degradation signal is the recognition via the N-terminal residues (the N-end rule pathway). The substrates bind directly to the E3 via their N-terminal residues. This signal is inherent in the primary structure of the protein (Varshavsky, 1992) and is responsible for the removal of erroneously transported and compartmentalised proteins in the cytosol after the destabilising N-terminal residue is produced by cleavage with a signal peptidase. The "destruction box" signal is responsible for the ubiquitin-related degradation of mitotic cyclins and certain cell-cycle regulatory proteins. The destruction box is a 9 amino acid sequence motif located approximately 40 to 50 amino acid residues from the N-terminus (Glotzer *et al.*, 1991). The "destruction box" is highly conserved except for the Arg1 and Leu4 (Yamano *et al.*, 1996). The destruction box containing cell cycle regulators are ligated to ubiquitin by the cyclosome/APC and are then degraded.

1.10 DEUBIQUITINATING ENZYMES (DUBS)

Ubiquitin molecules are synthesised as precursors either as a series of ubiquitin monomers or as a fusion to ribosomal proteins. Deubiquitinating enzymes (DUBs) function in maintaining the steady-state levels of free ubiquitin by cleaving ubiquitin

from ubiquitin-protein-fusions and/or isopeptide-bond-linked ubiquitin-protein conjugates (Chung and Baek, 1999; D'Andrea and Pellman, 1998; Wilkinson, 2000). DUBs are cysteine proteases that hydrolyse the amide bond immediately after the C-terminal residue of ubiquitin. There are two groups of DUBs; ubiquitin C-terminal hydrolases (UCHs) and ubiquitin specific proteases (UBPs). UCHs are 20 to 30 kDa enzymes that remove short peptide chains from the C-terminus of ubiquitin. The hydrolases contain a catalytic triad in their active site comprised of cysteine, histidine and aspartate residues, as well as an additional conserved glutamic acid residue within a consensus sequence (Chung and Baek, 1999; D'Andrea and Pellman, 1998). UBPs are ~100 kDa enzymes that cleave the isopeptide bond linking ubiquitin-ubiquitin or ubiquitin-protein and linear ubiquitin chains. These enzymes also contain conserved cysteine, histidine and aspartate residues within an active site residue, but differ in that the consensus sequence is divergent (D'Andrea and Pellman, 1998).

Ubiquitin is a product of the proteasome activity. The ubiquitin is released from the proteasome and recycled back into the pathway (Hough *et al.*, 1986; Swaminathan *et al.*, 1999). The regulatory particles (RP) of the proteasome has a DUB activity responsible for the removal of ubiquitin from protein substrates (Eytan *et al.*, 1993; Holzl *et al.*, 2000; Lam *et al.*, 1997b; Lam *et al.*, 1997a). In *D. melanogaster*, the RP subunit responsible for this activity is p37a.

1.11 UBIQUITIN-LIKE PROTEINS

Several proteins have been identified that are related to ubiquitin or function similarly to modify proteins by the covalent attachment of ubiquitin targeting them for degradation by proteasomes. Ubiquitin-like proteins are divided into two groups. Proteins belonging to the first group modify other proteins in a manner similar to ubiquitin by covalent conjugation of these small proteins to amino acid residues on the target proteins. The modes of conjugation and sites of attachment of these proteins onto target proteins resemble the ubiquitin modification process. These proteins have been designated ubiquitin-like proteins (UBLs). The second group of ubiquitin like proteins bears protein domains that are related to ubiquitin but are unrelated in sequence to each other. These proteins are not conjugated to other proteins and have been termed the ubiquitin domain proteins (UDPs).

1.11.1 Ubiquitin-like proteins (UBLs)

The conjugation pathways for the UBLs are very similar to that of ubiquitin. They form an isopeptide bond with an internal lysine of the substrate and require E1 and E2 enzymes. All known UBLs (except APG12 and URM1) are related by sequence homology to ubiquitin (Figure 6). APG12 and URM1 have been grouped with UBLs since they are small protein modifiers that are activated and conjugated to other proteins in a manner that resembles ubiquitin. Like ubiquitin, UBLs are expressed as inactive precursors and are activated by proteolytic cleavage of the C-terminal

extensions that prevents conjugation. This exposes the functional C-terminal Gly-Gly sequence. APG12, URM1 and FAT10 are expressed as mature proteins. An activating enzyme, E1, then activates the UBLs. The UBLs are then transferred to an E2 enzyme via a thiol-ester bond. The activation reactions of ubiquitin and UBLs are similar but the E1 and E2s involved in the two processes cannot replace one another. Most UBLs have a Gly-Gly sequence at their C-terminus as found in ubiquitin. This GG sequence is necessary for recognition and activation by specific activation enzymes E1s. The modification of cullins by the UBL protein RUB involves SCF/CBC-like E3 enzymes that are also involved in ubiquitination (Jentsch and Pyrowolakis, 2000). Other examples of UBLs include proteins such as SUMO/Smt3p/sentrin (~20% identity to ubiquitin), NEDD8/Rub1 (60% identity to ubiquitin), and APG12 (no sequence homology).



Figure 6: A sequence comparison of ubiquitin to ubiquitin-like proteins. The mammalian proteins UCRP and FAT10 have two ubiquitin-related domains fused in tandem (designated I and II). APG12 and URM1 are unrelated to ubiquitin and are therefore not shown. The alignment was completed using DNAMAN version 4.03 (Lynnon BioSoft).

SUMO

The transfer of SUMO to target proteins involves the E1 heterodimer AOS1-UBA2, a single E2 enzyme UBC9 (Johnson *et al.*, 1997; Johnson and Hochstrasser, 1997) and in yeast a specific E3 (Johnson and Gupta, 2001). Several SUMO molecules can attach to the substrate at distinct sites, but unlike ubiquitin multi SUMO chains do not form on the substrate (Jentsch and Pyrowolakis, 2000). Attachment of SUMO (sumoylation) does not target proteins for degradation although there is a link between the ubiquitin pathway and SUMO conjugation. Sumoylated proteins are found in the nucleus or the nuclear envelope, which may suggest that SUMO attachment to substrates, targets them to the nucleus. Sumoylated PML, a putative RING finger-containing E3 enzyme, is found in the nucleus (Kamitani *et al.*, 1998). The protein SP100 is also modified by SUMO resulting in the translocation from the cytoplasm or nucleus to the nuclear bodies (Muller and Dejean, 1999). Sumoylation of substrates RanGAP1, SP100, p53, $I_KB\alpha$ also increases their levels in the nucleus as compared with the cytoplasm (Yeh *et al.*, 2000). Sumo conjugation to the Ran GTPase-activating protein (RanGAP1) is required for the binding of RanGAP1 to Ran-GTP-binding protein RanBP2. This then targets the associated proteins to the nuclear pore complex (Mahajan *et al.*, 1997; Melchior, 2000). SUMO modification is important in *D. melanogaster semushi* mutants for the nuclear import of Bicoid transcription factor (Epps and Tanda, 1998). The evidence for a link between sumoylation and the ubiquitin pathway is demonstrated by SUMO conjugation to the $I_KB\alpha$ protein. $I_KB\alpha$ associates with NF-$_K$B/Rel transcription factors preventing their movement from the cytosol to the nucleus. Signal induced phosphorylation,

ubiquitination, and proteasomal degradation of $I_KB\alpha$ enables the NF-$_K$B/Rel proteins to enter the nucleus. $I_KB\alpha$ is also modified by SUMO conjugation to Lys21, the site for ubiquitin modification. Sumoylated $I_KB\alpha$ is thus resistant to signal-induced proteasomal degradation (Desterro *et al.*, 1998). Mdm2, the E3 ubiquitin ligase for the p53 tumour suppressor protein, is conjugated to SUMO at Lys446. This reduces Mdm2 self-ubiquitination, which increases Mdm2 ability to ubiquitinate p53 causing increased p53 degradation and resulting in inhibition of p53-mediated apoptosis (Buschmann *et al.*, 2000).

RUB

RUB (related to ubiquitin) is 50% identical to ubiquitin with homologues in all eukaryotes (Liakopoulos *et al.*, 1998). The *S. cerevisiae* protein is called Rub1 and the human homologue NEDD8. RUB is also related to the ubiquitin pathway. RUB conjugation to proteins (rubylation) involves a heterodimer E1 (ULA1/UBA1) and the E2 UBC12 (Hori *et al.*, 1999). The substrates are limited to a small number of proteins belonging to the cullin protein family. Rub1/NEDD8 conjugates to Cdc53/Cul-1, which are components of the SCF E3 ubiquitinating complex. This conjugation stabilises the interaction between Cdc53/Cul-1 with E2s such as Cdc34 or Ubc4 and to a variety of F-box proteins (Kawakami *et al.*, 2001; Liakopoulos *et al.*, 1998) and may play a general role in down regulating SCF activity (Glickman and Ciechanover, 2002).

APG12

This protein is unrelated in sequence to ubiquitin but functions similar to ubiquitin in that it forms a covalent protein complex with a lysine residue of Apg5. Conjugation involves the E1 enzyme Apg7, and the E2 enzyme Apg10 (Klionsky and Ohsumi, 1999). The conjugated protein plays a role in autophagosome vesicle formation and delivery of cytoplasmic material to the lysosome (Furukawa *et al.*, 2000). Both proteins form thioester linked complexes with Apg12. Apg7 is related in sequence to the E1 of ubiquitin, but Apg10 differs from the E2s. The Apg12 system is not directly linked to the ubiquitin system except for its similarity to the conjugation events in the ubiquitin system.

Besides the described UBLs, there are other identified proteins that have mechanistic similarities to ubiquitin with respect to its conjugation to other proteins, but lack sequence similarities to ubiquitin. The ubiquitin related modifier (URM1) is a 99 amino acid protein terminating with a double-glycine (Furukawa *et al.*, 2000) that also conjugates to target proteins. The identification of Apg12 and URM1 indicates that other UBLs may exist that cannot be identified by sequence homology to ubiquitin but rather need to be identified by functional similarities.

1.11.2 Ubiquitin domain proteins (UDPs)

The ubiquitin domain proteins (UDPs) are proteins with a distinct ubiquitin-like domain within a large protein. The ubiquitin-like domain does not always end in a

Gly-Gly sequence but shows sequence homology to ubiquitin and is not processed by DUB-like proteases (Jentsch and Pyrowolakis, 2000) to release the domain, however, the UDP proteins are linked to the ubiquitin pathway. UDPs play roles in mediating protein folding/degradation and in the regulation of signal-transduction cascades, including those that regulate apoptosis.

In addition to the ubiquitin-like domains, most UDPs associate with other characterised domains. Three of these UDPs (Rad23, Dsk2 and Bag-1) have been shown to bind to the 26S proteasome via the ubiquitin-like domain (Kleijnen *et al.*, 2000; Lambertson *et al.*, 1999; Luders *et al.*, 2000; Ortolan *et al.*, 2000). The ubiquitin-like domain binds to the polyubiquitin-binding site in the proteasomal subunit S5a (Mueller and Feigon, 2003). The replacement of the ubiquitin-like domain in Rad23 with ubiquitin demonstrates the structural importance of this ubiquitin-like domain and not its sequence identity to ubiquitin in the binding to the proteasome (Watkins *et al.*, 1993). In yeast the two UDPs Rad23 and Dsk2 overlap functionally with respect to their involvement in the spindle pole body duplication, and Rad23 has an additional role in DNA repair (Biggins *et al.*, 1996). Bag-1 has anti-apoptotic activity. It interacting with Bcl2 and recruits the Hsp70 chaperone to the proteasome.

The list of UDPs is growing and include proteins that play active roles in the ubiquitin system. Scythe, a *Xenopus* UDP, is involved in reaper-induced apoptosis pathways (Thress *et al.*, 1998). Parkin is a UDP that is defective in patients with

Parkinson's diseases. This protein has an ubiquitin-like domain with a di-glycine C-terminal end and two RING finger domains shown to have ubiquitin ligase E2 activity (Zhang *et al.*, 2000). Another UDP is the UIP28 protein. UIP28 is a member of a group of structurally related RING finger proteins that interact with the ubiquitin-conjugating enzyme UbcM4 (Martinez-Noel *et al.*, 1999). The ubiquitin-like domain of UIP28 does not have the di-glycine residues. Most UDPs play a role as adaptors linking the ubiquitin pathway to functions of the proteasome. There is also an indication that chaperones are required in this role.

One example of a UDP that functions similar to ubiquitin by attaching to target proteins is HUB (homologous to ubiquitin). HUB has a 22% sequence identity to ubiquitin but lacks the double-glycine amino acids at the C-terminal end (Luders *et al.*, 2003). Instead, the highly conserved HUB proteins have an invariant C-terminal double-tyrosine followed by a nonconserved single amino acid. The human homologue of HUB is UBL5 (McNally *et al.*, 2003) and the protein in yeast is HUB1. The solution structure of HUB1 indicated a folded molecule similar to that of ubiquitin (Ramelot *et al.*, 2003). HUB exhibits sequence similarity to ubiquitin but its mechanism of attachment does not involve covalent attachment to target proteins. HUB1 has been reported to function as a UBL (Dittmar *et al.*, 2002), but data from (Luders *et al.*, 2003) indicates HUB1 to be a UDP. HUB1 forms SDS-resistant complexes with cellular proteins. These complexes are not formed through covalent C-terminal conjugation of HUB1 to the substrates (Luders *et al.*, 2003). This result

argues strongly against a ubiquitin-like conjugation mechanism for HUB1, but rather suggest that they are formed by a different, possibly non-enzymatic mechanism.

1.12 AIMS

The aim of this study is to characterise the newly identified *D. melanogaster* homologue of the DWNN protein called SNAMA. The DWNN protein was first identified in a mutagenised Chinese hamster ovary (CHO) cell line. A CHO library was mutagenised with the defective retrovirus U3HygroTkNeo. The mutagenesis experiments trapped genes in CHO cells that were involved in the antigen processing and presentation pathway via MHC class I molecules to $CD8^+$ T cells (cytotoxic T lymphocytes (CTL)). One of these genes identified was named DWNN (Domain with no name). This gene contains a highly conserved DWNN domain found in a wide variety of plant and animal species. Since the process of target cell killing by CTL occurs by mechanisms closely related to those described for apoptosis, the CTL resistant CHO cell lines were also screened for resistance to apoptosis inducing chemical agents. CHO cells lacking the DWNN protein were found to be resistant to the apoptotic inducing agent staurosporine. This suggested that the gene mutagenised was not only involved in antigen presentation, but also in the apoptosis pathway.

Characterisation of this gene's role in apoptosis is of primary importance in the understanding of the mechanisms regulating apoptosis. Such knowledge allows for the development of therapeutic agents for a wide range of conditions impacted upon

by the apoptotic pathway. Some of these diseased states prevalent today include conditions such as allergy, autoimmunity and cancer. Knowledge of the apoptotic pathways enables identification of potential drug targets.

CHAPTER 2

MATERIALS AND METHODS

2.1 Media

2.2 Buffers and solutions

2.3 Chemicals and kits

2.4 Bacterial strains

2.5 Oligonucleotides

2.6 Genetic techniques

2.7 General molecular biology methods

2.8 Biochemical methods

2.9 Bioinformatics

2. MATERIALS AND METHODS

2.1 MEDIA

Media was prepared as listed in table 1.

Table 1: Media used in this study.

MEDIA	COMPOSITION
Apple juice agar	1.25% sucrose; 2% agar; 0.025% streptomycin; 0.15% methylparaben (3g dissolved in 30 ml ethanol made up to a litre); and 25% apple juice (v/v).
Larval glucose agar	7.5% glucose; 2.5% agar; 7.5% yeast; 0.15% methylparaben; and 0.025% streptomycin.
LB medium	1% tryptone; 0.5% yeast extract; and 1% NaCl.
LB agar	LB medium + 1.5% agar.
SOB	2% tryptone; 0.5% yeast extract; 10 mM NaCl; 2.5 mM KCl; 10 mM $MgCl_2$; and 10 mM $MgSO_4$.
SOC	SOB + 20 mM glucose.

2.2 BUFFERS AND SOLUTIONS

Buffers and solutions were prepared as listed in table 2.

Table 2: General molecular biology buffers and solutions.

BUFFER	COMPOSITION
50X Tris Acetate EDTA (TAE)	2 M Tris-base; 5.7% Glacial acetic acid; 0.5 M EDTA (pH 8.0).
6X DNA loading buffer	0.25% Bromophenol blue; 0.25% xylene cyanol; 30% glycerol.
10X Tris Borate EDTA	0.45 M Tris-base; 0.45 M Boric Acid; 0.01 M EDTA.
Tris EDTA buffer (TE)	10 mM Tris-HCl (pH 8.0); 1 mM EDTA (pH 8.0).
20X SSC	3 M NaCl; 0.3 M tri-sodium citrate.
Transformation buffer	10 mM K-MES[1] (pH 6.2); 100 mM KCl; 45 mM $MnCl_2.4H_2O$; 10 mM $CaCl_2.2H_2O$; 3 mM Hexamine Cobalt Chloride (Aldrich).
10X MOPS buffer	0.25 M MOPS (Promega); 50 mM Sodium Acetate; 10 mM EDTA (pH 7.0) adjusted with NaOH. Store in dark.
Lysis buffer	25 mM Tris-HCl pH7.5; 150 mM NaCl; 0.2% SDS; 4 mM Pefabloc® SC (Roche).

5X SDS sample buffer	125 mM Tris-HCl pH6.8; 30% glycerol; 10% 2-mercaptoethanol; 4 M urea; 0.0024% Bromphenol blue; 4% SDS
Grinding buffer	5% sucrose; 80 mM NaCl; 100 mM Tris-HCl pH8.5; 0.5% SDS; 50 mM EDTA.
0.1 M Sodium phosphate buffer pH7.2	Prepared from 0.2 M $NaH_2PO_4.H_2O$ and 0.2 M $Na_2HPO_4.7H_2O$.
Hoyer's Mount	30 g of gum Arabic; 50 ml distilled water, stir overnight. 200 g chloral hydrate is added while stirring. Add 20 g glycerol. Centrifuge at 13 000g for 3 hours to clear. Add lactate in a 1:4 ratio.
Western stripping buffer	2% w/v SDS; 62.5 mM Tris.HCl pH 6.8; and 100 mM β-mercaptoethanol.

[1] 1 M MES was prepared, and adjusted to pH 6.3 using KOH, sterile-filtered and stored at -20°C.

2.3 CHEMICALS AND KITS

All chemicals were of AnalaR grade, either from Sigma-Aldrich (South Africa) or BDH Ltd (Poole, Dorset) unless otherwise indicated. Agarose D1 LE (Low

electroendosmosis) was obtained from Techcomp Ltd (Hong Kong). Agar, tryptone and yeast extracts were from Difco laboratories (Detroit, Michigan, U.S.A.). Ampicillin was from Sigma-Aldrich (South Africa). Restriction endonucleases were from Promega. Deoxynucleotide triphosphates (dNTP) and Proteinase K were from Roche Applied Science (South Africa). Unsweetened apple juice (Woodland) and commercial bleach (containing 3.5% sodium hypochlorite) (Jik, Reckitt Benckiser) was obtained from a local department store.

2.4 BACTERIAL STRAINS

The genotype of the bacterial strains used for cloning are listed in table 3.

Table 3: Bacterial strains used in this study

STRAIN	**GENOTYPE**
Epicurian coli XL1 - Blue	recA1, endA1, gyrA96, thi-1, hsdR17, supE44, relA1, lac [F' proAB, lacl^{q}ZΔM15, Tn10 (Tet^{r})]
JM109	recA1, endA1, gyrA96, thi, hsdR17 (r_{K-},m_{K+}), relA1, supE44, Δ(lac-proAB), [F', traD36, proAB, lacl^{q}ZΔM15]

2.5 OLIGONUCLEOTIDES

Oligonucleotides were purchased from Integrated DNA Technologies. The L18F and L18R oligonucleotides were made at the Protein and Nucleic Acid Facility (Cambridge University). Table 4 lists the sequence of the oligonucleutides used in this study.

Table 4: Oligonucleotides used for PCR and RT-PCR

OLIGONUCLEOTIDE	**SEQUENCE**
DWNN1	5' GCATGCATGTCGGTACACTAT 3'
DWNN2	5' AAGCTTGGCGATGGGGATGCG 3'
DWNN4	5' CGTCGGCTCGAGTATGTGCTTGCAGA 3'
L18F	5' CAAAAATGTCTCGCAGTGCTAG 3'
L18R	5' CGTTTCGTTTAGAATAAACGTTGAT3'
M13 Forward	5' GTAAAACGACGGCCAGTG 3'
Pry4	5' CAATCATATCGCTGTCTCACTCA 3'
Plac4	5' ACTGTGCGTTAGGTCCTGTTCATTGTT 3'

2.6 GENETIC TECHNIQUES

2.6.1 Fly stocks and genetic manipulation

All flies were reared at 25°C on standard cornmeal agar food plus yeast. *Canton-S* was used as the wild-type strain. For *SNAMA* gene knockout mutants the *D. melanogaster* EP0503 P-element insertion line *P{PZ}l (2)rQ313*^{rQ313} , abbreviated as *l(2)rQ313* was used. *l(2)rQ313* has a P{PZ} insertion at position 60B8-60B11. This insertion disrupts the expression of the *SNAMA* gene. It was obtained from the EP flystation by Exelixis Inc. PCR was used to identify the site of integration of the P-element (see section 2.8.3). The *CyO* balancer chromosome on this line has an effective lethal phase late embryo-early larva (Kidwell, 1972). To select for homozygous SNAMA mutations, *l(2)rQ313* was crossed with *Cyo pact GFP* (w; Sco/ln(2LR)O,Cy dp[lvl] pr cn[1] y[+] P[w+mC] from the Cambridge *Drosophila melanogaster* stock 12-135 resulting in w;*l(2)rQ313*/GFP;Cyo. The homozygous SNAMA knockout mutants did not express the GFP tag and could then be selected. To establish that *l(2)rQ313* (loss of function mutant) is due to the P-element insertion, the P-element was mobilised by crossing with flies containing Δ2-3 transposase. These excision lines were scored by loss of eye colour and the mobilisation of the P-element was confirmed by PCR using primers flanking the P-element on the SNAMA gene.

2.6.2 Generation of the P-element excision line

Females from *w;Cyo/GFP;Sco* were crossed with *Sb [ry$^+$ Δ2-3]/ TM6b* to mobilize the P-element. Males from this cross which are *w;Cyo/+;Sb [ry$^+$ Δ2-3]/+* have phenotype Curly (Cy) and Stubble (Sb) were crossed with females from *w;l(2)rQ313/GFP;Cyo*. The males which are *w;Cyo/l(2)rQ313; Sb [ry$^+$ Δ2-3]/+* have phenotype Cy, w+ and Sb were then crossed with females from *w;Cyo/GFP;Sco*. This will result in transposition where males are *w;excision/Cyo* with phenotype w (white eyes) and Cy. The excision was confirmed by PCR of genomic DNA to indicate the removal of the P-element (see section 2.8.3). The primer pair used in the PCR flanked the P-element and is complimentary to the genomic DNA. A PCR product would only be possible if the P-element was removed since the large P-element would disrupt amplification. Males *w;excision/Cyo* were crossed with wild type females to set up a stock.

2.7 GENERAL MOLECULAR BIOLOGY METHODS

2.7.1 Small scale preparation of plasmid DNA

Plasmid DNA was purified from *E. coli* cultures by means of two methods:

(i) Qiagen mini plasmid purification kit

(ii) Method described in (Ausubel *et al.*, 1995).

(i) Qiagen mini plasmid DNA purification kit

E. coli cells were grown in 3 ml cultures overnight and plasmid DNA was purified following the manufacturers instructions. The purification protocol of the kit is based

on a modified alkaline lysis procedure, followed by binding of the plasmid DNA to Qiagen anion-exchange resin under appropriate salt and pH conditions. RNA, protein dyes, and low molecular weight impurities are removed by a medium salt wash. Plasmid DNA is eluted in a high salt buffer, and concentrated and desalted by isopropanol precipitation.

(ii) Plasmid DNA preparation by alkaline lysis method

Plasmid DNA was prepared from small (1 ml) overnight cultures of *E. coli* as described in (Ausubel *et al.*, 1995). Briefly, cells were lysed by the alkali method (Birnboim and Doly, 1979; Ish-Horowicz and Burke, 1981). The plasmid DNA was resuspended in TE buffer (pH 8.0).

2.7.2 Large scale preparation of plasmid DNA

Large amounts of DNA were prepared using the Qiagen plasmid Maxi purification kit following the manufacturers instructions. Typically 100 ml cultures were grown in LB medium with appropriate antibiotic selection and the Qiagen-tip 500 was used to purify DNA from RNA, proteins and other cellular contaminants in the bacterial lysate. The purification protocol of the kit is based on a modified alkaline lysis procedure, followed by binding of the plasmid DNA to Qiagen anion-exchange resin under appropriate salt and pH conditions. RNA, protein dyes, and low molecular weight impurities are removed by a medium salt wash. Plasmid DNA is eluted in a high salt buffer, and concentrated and desalted by isopropanol precipitation.

2.7.3 Digestion of DNA fragments with restriction enzymes

Digestion with restriction enzymes was carried out as described (Ausubel *et al.*, 1995). Reaction conditions were those recommended by the manufacturer of the enzymes. All reactions were incubated at 37°C overnight. Adding 1X DNA loading buffer to the digestion after the overnight incubation stopped the reactions.

2.7.4 Purification of restriction enzyme-digested DNA fragments

The fragments of interest were carefully removed from the agarose gel after electrophoresis and purified using the Concert™ Matrix Gel Extraction System (Life Technologies Gibco-BRL) following the manufacturers instructions provided with the kit. This method employs a silica resin to which single and double-stranded DNA bind without binding any DNA contaminants. The agarose gel is dissolved by sodium perchlorate and the DNA is adsorbed onto the silica support (Vogelstein and Gillespie, 1979). Adsorption to the silica is influenced by the buffer composition and temperature. Agarose and electrophoresis buffers are removed with alcohol-containing wash buffers. The DNA of interest is then eluted with TE buffer (pH 8.0).

2.7.5 Preparation of competent *E. coli* cells

E. coli XL1- Blue cells were made competent to take up foreign supercoiled DNA by modification of the method of Hanahan (Maniatis *et al.*, 1989). Exponentially growing cells (OD_{550} =0.4) were pelleted and resuspended in two changes of TFB (Table 2). To make the cells competent, DMF (7 µl/200 µl cells) was added. After 5

minutes 0.21 M DTT (7 µl/200 µl cells) was added. After 10 minutes 7 µl DMF/200 µl cells was added. All steps were carried out at 4°C. Efficiency of transformation was typically $5x10^6$ cfu/µg.

2.7.6 Transformation of *E.coli* cells with plasmid DNA

Competent *E. coli* XL1-Blue cells were transformed (Maniatis *et al.*, 1989) with 30-50 ng DNA per 200 µl aliquots and allowed to stand on ice for 30 minutes, heat shocked for 90 seconds at 42°C and returned to ice for 2 minutes. 800 µl cold SOC was added and the cells were incubated at 37°C for 45 minutes to 1 hour. Cells were then plated onto LB agar selection plates with appropriate antibiotics, or pelleted and resuspended in 1 ml SOB + 20% glycerol in eppendorf tubes, snap-frozen in liquid nitrogen and stored at –70°C.

2.7.7 Isolation of total RNA from *D. melanogaster* embryos

Total RNA was isolated from *D. melanogaster* embryos by either of two methods:

(i) High Pure RNA isolation Kit – Roche Molecular Biochemicals.

(ii) Method using TRIzol® LS reagent – Life Technologies Gibco-BRL.

(i) Isolation of RNA by the High Pure RNA isolation Kit

Embryos were collected, dechorionated in commercial bleach, and transferred to a homogeniser. Embryos were homogenised in 400 µl lysis/binding buffer provided in the kit. The homogenate was then loaded onto columns (kit) packed with glass fibre

fleece. The RNA binds specifically to the surface of the glass fibres since the binding conditions have been optimized for RNA binding. DNase I is then applied directly onto the glass fibre fleece and digests the contaminating DNA. The bound RNA was then washed and eluted in water. The RNA was heated at 65°C for 10 minutes. Concentration was determined by spectrophotometry. The RNA was stored at –20°C until the various developmental stages were collected. RNA prepared by this method was used in the RT-PCR reactions.

(ii) Total RNA isolation with TRIzol® LS reagent

TRIzol® LS reagent is based on the guanidinium method of (Chomczynski, 1993). Embryos or adult flies were collected and transferred to a homogeniser. Embryos were homogenised in 1.5 ml tubes with 0.75 ml TRIzol/100 mg embryos. For isolation of RNA from adult flies, 50 flies were homogenised in 500 µl TRIzol reagent. 150 µl chloroform was then added. Samples were spun at 13 000 x g for 10 minutes at 4°C. The RNA was then precipitated from the supernatant (aqueous layer) after the addition of 350 µl isopropanol and incubation at room temperature for 2 minutes. The RNA was recovered by spinning samples at 13 000 x g for 15 minutes at room temperature. The pellet was washed with 70% cold ethanol and dissolved in 50 µl RNase free water. Samples were heated at 65°C for 10 minutes, and then put on ice immediately. Concentrations were determined by spectrophotometry. 50 µg aliquots per tube were prepared by precipitating the RNA from the relevant volume with 2,5 volumes of 100% cold ethanol. The RNA was stored in ethanol at –20°C.

2.7.8 Preparation of whole embryo extracts

To make small (~ 1 ml) preparations of total proteins, the embryos were collected and homogenised in 5X SDS sample buffer using a glass homogeniser. The homogenate was then transferred to microfuge tubes and large particulate matter was removed by spinning at 10 000 x g for 3 minutes. The supernatant was then retained for protein estimation and SDS-PAGE. Protein estimation was done using the Bio-Rad protein Assay (Bio-Rad Laboratories, U.S.A.). The assay is a dye-binding assay based on the differential colour change of a dye in response to various concentrations of proteins (Bradford, 1976). After protein estimation, SDS sample buffer was added and samples were loaded onto the polyacrylamide gel.

2.7.9 Preparation of genomic DNA

Genomic DNA was prepared using a protocol based on the method described by (Bender *et al.*, 1983). 20 flies were collected and homogenised in 100 µl grinding buffer. After 30 minutes incubation at 70°C, 35 µl of 8 M KOAc was added and incubated on ice for 30 minutes to pellet the DNA (Hamilton and Zinn, 1994).

2.8 BIOCHEMICAL METHODS

2.8.1 *In situ* hybridisation to *D. melanogaster* embryos

In situ hybridisation was done to detect the localisation of *SNAMA* RNA transcripts in whole mount *D. melanogaster* embryos using the method described by (Tautz and

Pfeifle, 1989) with modifications to the protocol as followed in Dr Alfonso Martinez-Arias laboratory (Cambridge) modified by Charlie Oh and Bruce Edgar. Solutions were prepared as described in table 5.

Table 5: Solutions used for *in situ* hybridisation study.

SOLUTION	**COMPOSITION**
Phosphate buffered saline (PBS)	7 mM Na_2HPO_4; 3 mM NaH_2PO_4; 150 mM NaCl.
Phosphate buffered saline with Tween 20 (PBT)	PBS; 0.1% Tween 20.
Embryo wash (10x)	0.4% Triton X-100; 7% NaCl.
Fixative A	PBS; 4% formaldehyde.
Fixative B	PBS; 10% formaldehyde.
Hybridisation solution (HS)	50% deionised formamide; 5X SSC; 100 µg/ml sonicated and boiled salmon sperm DNA; 100 µg/ml calf liver tRNA; 50 µg/ml heparin; 0.1% Tween 20.
Staining solution	100 mM Tris (pH 9.5); 50 mM $MgCl_2$; 100 mM NaCl; 1 mM Levamisole; 0.1% Tween 20.

Collection and preparation of embryos

Embryos were collected from apple juice agar plates by washing them with water into a small basket. The embryos were then dechorionated in commercial bleach for 2-3 minutes and washed thoroughly with alternate changes of water and 1X embryo wash to remove the bleach. After the last wash with water the basket was dabbed briefly on a paper towel to remove excess water. The embryos were then collected into a 2 ml tube containing a 1:3 mixture of Fixative A: Heptane.

Fixation and devitellinisation

The embryos were fixed by shaking vigorously for 20 minutes. After removing the aqueous phase and the heptane, the embryos were devitellinized by shaking vigorously in methanol at room temperature. The methanol was then removed, and new methanol was added to wash the embryos. The methanol wash step was repeated 3 times. The embryos were then stored in methanol at –20^{o}C.

Preparation of probes

The probe was prepared from the POT2 vector (Figure 18) containing the cDNA for *SNAMA*. The vector was obtained from the BDGP EST sequencing project. This vector was digested with *EcoRI* and *XhoI* (see section 2.7.3) releasing a 914 bp *SNAMA* fragment. For a positive control, the *dpp* gene was used. The vector containing the *dpp* gene was digested with *EcoR1* and *Hind III* releasing the 3200 bp fragment. These digests were run on a 1.5% TBE agarose gel. The bands at 914 bp for *SNAMA* and 3200 bp for *dpp* were excised. These DNA fragments were purified

out of the gel fragments using the Concert™ Matrix Gel Extraction System (Life Technologies) (see section 2.7.4).

Labeling of probe with digoxigenin-11-dUTP

DNA fragments to be used as probes were labelled with digoxigenin-11-dUTP by the random primed technique using the Dig-High Prime enzyme mix (Roche Molecular Biochemicals) as per the manufacturer's protocol. 1 µg of DNA was made up to 16 µl with deionised water, boiled for 10 minutes and chilled immediately for 5 minutes in an ethanol bath at –20°C. The tubes were then placed on ice and 4 µl Dig High Prime enzyme mix was added. The reaction was then incubated at 37°C overnight. The labelling reaction was stopped by heating at 65°C for 10 minutes.

Removal of unincorporated free DIG-dUTP

The unincorporated free DIG-11-dUTP nucleotides were removed using the High Pure PCR Product Purification Kit (Roche) as per the manufacturers protocol. The kit contains High Pure filter tubes pre-packed with special glass fibres. In the presence of the chaotropic salt, guanidine thiocyanate, the labelled DNA binds selectively to the surface of the glass fibres in the column. The wash steps remove the unincorporated free nucleotides. The bound DIG-labelled DNA is then eluted using a low-salt buffer.

Yield of the DIG-High Prime labelling reaction

The yield of DIG-labelled DNA was then estimated using the DIG Quantification and DIG Control Teststrips (Roche) following the manufacturers instructions. Briefly, a series of dilutions of the DIG-labelled DNA is applied to the squares on the DIG Quantification Teststrips. DIG Control Teststrips are already loaded with 5 defined dilutions of a control DNA and are used as standards. The teststrips are then subjected to immunological detection with Anti-Digoxigenin-AP conjugate and the premixed colour substrates NBT/BCIP. The DIG labelling efficiency is then determined by visually comparing the signal intensities of the spots on the quantification teststrips with those on the control teststrip. Before hybridisation, 50 ng of the labelled probe in 40 µl HS was boiled for 10 minutes and immediately chilled in an ethanol bath (-20°C) for 5 minutes, made up to 100 µl with HS and used immediately on the prehybridised embryos.

Hybridisation and staining

Before hybridisation the embryos were aliquoted into 50 µl volumes, rinsed in 1 ml methanol, again in PBT-methanol (1:1), in 1 ml PBT and then "postfixed" (a second fixation step) by shaking in Fixative B for 20 minutes at room temperature. The embryos were then rinsed twice, washed for 5 minutes and rinsed once again in PBT. They were then incubated for 3 minutes in PBS containing 50 µg/ml Proteinase K. The digestion was stopped by removing the Proteinase K solution and adding PBT containing 2 mg/ml glycine for 2 minutes. Embryos were then washed 2 x 5 minutes in PBT. Embryos were re-fixed in 5% formaldehyde in PBS for 20 minutes. This was

followed by 5 x 5 minute PBT washes. The PBT was replaced with 1:1 mixture of PBS: Hybridisation buffer (HS) and the tube was rotated for 20 minutes at room temperature followed by a further 30 minutes in HS at room temperature. To pre-hybridise, fresh HS was added and the tubes containing embryos were immersed in a water bath at 48°C for an hour. To hybridise, 100 µl of probe was added after the removing of HS. The tubes were then immersed in a water-bath at 48°C and left overnight without agitation. The following day 500 µl of HS (preheated to 48 °C) was added and the tubes were returned to the 48 °C water-bath for a further 15 minutes. This was followed by several 500 µl washes in decreasing concentration of HS in PBT: 70% at 48 °C, 50% at 48 °C, and 30% at 48 °C, rinsed twice in PBT, then washed three times for 5, 10, and 15 minutes in 1 ml PBT at room temperature. At this stage, the embryos were transferred to a multi-well (tissue culture) dish and incubated at room temperature with agitation for 2 hours in pre-absorbed alkaline phosphatase-conjugated anti-Digoxigenin antibody (supplied in the DIG High Prime Labelling and Detection Kit I from Roche Applied Science) diluted 1/2000. To detect the alkaline phosphatase linked to the anti-DIG antibody, the embryos were rinsed 4 x 20 minutes in PBT, followed by 2 x 5 minute washes with the staining solution. To the last rinse 8 µl of the colorimetric substrate premix NBT/BCIP (supplied in the kit) was added and the embryos were incubated without agitation in the dark at 4°C. Staining takes 30 minutes to several hours. Washing in PBT stopped the reaction.

Mounting

The embryos were washed thoroughly in PBS and then in increasing concentration (30%, 50%, and 80%) of glycerol in PBS. In this form, the embryos could be placed on a microscope slide with a coverslip and sealed with nail varnish.

2.8.2 Northern hybridisation

Agarose gel electrophoresis and transfer of RNA to membrane

RNA was detected as described previously (Maniatis *et al.*, 1989). Total RNA from various stages of embryo development was extracted as described in section 2.7.7 (ii). The ethanol was removed from the RNA samples by suction followed by a 70% ethanol wash step. After the wash step the RNA was dried at 37°C for 2 minutes. To this, 20 µl RNA loading buffer was added. The samples were heated at 65°C for 10 minutes and then placed on ice. The RNA populations were then separated according to size by electrophoresis through 1% agarose gels containing 6% formaldehyde cast in 1X MOPS buffer. The RNA was transferred onto Hybond™ Nylon hybridisation transfer membranes (Amersham Pharmacia biotech) by capillary transfer in 10X SSC. To fix the RNA, the slightly damp membrane was exposed to 120 000 µjoules/cm^2 UV light for 30 seconds using the Stratagene UV crosslinker.

Labelling of DNA probe with [α-^{32}P]dCTP

The *redi*prime™ II (Amersham Pharmacia biotech) random prime labelling system was used following the manufacturer's instructions. Typically 25 ng linearised and denatured DNA was labelled with Redivue [α-^{32}P] dCTP (Amersham Pharmacia

biotech) using the individually dispensed Rediprime reaction mix. The reaction mix consists of a buffered solution of dATP, dGTP, dTTP, exonuclease free Klenow enzyme and random primers in a dried, stabilised form. Reactions were labelled overnight at room temperature. After labelling, the unincorporated free nucleotides were removed using the High Pure PCR Product Purification Kit (Roche Applied Science) as per the manufacturers protocol (see section 2.8.1). 14 µl of the labelled probe per 5 ml hybridisation buffer was used as recommended by the manufacturers.

Hybridisation and detection

Blots were hybridised using the Rapid-hyb buffer (Amersham Pharmacia biotech) following the manufacturers protocol. This rate enhanced hybridisation buffer allows single copy genes to be detected after only a 2 hour hybridisation with ^{32}P labelled DNA probes at 65°C. The non-specifically bound radiolabels were removed by 2 washes with 0.1X SSC buffer at 65°C. Blots were wrapped in UV-transparent wrap and autoradiography was performed at –70°C.

Stripping of blots to be reprobed

The probe was removed from the blot by pouring a boiling solution of 0.5% (w/v) SDS solution. The membrane was allowed to cool to room temperature before it was pre-hybridised and hybridised with a new probe. Typically membranes were probed twice.

2.8.3 The polymerase chain reaction (PCR)

The site of the integration of the P-element in *l(2)rQ313* and the mobilization of the P-element in the excision crosses was confirmed by PCR. The DWNN specific primers DWNN2 and DWNN4, and the P-element primers Plac4 (P-element 5' flank) and Pry4 (P-element 3' flank) were used. DNA was amplified using the Expand High Fidelity PCR System (Roche Applied Science). The system is composed of a unique enzyme mix containing thermostable *Taq* DNA polymerase and a proofreading polymerase (Barnes, 1994). The final amounts used were:

DNA sample	400 ng
Primers	0.3 µM
10X PCR buffer with 15 mM $MgCl_2$	1X
dNTP mix	2.5 mM
Expand High Fidelity enzyme mix	0.1 U/µl

The cycle conditions were: denaturation at 94°C for 30 seconds, annealing at 60°C for 30 seconds and extention at 68°C for 1 minute over 35 cycles.

2.8.4 Reverse transcription PCR (RT-PCR)

A quantitative RT-PCR was performed using the Access RT-PCR system (Promega) using volumes suggested by the manufacturer. The *L18* gene (Ntwasa *et al.*, 1994) was used as a control for equal amounts of RNA loading. The following amounts were used:

RNA sample	2 µg
Primer 1 (DWNN1/L18F)	1 µM
Primer 2 (DWNN2/L18R)	1 µM
AMV/*Tfl* 5X reaction buffer	1X

dNTP mix	0.2 mM
$MgSO_4$	1 mM
AMV reverse transcriptase	0.1 U/µl
Tfl DNA polymerase	0.1 U/µl
Nuclease-free water	Final volume of 50 µl

The cycle conditions were as follows: 48°C for 45 minutes; 94°C for 2 minutes; followed by 40 cycles of: 94°C for 30 seconds, 55°C for 1 minute, 68°C for 1 minute; and a final extension at 68°C for 7 minutes.

2.8.5 Western blotting

Whole embryo extracts (section 2.7.8) were electrophoresed on a 6% polyacrylamide gel and transferred with Towbin buffer to a Hybond–P: PVDF membrane (Amersham Bioscience) (Towbin *et al.*, 1979). Non-specific binding to the membrane was prevented by blocking with the SuperBlock™ Dry Blend Blocking buffer (Pierce). The blot was incubated with either anti-SNAMA or anti-Dmp53 antibody. The anti-SNAMA antibody used was prepared by immunising rabbits against the DWNN domain of the SNAMA protein and the anti-Dmp53 antibody was obtained from Santa Cruz Biotechnology (p53 (dD-21), 200 µg/ml). The p53 (dD-21) antibody is an affinity purified goat antibody raised against a peptide mapping within an internal region of p53 of *D. melanogaster* origin.

The primary anti-SNAMA antibody used was a 1:10 000 dilution for 1 hour, and a 1:10 000 dilution of peroxidase conjugated goat anti-rabbit immunoglobulin secondary antibody (Sigma-Aldrich) for 1 hour. A 1:3000 dilution of anti-Dmp53 and

a 1:10 000 dilution of peroxidase conjugated rabbit anti-goat immunoglobulin secondary antibody (Santa Cruz Biotechnology) was used. To block the Dmp53 antibody binding site, Dmp53 blocking peptide (Santa Cruz Biotechnology) was used at a concentration of 20 ng/ml. Between steps, the membrane was washed in PBS + 0.1% Tween 20 twice for 10 minutes each time except after the secondary antibody when it was washed three times for 5 minutes in PBS + 1% Tween 20. Bound primary antibody was detected using the SuperSignal® West Pico chemiluminescent substrate (Pierce) after a 30 second exposure to X-ray film. Blots were stripped of blocking reagent, antibody and SuperSignal® substrate by incubation for 30 minutes at 70°C in Western stripping buffer. Membranes were then washed twice for 10 minutes per wash before they could be reprobed as indicated before.

To activate the Dmp53 pathway larvae were allowed to develop on media supplemented with the topoisomerase inhibitor Camptothecin. Camptothecin has been shown to activate the p53 pathway (Takada *et al.*, 2003). Larvae media was prepared as before and supplemented with either 1mM Camptothecin, or DMSO. The DMSO control was set up since Camptothecin was dissolved in DMSO. Larvae were allowed to develop overnight on this media and proteins were collected as before. A Western blot was performed and probed with the anti-SNAMA and anti-Dmp53 antibody.

2.8.6 Acridine orange staining

To detect cells undergoing apoptosis embryos were stained with acridine orange (Abrams *et al.*, 1993). They were dechorionated in bleach, washed with embryo wash, and then rinsed with water. The embryos were then placed in equal volumes of heptane and 5 µg/ml acridine orange (Sigma-Aldrich) in a 0.1 M sodium phosphate buffer, pH 7.2. After shaking for 5 minutes**,** embryos at the interphase were removed and placed on a cavity slide with Series 700 halocarbon oil (Sigma-Aldrich). The samples were viewed using a confocal microscope.

2.8.7 Cuticle preparation

Wild type embryos were collected on apple juice agar plates. They were allowed to age for 24-36 hours at 25°C. Unhatched embryos were rinsed into a nytex screen and dechorionated in commercial bleach until the embryos float to the surface. Embryos were rinsed and transferred to a vial with a 5 ml PBS/5 ml Heptane mixture. Embryos in upper heptane solution were then transferred to another tube, the volume was adjusted to 0.5 ml and an equal volume of 4% paraformaldehyde in PBT was added. Tubes were shaken for 20 minutes after which the lower aqueous layer was removed. An equal volume of methanol was added and tubes shaken for 15 seconds to devitellenize embryos. At this step most of the embryos settled to the bottom of the tube. Liquid was removed and embryos were washed 3X with methanol. They were then incubated in 1:4 glycerol: acetic acid for 1 hour at 65°C. Embryos were mounted

on glass slides in a drop of Hoyer's Mount and dried overnight at 60°C. Cuticles were examined using dark field microscopy.

For the P-element mutant *P{PZ}l (2)rQ313*^{rQ313} unhatched embryos were picked from apple juice agar plates into water in a small dish. The embryos were dechorionated with undiluted bleach added to the dish and washed with embryo wash and water. The embryos were incubated in 1:4 glycerol: acetic acid for 1 hour at 60°C. Embryos were cooled to room temperature, mounted on glass slides in a drop of Hoyer's mount and viewed using dark field microscopy.

2.9 BIOINFORMATICS

Sequence alignments and phylogenetic trees were completed using the DNAMAN programme version 4.03 (Lynnon BioSoft). This programme was used for the generation of the phylogenetic tree. This tree is setup with the distance matrix using the Neighbor-Joining method (Saitou and Nei, 1987). Protein domains of SNAMA were identified by the SMART (Simple Modular Architecture Research Tool) program (Schultz *et al.*, 1998), Pfam protein family database (Bateman *et al.*, 2004) and ScanProsite (Hulo *et al.*, 2004). For promoter prediction, the BDGP 1999 Neural Network Promoter Prediction (NNPP) version 2.2 programme was used (Reese, 2001). Pair sequence alignments of the DWNN domain from various organisms was done using the ALIGN Query on the Genestream server at Genestream Resource Centre in France (Pearson *et al.*, 1997). The ubiquitin-like domain and the secondary

structure of DWNN was analysed using FUGUE v2.s.07 (Shi *et al.*, 2001) and PSIPRED (Jones, 1999). A comparison of the secondary structure of ubiquitin to the DWNN domain of SNAMA was made using the Network Protein Sequence Analysis tool (NPSA) (Combet *et al.*, 2000). The web addresses of the bioinformatics tools used are listed in table 6.

Table 6: Web addresses for the various bioinformatics tools used.

TOOL	WEB ADDRESS
ALIGN Query	www2.igh.cnrs.fr/bin/align-guess.cgi
BLAST	http://www.ncbi.nlm.nih.gov/BLAST
FlyBase	http://flybase.bio.indiana.edu
FUGUE	http://www-cryst.bioc.cam.ac.uk/fugue
GeneBank	http://www.ncbi.nlm.nih.gov
NNPP	http://www.fruitfly.org/seq_tools/promoter.html
NPSA	http://www.cmpharm.ucsf.edu/~nomi/nnpredict.html
Pfam	http://pfam.wustl.edu/hmmsearch.shtml
PSIPRED	http://bioinf.cs.ucl.ac.uk/psipred
ScanProsite	http://ca.expasy.org/tools/scanprosite
SMART	http://smart.embl-heidelberg.de

CHAPTER 3

IDENTIFICATION AND SEQUENCE ANALYSIS OF *SNAMA*, THE DWNN-CONTAINING HOMOLOGUE, IN *Drosophila melanogaster*

3.1 Introduction

3.2 Sequence analysis of the *SNAMA* gene

3.3 The domain organisation of SNAMA

3.1 INTRODUCTION

This research describes the *Drosophila melanogaster* homologue of a gene product identified by Dr Allison George. She used a system of promoter trap retrovirus to induce mutations in Chinese hamster ovary (CHO) cells (Chang *et al.*, 1993; Hubbard *et al.*, 1994; George, 1995), to create cell lines that are defective in the processing or presentation of antigen by the major histocompatibility complex (MHC) class I molecules. She was interested in identifying novel components of the cellular pathways involved in the recognition and killing of target cells by cytotoxic T lymphocytes (CTL). Dr George created CHO cell lines that were susceptible to lysis with specific CTL and were susceptible to infection with the defective murine retrovirus (the insertional mutagen). These cells were then infected with the promoter-less retrovirus to create loss of function mutants in the immunity pathways. Clones resistant to CTL killing were then selected and their DNA analysed to identify which genes have been disrupted. These mutational experiments generated approximately 40 mutant clones that are defective in the antigen processing or presentation pathway (George, 1995), three of which showed 100% resistance to CTL killing compared to the parental cell line, whilst the rest of the clones tested showed reduced levels of resistance. The process of target cell killing by CTL occurs by mechanisms closely related to those described for apoptosis.

Foreign viral proteins within the cell cytoplasm of antigen presenting cells are processed and presented to the cell surface by MHC class I molecules. Upon

presentation of these epitopes to the cell surface the specific CTL engages with the antigen presenting cell. The antigen presenting cell is then lysed by one of two mechanisms: either the secretion of soluble cytolytic proteins of lymphocytic origin, or the engagement of the Fas molecule with its ligand FasL to induce apoptosis. Although the induction of apoptosis by the FasL/fas interaction and by apoptosis inducing chemical agents follows different pathways, the activation of caspase is a central event in both pathways. Thus the CTL resistant cell lines were further screened for resistance to apoptosis induction by various apoptosis inducing agents. Three CTL resistant cell lines were found to show resistance to apoptosis induction by staurosporine (the chemical inducer of apoptosis) (Pretorius, 1999).

One of these cell lines labelled Mut 7(3xHA8)3.5hrs was analysed by inverse PCR to obtain the 5′ sequence adjacent to the site of retroviral integration (Pretorius, 1999). This sequence matched 21C4 in the human EST (expressed sequence tag) database without any entries in the DNA or protein databases. The full-length 21C4 sequence of 929 bases was compared with the protein database using BLASTX 2.0 (basic local alignment search tool) to identify open reading frames from homologies. The matching related protein sequences were aligned using CLUSTAL X multiple sequence alignment software. The sequence alignment showed a significant match to a highly conserved N-terminal domain called Domain With No Name (DWNN), found in a wide variety of eukaryotic organisms but is absent in prokaryotes (Figure 7).

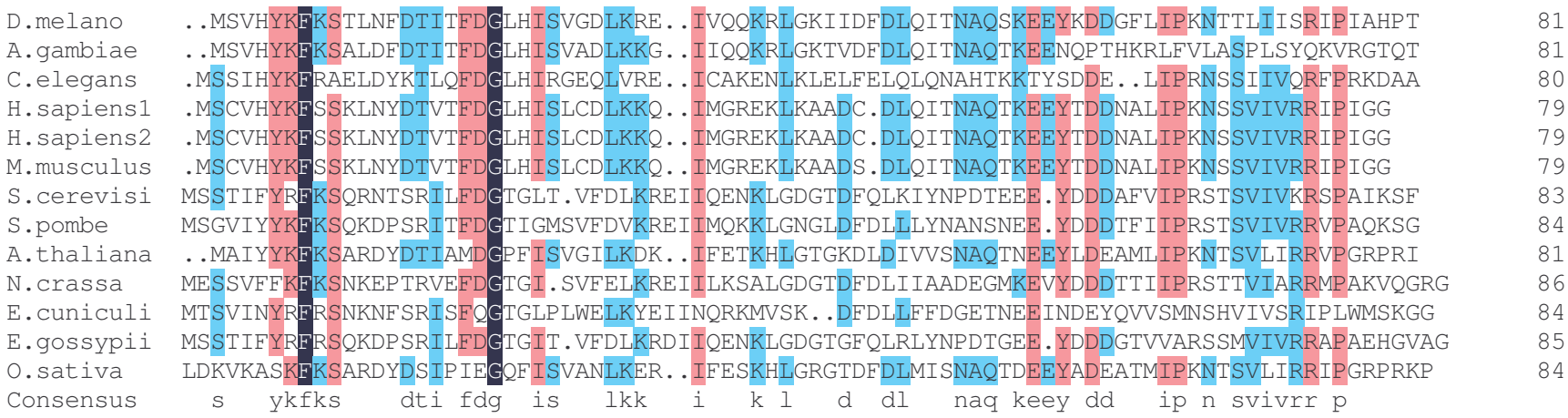


Figure 7: Amino acid sequence and alignment of the conserved DWNN domains with homologous DWNN domains from a variety of eukaryotic organisms. In *H.sapiens* two transcripts have been identified and have been labelled 1 and 2. The high level of conservation indicates that this gene codes for an important and well-adapted protein. The alignment was completed using DNAMAN version 4.03 (Lynnon BioSoft). Identical residues are shaded in black, residues with greater/equal to 75% homology level to the consensus are shaded in pink and residues with a homology level greater/equal to 50% are shaded in blue.

The gene is also found in organisms with primitive genomes. The DWNN gene is found in protozoa species that have a minimum complement of genes essential for viability. This level of conservation between such diverse groups of organisms is rare and indicates that the function and structure of the gene product is both important to the organism and is a well-adapted protein.

The DWNN homologues have been isolated in human (RBBP6) (Sakai *et al.*, 1995), mouse (PACT) (Simons *et al.*, 1997) and in yeast (Mpe1) (Vo *et al.*, 2001) (see section 1.5.2). This chapter describes the identification of the DWNN protein in *Drosophila melanogaster.* This multicellular organism provides an environment that is easy to manipulate in which to investigate the biological function of DWNN. The *D. melanogaster* DWNN protein has been named SNAMA. The word SNAMA is from the Xhosa language meaning "something that sticks like glue". A pair sequence alignment (Pearson *et al.*, 1997) of proteins containing the DWNN domain was completed (Table 7). This supported the alignments that show that the DWNN proteins are highly conserved through diverse species. These proteins are homologous to SNAMA. The yeast homologue Mpe1 is involved in 3′ end mRNA processing while the mammalian forms have been shown to interact with p53 and Rb. None of the identified homologues of SNAMA have as yet been fully characterised and thus this study will characterise the *D. melanogaster* homologue. The highest level of conservation between the proteins is within the first 76 amino acids, the DWNN domain region. The homologues lack some of the domains associated with SNAMA.

Table 7: A species comparison of the DWNN sequences. DWNN sequences from the indicated species were compared in pairs using the Align query at the Genestream Resource Centre and the percentage identities between them were recorded.

	D.melanogaster	*A.thaliana*	*H.sapien*	*M.musculus*	*S.pombe*	*O.sativa*	*S.cerevisiae*	*C.elegens*	*A.gambiae*
D.melanogaster	---	51.3	64.9	64.9	53.2	48.7	48.7	46.8	63.6
A.thaliana	51.3	---	46.8	46.8	48.1	64.1	41.0	33.8	42.1
H.sapien	64.9	46.8	---	98.7	44.3	44.9	42.3	51.3	51.9
M.musculus	64.9	46.8	98.7	---	44.3	44.9	42.3	51.3	51.9
S.pombe	53.2	48.1	44.3	44.3	---	43.0	64.6	39.2	41.5
O.sativa	48.7	64.1	44.9	44.9	43.0	---	41.0	28.2	41.0
S.cerevisiae	48.7	41.0	42.3	42.3	64.6	41.0	---	37.2	35.9
C.elegens	46.8	33.8	51.3	51.3	39.2	28.2	37.2	---	37.2
A.gambiae	63.6	42.1	51.9	51.9	41.5	41.0	35.9	37.2	---

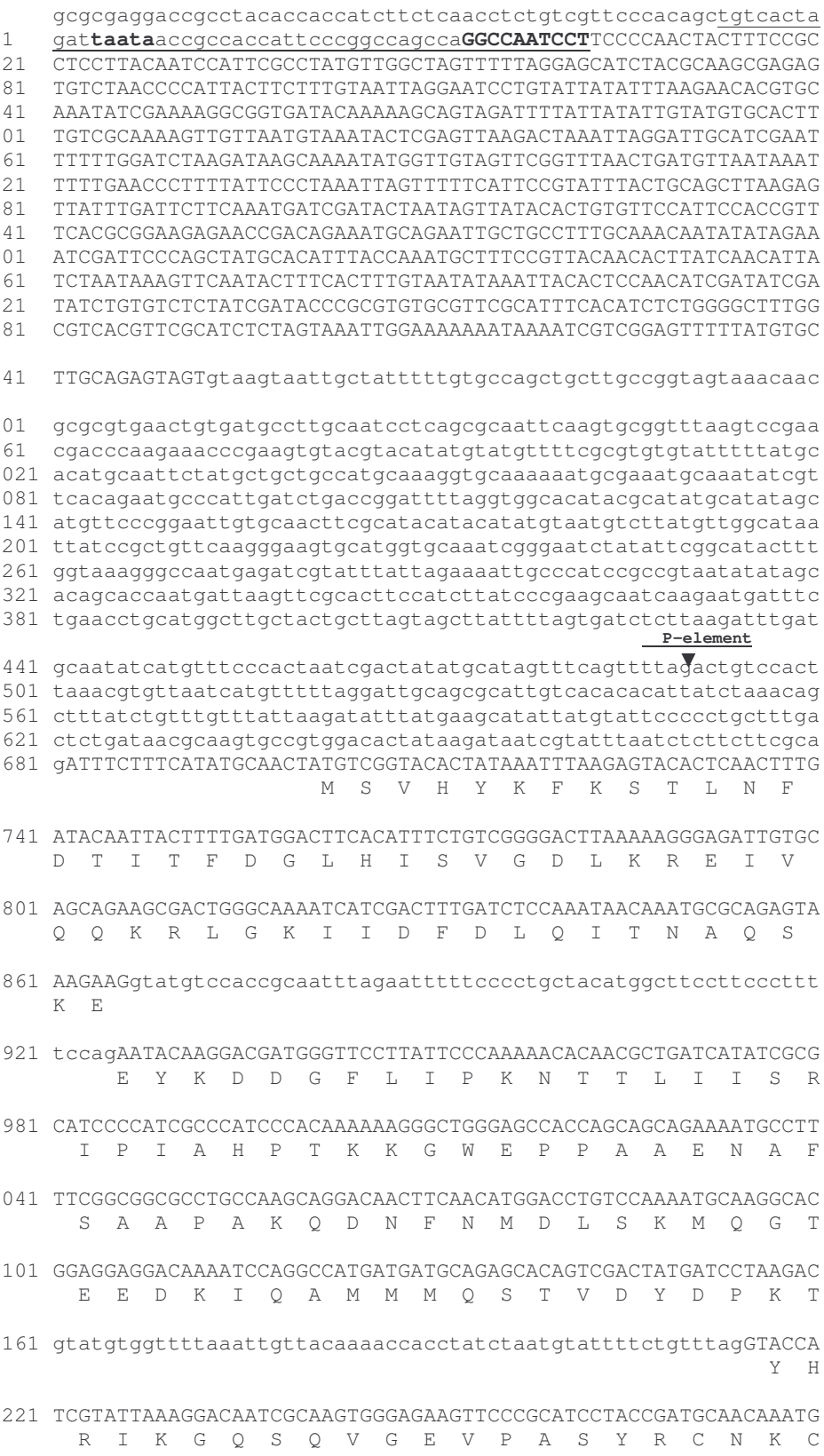
Both the human and the *D. melanogaster* proteins have a high percentage identity (64.9%). High levels of protein sequence similarity between different organisms suggest a common structure. Structural information about homologous domains will be used in molecular modelling techniques, which may give insight into the function of these proteins. This together with experimental data will elucidate the function of this recently identified protein.

3.2 SEQUENCE ANALYSIS OF THE *SNAMA* GENE

A search of FlyBase using the mammalian short DWNN gene revealed a perfectly matched gene CG3231 (FlyBase ID FBgn0027522) and a genomic scaffold (GenBank accession number AE003463.2) in *Drosophila melanogaster*. The gene is localised at 60B9 on chromosome 2R (Stapleton *et al.*, 2002a). The protein and DNA sequence for SNAMA corresponds to a gene with GenBank accession number AF132177. The *SNAMA* sequence was obtained from the 1999 EST (Expressed Sequence Tag) project of BDGP (Rubin *et al.*, 2000; Stapleton *et al.*, 2002a; Stapleton *et al.*, 2002b) EST clone BcDNA: LD21643.

The sequence was examined visually and aligned onto the genomic scaffold (Figure 8). The transcriptional unit of SNAMA consists of an open reading frame of 1231 amino acids with a predicted molecular mass of 139 kDa. Although the predicted translational start at position 1702 bp of the genomic sequence is not in the best context of *D. melanogaster* consensus sequence C/AAA/CATG (Cavener, 1987),

there is no other upstream methionine codon and there is an upstream in frame stop codon (Mather *et al.*, 2005). The translational initiation site identified for *SNAMA* is CAACTATG. This highly basic protein (pI 10.16) contains the 76 amino acids conserved DWNN domain. The overall putative protein is highly charged with 27.4% basic residues and 15% acidic residues. The C-terminal lysine rich region by itself has more than 59% basic residues making it unusually charged.


P-element

```
2281 CAAGAAAAGCGGACACTGGATCAAAAACTGTCCCTTTGTGGGGGGAAAGGACCAGCAAGA
       K  K  S  G  H  W  I  K  N  C  P  F  V  G  G  K  D  Q  Q  E

2341 GGTCAAACGGAATACTGGTATTCCGCGGTCTTTCCGCGACAAGCCAGATGCGGCTGAGAA
       V  K  R  N  T  G  I  P  R  S  F  R  D  K  P  D  A  A  E  N

2401 CGAATCAGCCGATTTTGTGCTGCCTGCTGTACAAAACCAAGAGATACCGGAGGATCTGAT
       E  S  A  D  F  V  L  P  A  V  Q  N  Q  E  I  P  E  D  L  I

2461 ATGCGGCATATGCCGAGATATATTCGTCGATGCTGTCATGATACCCTGCTGCGGAAGTTC
       C  G  I  C  R  D  I  F  V  D  A  V  M  I  P  C  C  G  S  S

2521 CTTTTGTGACGACTGTGTGCGAACCTCCTTATTGGAGTCAGAGGATAGTGAGTGCCCCGA
       F  C  D  D  C  V  R  T  S  L  L  E  S  E  D  S  E  C  P  D

2581 TTGCAAGGAGAAGAACTGTTCGCCTGGCTCCCTGATACCTAATCGGTTCTTGAGGAATTC
       C  K  E  K  N  C  S  P  G  S  L  I  P  N  R  F  L  R  N  S

2641 GGTGAACGCCTTTAAAAATGAGACTGGGTATAACAAAAGCGCGGCTAAGCCAGgtaagtt
       V  N  A  F  K  N  E  T  G  Y  N  K  S  A  A  K  P

2701 cattatactcgtttgcatcgcttaatgcaaatgaatcgaatttctcttttagCTGCAGT
                                                          A  A  V

2761 AAAAAATGAGGAAAAACCTCCTGTTGAAAAAGAAGTGGAGAAAAAGCCAGTCGCGGAGGT
       K  N  E  E  K  P  P  V  E  K  E  V  E  K  K  P  V  A  E  V

2821 GGAACCCGAAGAGACTGAGGTGAAACCTGAAAAGCAAAAAGAATCCGAAACCAATGGCAG
       E  P  E  E  T  E  V  K  P  E  K  Q  K  E  S  E  T  N  G  S

2881 TAATCCGCCAAAATCGGAATCTCCAGAGCCTCCCGCAACCACAGAACCATCACAGAAGGA
       N  P  P  K  S  E  S  P  E  P  P  A  T  T  E  P  S  Q  K  E

2941 GAAAGATAAATATGATTCAGACTACGAGGATAACATTACCATAAAAATGCCCCAGCCTGC
       K  D  K  Y  D  S  D  Y  E  D  N  I  T  I  K  M  P  Q  P  A

3001 AGCTGATTCTACAACAGTGCCCAGCAAAgtgagacatacatatttgacacatctgcaaca
       A  D  S  T  T  V  P  S  K

3061 tagatacataattattgccttaatgtttcagAGATCCCCCAGTTATTCCCACAGAAGTGA
                                     R  S  P  S  Y  S  H  R  S  E

3121 ATCCTCTCATCGACGGGACAGGTCGGATTATGTTTCCGATCACGATCACAAGCACCAACG
       S  S  H  R  R  D  R  S  D  Y  V  S  D  H  D  H  K  H  Q  R

3181 TCCATCAAAATCGGAGTCTGTTAACAAGGATCGCAGTCTCCTGCCCTTGCCCATTGGCAC
       P  S  K  S  E  S  V  N  K  D  R  S  L  L  P  L  P  I  G  T

3241 CCTGCCTAGCTACCAGGGCCACATGATGGCCGAATCAGAAGAAGCTCGTCGATCGAGTGC
       L  P  S  Y  Q  G  H  M  M  A  E  S  E  E  A  R  R  S  S  A

3301 CTATAAGCCCCCTTATATGCAAATGCAGCGGGGCCCACCTCCTATGCACATGATGAGTCA
       Y  K  P  P  Y  M  Q  M  Q  R  G  P  P  P  M  H  M  M  S  H

3361 CCACATGCCAGCCTACAACAACGGGTTTAACAACATGGGACAGAGGCCTCCCCTCAGgta
       H  M  P  A  Y  N  N  G  F  N  N  M  G  Q  R  P  P  L  S

3421 ggtttcttttccacccgtccggatcccgctaccatgtggatcgagtcaagtgtcgaggcc
3481 actggaactccgaggtgttgacgagtccatacgcagccggtcctgcctgatctcaatatt
```

```
3541 tttaagttccaacagagactatacacacacactagttaagagtaatttaagctctgggct
3601 ctgcgcctttatgctttgttgaatattggttattcctaattaaaacagtattgcttacaa
3661 tcaacgttaagctttttgaattgtaattaacttctaaaatttcagCTATGTGCCGTATCA
                                                    Y  V  P  Y  Q

3721 AAACCAATCCGTACACCCAATGCGTGCGCCGTACGGATCTGCAGGCGGAGGTATGAATAT
       N  Q  S  V  H  P  M  R  A  P  Y  G  S  A  G  G  G  M  N  M

3781 GAATATGTCACAACCATTTCAGTCCCCAAATTTAGCCTCGATATACCAAGGGGTGGCAGC
       N  M  S  Q  P  F  Q  S  P  N  L  A  S  I  Y  Q  G  V  A  A

3841 GAAGGTCGGTTCCGGgtgagtatctttagggcgttgtagaagcgtgggatgacaagtaat
       K  V  G  S  G

3901 cccaatgttttgcagTCCCATTGACGATCCGTTGGAGGCCTTCAATCGCATCATGAAGGA
                       P  I  D  D  P  L  E  A  F  N  R  I  M  K  E

3961 GAAGGAGCGGAAGAAGGTGGACCGCTTTCGAAGCTCTGACCGCCACAGGTCAAGGTCCCC
       K  E  R  K  K  V  D  R  F  R  S  S  D  R  H  R  S  R  S  P

4021 GGATAGACAAAGGCACCGCTTTAAGTCTCCCATGTACGAAAAGGACAACTCCAGGGATAA
       D  R  Q  R  H  R  F  K  S  P  M  Y  E  K  D  N  S  R  D  N

4081 TCTCAAGGACAAAAGACCGCGATCCCGGGAAAGGAAGCGAGAACATAGCTACGAACGGCA
       L  K  D  K  R  P  R  S  R  E  R  K  R  E  H  S  Y  E  R  H

4141 TATACGCCACCCTCGTTCTAGTCGCCAGCCGAATGATGGCTCTAAGTCCCCAGGTGGCAG
       I  R  H  P  R  S  S  R  Q  P  N  D  G  S  K  S  P  G  G  R

4201 AATCAAAAGgtagttttaccaccagccctttttccatgtgcttacattcatggtctactct
       I  K  R

4261 ctgctttttagATCTGGACATCGTCGCTCTGCATCTCCAAAGCCGGGCTACAAGAGTGAT
                   S  G  H  R  R  S  A  S  P  K  P  G  Y  K  S  D

4321 TACAGAGACAAGCCGTACAACAAGCCTAGTGCTCCCAAAACGGAGGCAGTTGAGCCTCCT
      Y  R  D  K  P  Y  N  K  P  S  A  P  K  T  E  A  V  E  P  P

4381 CCCCCCGGATTCGAGCCGTTGCAGCTGACGGATGAAGACGGCTACAGGAACAAGCACCCG
      P  P  G  F  E  P  L  Q  L  T  D  E  D  G  Y  R  N  K  H  P

4441 ACCAGTTCGGAAGCATCACAAAGCAGCAAGGGTGATAGCAGCAAGAAGAGAGGGGAAAAC
      T  S  S  E  A  S  Q  S  S  K  G  D  S  S  K  K  R  G  E  N

4501 AGGCACGAAGAGGCGCCACGAAAGAGGCACAGGTCTCGCAGCATTAGCAAGGAACCGAAG
      R  H  E  E  A  P  R  K  R  H  R  S  R  S  I  S  K  E  P  K

4561 CCGAATGACAGCAACTACAGGAGCCTGACTCCACCAGCAAAGATCACCACACCGAAAATG
      P  N  D  S  N  Y  R  S  L  T  P  P  A  K  I  T  T  P  K  M

4621 ACTGCTGCCCAGTTGAGGCAACGCGAAAGTTCACCGAAGACGCCGGAAAAGAGTCACGAC
      T  A  A  Q  L  R  Q  R  E  S  S  P  K  T  P  E  K  S  H  D

4681 GATTATCTGACCGCGAAGGCCAGAATTATGGCCTCCCAGCCCGTCATCAACGACACGGAA
      D  Y  L  T  A  K  A  R  I  M  A  S  Q  P  V  I  N  D  T  E

4741 ATGGAGACCAATGTGGGCAAGGAGAACAAGGCCAAGAGTCCGTTGTCAAAAGATCGCAAG
      M  E  T  N  V  G  K  E  N  K  A  K  S  P  L  S  K  D  R  K
```

```
4801 AAGAAGAAGAAGGACAAGGACAAGGCTGAGCGCAAGAAAAACAAGAAGGACAAGCGCGCT
      K  K  K  K  D  K  D  K  A  E  R  K  K  N  K  K  D  K  R  A

4861 AAGAAGGAGAAAGGGGATCGCCAGAAGAAGAGCTCCTCAGTTAATCGATCTGACTCGGAT
      K  K  E  K  G  D  R  Q  K  K  S  S  S  V  N  R  S  D  S  D

4921 ATTAACAACAGCTCACTAATGAACGAGTCAAATTATAAAGTATTGTCTCCCAGGGCTCAA
      I  N  N  S  S  L  M  N  E  S  N  Y  K  V  L  S  P  R  A  Q

4981 AGTCCCAGCATTGAGATCAATGCGGCTCAACTTTCCCCTACTCACAACGCTACTGAAAAC
      S  P  S  I  E  I  N  A  A  Q  L  S  P  T  H  N  A  T  E  N

5041 GTTAATCCGAAGAGTCATTCCATCCTTACTGTGGGTGCTGCTAGCGACGATAATCTTGGC
      V  N  P  K  S  H  S  I  L  T  V  G  A  A  S  D  D  N  L  G

5101 CCAAGAAGCAAACTCAGCGAGGCTAATTCTGTCAATCTATCCAAATGGGAAATCGACGAG
      P  R  S  K  L  S  E  A  N  S  V  N  L  S  K  W  E  I  D  E

5161 AATATCTTAGGTTTGGAGGATTCCTCCAAAAAAGCTGCCGGGGCCTCCGACGATCCGTCG
      N  I  L  G  L  E  D  S  S  K  K  A  A  G  A  S  D  D  P  S

5221 GAAATAACTTCAGACGTCCTGCGAAAGGCTGAGAACGCAATATTTGCAAAGGCTATTAAT
      E  I  T  S  D  V  L  R  K  A  E  N  A  I  K  A  K  A  I  N

5281 GCCATCAGGCCTATGGAGTTTCAAGTTATTATCAATTCCAAGGACAACAGCAAGGACCGC
      A  I  R  P  M  E  F  Q  V  I  I  N  S  K  D  N  S  K  D  R

5341 TCCGTAGTTCGAAGTGACAAGGATCGCTCCTCCTCACCCAGGCGTAACAACAGCAGCAGG
      S  V  V  R  S  D  K  D  R  S  S  S  P  R  R  N  N  S  S  R

5401 TCGGTAAAGGATAGGCTGGGCACCAAGATTTCCAATGATAGAAGCCGTTCGCGAGACAAG
      S  V  K  D  R  L  G  T  K  I  S  N  D  R  S  R  S  R  D  K

5461 TCGAAGGGCAGGCGCCGGGCCGCAAGGAGCTCCGACGACGATGCGAACCGCGGCAGGTCG
      S  K  G  R  R  R  A  A  R  S  S  D  D  D  A  N  R  G  R  S

5521 GATCGTCATGGCAGCCGGAAGAGGGACAACAGATCCCGCGACAGGGCGGCGCCTTCAGAG
      D  R  H  G  S  R  K  R  D  N  R  S  R  D  R  A  A  P  S  E

5581 AAGAGGCAGGAGCGTTCGTACAAGCGAAGCTCGCCGGAGGACGACAAGCTGAGGCGCCAG
      K  R  Q  E  R  S  Y  K  R  S  S  P  E  D  D  K  L  R  R  Q

5641 AACAAGGAACAGTCCGAATCCAAGCACGGAAAGCATGATCAAAACAATAGCGACGACTCG
      N  K  E  Q  S  E  S  K  H  G  K  H  D  Q  N  N  S  D  D  S

5701 GATCGGAGGGCGGCCAAAAACACCAAGTCCAGCGACAGCCGAGTGGTCTCCTCTGTAACA
      D  R  R  A  A  K  N  T  K  S  S  D  S  R  V  V  S  S  V  T

5761 GCCGTGGTTGCTCCTCCCAAACCCTGTCGTCCAGACAACCCGTTCCGCAAGTTCGTCGAC
      A  V  V  A  P  P  K  P  C  R  P  D  N  P  F  R  K  F  V  D

5821 ACCAGTTCGTCGAGCAGCTTAGTTGTAAAATATGATAACACGATACAGAAGGAGGGCGCG
      T  S  S  S  S  S  L  V  V  K  Y  D  N  T  I  Q  K  E  G  A

5881 TCCTCGGACAACGGCATGGAGCACAGGAAGCAGAGGGATAAGAAGCTGAAGAAACATTCA
      S  S  D  N  G  M  E  H  R  K  Q  R  D  K  K  L  K  K  H  S

5941 AAATATTCGTCAACCGATTCGTTGAAGAGCGAGAAGCGCAAGGATCCGAAGAGCAAAAAG
      K  Y  S  S  T  D  S  L  K  S  E  K  R  K  D  P  K  S  K  K
```

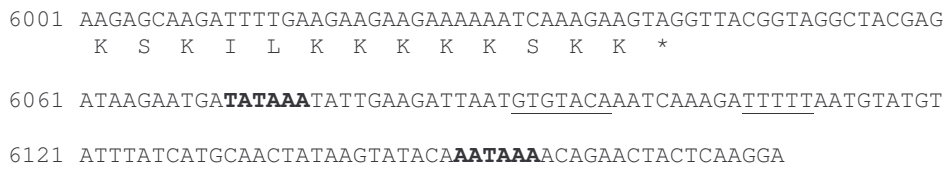


Figure 8: The sequence of the 4.595 kbp *SNAMA* gene. The amino acid sequence of the region that encodes SNAMA is also shown. Introns are indicated by lowercase letters. The potential polyadenylation sites are indicated by a bold typeface and uppercase letters. The predicted putative promoter is underlined indicating the transcriptional start site in large font. The position of the P-element insertion in *l(2)rQ313* is also indicated.

3.2.1 The 5′ untranslated region

The 5′ untranslated region (5′ UTR) has not been verified experimentally, but a putative promoter sequence and a transcriptional start site has been predicted with a score of 0.97 at position 1608 bp upstream from the translational start site using the BDGP Neural Network Promoter Prediction programme located on the web at http://www.fruitfly.org/seq_tools/promoter.html (Reese, 2001). The programme identifies elements such as the TATA box, GC box and CAAT box, the transcriptional start site and other common eukaryotic 5′ elements. There are no exact sequence patterns to identify promoters and the criteria applied to identify these sites are very general and so the predicted promoter site should be verified by biochemical methods. A typical promoter includes an A/T rich region called the TATA box and one or more sequence elements of 8 to12 base pairs (Maniatis *et al.*, 1987). These elements are known as the upstream promoter elements (UPEs) and include the CCAAT box (consensus GGC/TCAATCT). The UPEs act regardless of their orientation with respect to the TATA box. A UPE sequence has been identified as GGCCAATCCT within the predicted promoter region for *SNAMA* (Figure 8). This sequence has been predicted as the transcriptional start site positioned 1.6 kbp from the translational start site. The TATA box is required for the accurate transcription of the gene and the UPEs increase the rate of transcription. Most promoters contain the TATA box 25 to 30 bases from the transcriptional start site. A TAATA sequence is present 25 bp upstream the predicted transcriptional start site and could be the TATA

box. The SNAMA transcript would therefore consist of a 5′ UTR of 778 nucleotides and an mRNA sequence of 4.595 kbp.

3.2.2 The 3′ untranslated region

The 3′-end processing involves endonucleolytic cleavage of the primary transcript followed by polyadenylation of the upstream RNA fragments. The cleavage process is determined by specific sequence elements in the RNA, one upstream and one downstream of the cleavage site (Proudfoot, 1989). A review of the elements and mechanisms involved in polyadenylation is given by (Wahle and Keller, 1992). These elements are the polyadenylation signal AATAAA upstream of the cleavage site, followed by a GT-rich and a T-rich region downstream the cleavage site. The polyadenylation signal is essential for the process of attaching the poly(A) tail to mRNA. The AATAAA sequence is found 10 to 35 nucleotides upstream of the poly(A) site in most eukaryotic mRNAs (Gil and Proudfoot, 1987). An unusual polyadenylation site (TATAAA) is found in hepatitis B viruses and it has been suggested that this signal is used to control the amount of mRNA produced (Russnak and Ganem, 1990). The consensus for the GT-rich region is YGTGTTYY (where Y is a pyrimidine) and the T-rich and the GT-rich regions are found adjacent or in close proximity to each other downstream of the poly(A) site.

Two potential polyadenylation signals were identified at position 6072 (ATAAA) and 6147 (AATAAA) downstream of the *SNAMA* translation termination codon

(Figure 8). The first polyadenylation signal TATAAA is followed by a GT-rich region GTGTACA at position 6091 and a T-rich region TTTTT located at position 6107. The GT-rich region and the T-rich region are in close proximity to each other (11 nucleotides apart). All the criteria for the 3′ untranslated region (3′ UTR) has been met by the first potential polyadenylation signal suggesting that this signal is the functional one. No GT-rich region or T-rich region is observed for the second potential polyadenylation site. Additional information is needed to confirm the functional polyadenylation site. Using the first polyadenylation site as the probable signal, cleavage would take place at a position between 6077-6090 resulting in a 3′ UTR of approximately 42 bp. It should however be noted that alternative polyadenylation signals in *Drosophila melanogaster* genes have been identified which results in different sized RNA (Bernstein *et al.*, 1986) and it is plausible that both polyadenylation sites are functional in *SNAMA*. The second polyadenylation site would result in a 3′ UTR 124 bp long.

3.2.3 Splice junctions

The consensus sequence for splicing of introns in many species is the 5′ splice sequence MAG|GTRAGT (where | indicates the splice point, M indicates A or C and R is a purine) and the 3′ splice consensus sequence YAG|G or AG where Y is a pyrimidine (Mount, 1982; Mount *et al.*, 1992). In addition a conserved sequence upstream the 3′ splice site in the intron is required. This sequence is the branchpoint and is an important 3′ splice signal crucial to the formation of a lariat structure (Keller

and Noon, 1985). The consensus sequence of the *D. melanogaster* branchpoint has been identified as YTRAY (where Y is a pyrimidine and R is a purine). This sequence occurs very often as CTAAT (Mount *et al.*, 1992). The T and A at position 2 and 4 respectively are invariant, and A represents the position of attack by the 5′ splice site guanidine. There exist a degree of variability at position 3 and 5 where any of the four nucleotides can occur (Keller and Noon, 1985).

The *SNAMA* cDNA was aligned to the genomic DNA considering all of the above points. Using this criterion the *SNAMA* transcriptional unit was found to consist of 9 exons and 8 introns. All the splice junctions are consistent with the consensus sequence of the presumptive splice junctions in *D. melanogaster* (Mount *et al.*, 1992) (Table 8). All the introns are very short except for intron 1 (829 bp) and intron 6 (288 bp). The possible branchpoint sequences have been identified and are within the 3′ splice region of the intron (Table 9).

Table 8: *SNAMA* intron and exon splice junctions. Splice junctions were identified by comparing the genomic sequence (GenBank accession number AE003463.2) with the cDNA sequence (GenBank accession number AF132177).

(5′ splice junction)	Position	(3′ splice junction)	Position	Size (bp)
1. ATGTGC TTGCAGAGTAGTgtaa	852	1. tttaatctcttcttcgcagATTTCT	1681	829
2. ATGCGCAGAGTAAAGAAGgtat	1866	2. ggcttccttccctttccagAATAC	1925	59
3. CGACTATGATCCTAAGACgtatg	2160	3. ctaatgtattttctgtttagGTACC	2214	54
4. AAAGCGCGGCTAAGCCAGgtaa	2693	4. aatcgaatttctcttttagCTGCA	2753	60
5. ACAACAGTGCCCAGCAAAgtga	3028	5. attgccttaatgtttcagAGATCC	3091	63
6. ACAGAGGCCTCCCCTCAGgtag	3417	6. attaacttctaaaatttcagCTATG	3705	288
7. AGCGAAGGTCGGTTCCGGgtga	3855	7. taatcccaatgttttgcagTCCCA	3915	60
8. AGGTGGCAGAATCAAAAGgtag	4209	8. gtctactctctgctttttagATCTG	4271	61

Table 9: Introns from the *SNAMA* gene. The last 50 bp of the intron showing the putative branchpoint sequences (underlined). Distance is the number of base pairs between the end of the sequence and the 3′ splice point.

Intron Sequence	Position	Distance
1.caagtgccgtggacactataagataatcgtatttaatctcttcttcgcag	852	13
2.ccgcaatttagaatttttcccctgctacatggcttccttccctttccag	1866	39
3.gtggttttaaattgttacaaaaccacctatctaatgtattttctgtttag	2160	15
4.tatactcgtttgcatcgcttaatgcaaatgaatcgaatttctcttttag	2693	28
5.atttgacacatctgcaacatagatacataattattgccttaatgtttcag	3028	7, 43
6.tacaatcaacgttaagcttttgaattgtaattaacttctaaaatttcag	3417	25
7.tttagggcgttgtagaagcgtgggatgacaagtaatcccaatgttttgcag	3855	46
8.ccagccctttccatgtgcttacattcatggtctactctctgcttttag	4209	21

3.3 THE DOMAIN ORGANISATION OF SNAMA

The mammalian DWNN domain proteins exist in two forms, as a small protein containing the DWNN domain sequence with a short C-terminal extension and as a multi–domain protein. There are two human transcripts; a short one encoding a 118 amino acid protein consisting of the 76 amino acid residue DWNN domain with a short C-terminal extension (approximately 13kDa) called DWNN-13 and a second much larger transcript of 1792 amino acids (200 kDa) referred to as DWNN-200 (Dlamini, Z., personal communication). DWNN-13 has a glycine dipeptide and a small C-terminal extension following the conserved proline found in most of the DWNN proteins. The identification of the smaller DWNN-13 protein with the di-glycine at the C-terminal end of the protein (positions 78 and 79) suggests that the DWNN domain in humans may be transferred to other proteins and act as a molecular tag in a manner similar to the transfer of ubiquitin to proteins. The ubiquitin protein contains a glycine at position 75 and position 76, which is important for ubiquitin conjugation to other proteins and to the process of de-ubiquitination (Pickart *et al.*, 1994). It is at this position that cleavage takes place.

The multi-domain DWNN proteins are found associated with other characterised domains (Figure 9). The plant and fungal DWNN proteins contain a Zn^{2+} finger motif of the CCHC and a RING finger-like (RFL) C3HC4 type sequence. In addition, the *D. melanogaster* and *C. elegans* DWNN domain proteins also contain regions with homology to the mouse PACT protein (**P**53 **A**ssociated **C**ellular protein- **T**estes

derived) that binds to p53. The PACT protein is most closely related to the human RBBP6 (retinoblastoma binding protein 6) protein that binds to pRB. The mouse PACT was identified as a protein with the ability to bind p53 (Simons *et al.*, 1997). RBBP6 is the truncated human homologue of PACT. A closer look at these two proteins revealed that they are in fact truncated mouse and human homologues of SNAMA, respectively. Adjacent to the stop codon of SNAMA is a basic lysine-rich region also found in PACT.

SNAMA has been identified through homology searches using sequence identified from the promoter trap mutagenesis screen for apoptotic genes in CHO cells. Examination of the *SNAMA* sequence using the SMART (Simple Modular Architecture Research Tool) programme (Schultz *et al.*, 1998), Pfam protein family database (Bateman *et al.*, 2004) and ScanProsite (in the ExPASy Molecular Biology Server) (Hulo *et al.*, 2004) revealed that the protein contains a zinc finger (CCHC) between amino acids 151 and 168 and a RFL motif between amino acids 217 and 258 (Figure 10). This cysteine rich RING finger-like (RFL) domain is discussed later.

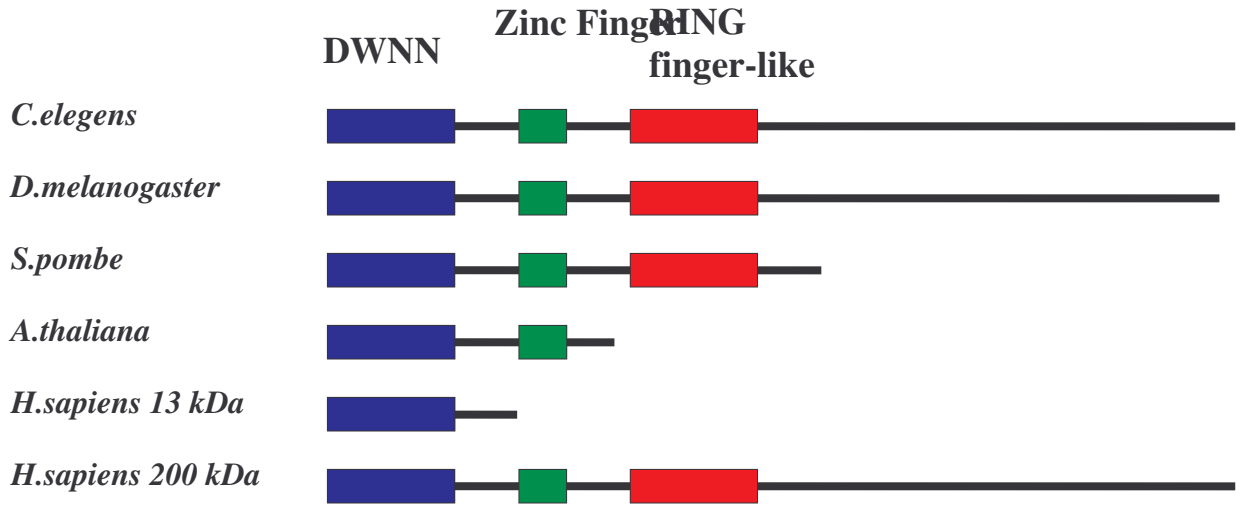


Figure 9: The domain arrangement of SNAMA homologous proteins. DWNN proteins are found associated with other characterised domains including the zinc finger domain (CCHC) and the cysteine rich RING finger-like domain. The DWNN human homologue encodes 2 overlapping proteins of 13 kDa (DWNN-13) and 200 kDa (DWNN-200). DWNN-13 is the DWNN domain with a short C-terminal extension and DWNN-200 contains the associated domains.

1 MSVHYKFKST LNFDTITFDG LHISVGDLKR EIVQQKRLGK IIDFDLQITN AQSKEEYKDD
61 GFLIPKNTTL IISRIPIAHP TKKGWEPPAA ENAFSAAPAK QDNFNMDLSK MQGTEEDKIQ
121 AMMMQSTVDY DPKTYHRIKG QSQVGEVPAS YRCNKCKKSG HWIKNCPFVG GKDQQEVKRN
181 TGIPRSFRDK PDAAENESAD FVLPAVQNQE IPEDLICGIC RDIFVDAVMI PCCGSSFCDD
241 CVRTSLLESE DSECPDCKEK NCSPGSLIPN RFLRNSVNAF KNETGYNKSA AKPAAVKNEE
301 KPPVEKEVEK KPVAEVEPEE TEVKPEKQKE SETNGSNPPK SESPEPPATT EPSQKEKDKY
361 DSDYEDNITI KMPQPAADST TVPSKRSPSY SHRSESSHRR DRSDYVSDHD HKHQRPSKSE
421 SVNKDRSLLP LPIGTLPSYQ GHMMAESEEA RRSSAYKPPY MQMQRGPPPM HMMSHHMPAY
481 NNGFNNMGQR PPLSYVPYQN QSVHPMRAPY GSAGGGMNMN MSQPFQSPNL ASIYQGVAAK
541 VGSGPIDDPL EAFNRIMKEK ERKKVDRFRS SDRHRSRSPD RQRHRFKSPM YEKDNSRDNL
601 KDKRPRSRER KREHSYERHI RHPRSSRQPN DGSKSPGGRI KRSGHRRSAS PKPGYKSDYR
661 DKPYNKPSAP KTEAVEPPPP GFEPLQLTDE DGYRNKHPTS SEASQSSKGD SSKKRGENRH
721 TEAPRKRHRS RSISKEPKPN DSNYRSLTPP AKITTPKMTA AQLRQRESSP KTPEKSHDDY
781 LTAKARIMAS QPVINDTEME TNVGKENKAK SPLSKDRKKK KKDKDKAERK KNKKDKRAKK
841 EKGDRQKKSS SVNRSDSDIN NSSLMNESNY KVLSPRAQSP SIEINAAQLS PTHNATENVN
901 PKSHSILTVG AASDDNLGPR SKLSEANSVN LSKWEIDENI LGLEDSSKKA AGASDDPSEI
961 TSDVLRKAEN AIFAKAINAI RPMEFQVIIN SKDNSKDRSV VRSDKDRSSS PRRNNSSRSV
1021 KDRLGTKISN DRSRSRDKSK GRRRAARSSD DDANRGRSDR HGSRKRDNRS RDRAAPSEKR
1081 QERSYKRSSP EDDKLRRQNK EQSESKHGKH DQNNSDDSDR RAAKNTKSSD SRVVSSVTAV
1141 VAPPKPCRPD NPFRKFVDTS SSSSLVVKYD NTIQKEGASS DNGMEHRKQR DKKLKKHSKV
1201 GSTDSLKSEK RKDPKSKKKS KLLKKKKKSK K

DWNN domain

Zinc finger CCHC motif

Cysteine rich RING finger-like (RFL) domain

Nuclear Bipartite nuclear sequence

Lysine rich sequence

Glutamic acid rich region

SR and **RS** dipeptides are indicated in bold

Figure 10: The amino acid sequence of *Drosophila melanogaster* SNAMA protein showing conserved domains and motifs. The DWNN domain and downstream sequence elements identified by the Pfam, SMART and PROSITE databases have been highlighted.

A domain near the amino terminus (amino acid 234-252) has been implicated in cell growth control and has been referred to as the cell division sequence motif (Figge and Smith, 1988). ScanProsite revealed sequences including four possible nuclear bipartite targeting sequences (residues 713-730; 818-834; 1195-1211 and 1211-1227). There are also two lysine rich regions between amino acids 805 and 848 that encode 21 lysine residues and a 44 amino acid sequence adjacent to the stop codon from amino acid 1188 that has 21 lysine residues. The peptide has a region rich in arginine-serine/serine-arginine dipeptides that have been found in proteins that are involved in RNA processing. The zinc finger motif has been implicated in the interactions between proteins and single-stranded nucleic acids (Laity *et al.*, 2001; Berg and Shi, 1996). These motifs indicate that the SNAMA protein may play key roles in cell cycle or transcriptional regulation events.

Database searches also revealed regions with homology to the human RBQ-1 (RBBP6) (retinoblastoma binding protein 6) proteins (Sakai *et al.*, 1995), mouse PACT (P53 Associated Cellular protein-Testes derived) (Simons *et al.*, 1997), mouse P2P-R (proliferation potential protein-related) (Witte and Scott, 1997) and the yeast Mpe1 (Vo *et al.*, 2001). These four proteins are in fact the human, mouse, and yeast homologues of SNAMA, respectively and have been previously discussed in section 1.5.2. The RBQ1, first identified in humans, was a truncated version of P2P-R. RBQ1 was truncated as a result of a possible mutational event occurring in the cell line in which this protein was identified. The yeast homologue is involved in 3′ end mRNA processing (Vo *et al.*, 2001) while the mammalian forms interact with p53 and

Rb(Simons *et al.*, 1997). The researchers that identified these genes failed to identify and characterise the novel DWNN domain.

The aforementioned evidence indicates that SNAMAs homologues play a role in the ubiquitin pathway and in apoptosis. This together with the domain arrangement of SNAMA makes it possible to speculate that SNAMA itself may be involved in the ubiquitin pathway and in apoptosis. The diverse arrangements of domains may also suggest that this protein plays a role in transcriptional regulation or cellular proliferation. The zinc fingers may suggest the presence of a DNA binding domain. However, it is necessary to be cautious about interpreting the function of a zinc finger. Zinc fingers may be involved in the binding of RNA rather than DNA or may not have any nucleic acid binding activity. This thesis attempts to elucidate SNAMAs role in apoptosis and its ability to bind Dmp53.

3.3.1 The DWNN domain of the SNAMA protein

SNAMA contains the highly conserved N-terminal domain DWNN. This domain has a degree of identity to the ubiquitin protein although it lacks the conserved di-glycine residues at the C-terminal end that is essential for recognition by hydrolases and for modifying other proteins. The glycine is replaced by a conserved proline that is present in many members of this family of proteins (Figure 7). Proline has a major effect on the structure of a protein because of the special restrictions imposed by the ring structure of this amino acid. A pair sequence alignment between ubiquitin and

ubiquitin-like proteins using the ALIGN Query (Pearson *et al.*, 1997) on the Genestream server at Genestream Resource Centre in France indicated a 23.5% identity between DWNN and ubiquitin. This level of identity is significant because the ubiquitin-like protein BemI has a 20.7% identity to ubiquitin (Table 10). This level of identity indicates that DWNN domains may function like ubiquitin to tag proteins for transport or degradation. This idea is supported by the fact that the SNAMA protein has a RING finger domain often found having E3 ubiquitin ligase activity. The DWNN domain of SNAMA lacks the double glycine motif required for ubiquitin precursor processing and for ubiquitin conjugation. This glycine is replaced by a conserved proline, present in virtually all members of the DWNN family. Interestingly, the two human forms and the mouse orthologue have a di-glycine peptide at position 78 and 79 causing speculation that cleavage could happen here (Figure 7). Alternatively the proline at position 76 in the DWNN domain of SNAMA could indicate an entirely new enzymology pathway if in fact DWNN is used as a molecular tag for other proteins. Alternatively, SNAMA could be an ubiquitin domain protein (UDP). UDPs have ubiquitin-like domains within a large protein. This ubiquitin-like domain does not always end in a Gly-Gly sequence but shows sequence homology similar to ubiquitin (section 1.11.2). The list of UDPs is growing and includes proteins such as HUB (Luders *et al*., 2003), UIP28 (Martinez-Noel *et al*., 1999), Scythe (Thress *et al*., 1998), Parkin (Zhang *et al*., 2000), Rad23 (Watkins *et al*., 1993), Dsk2 (Biggins *et al*., 1996) and Bag-1 (Luders *et al*., 2000).

Table 10: Sequence relationships among ubiquitin-like proteins. Sequences were compared in pairs using the ALIGN Query on the Genestream server. The percentage amino acid sequence identity between each pair is recorded on the table. The DWNN domain of SNAMA has a 23.5% identity to the human ubiquitin protein. This percentage is substantial since the ubiquitin like protein Bem1 has a 20.7% identity to ubiquitin.

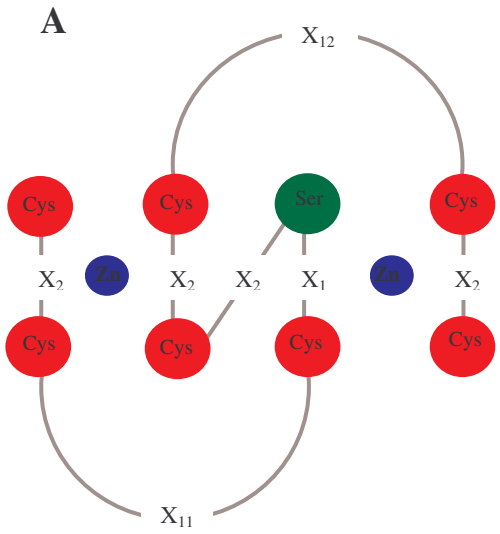
	Ubiquitin	Rub1	Nedd8	Ivcba2	DWNN	Bem1	Bem1 yeast	Mpe1p
Ubiquitin	---	63.2	57.9	26.0	23.5	20.7	14.3	17.8
Rub1	63.2	---	85.5	28.2	19.5	16.0	13.6	16.7
Nedd8	57.9	85.5	---	23.4	19.5	11.1	16.5	14.4
Ivcba2	26.0	28.2	23.4	---	22.9	18.2	13.6	24.2
SNAMA	23.5	19.5	19.5	22.9	---	22.6	12.5	42.2
Bem1	20.7	16.0	11.1	18.2	22.6	---	11.1	21.3
Bem1 yeast	14.3	13.6	16.5	13.6	12.5	11.1	---	15.4
Mpe1P	17.8	16.7	14.4	24.2	42.2	21.3	15.4	---

3.3.2 The cysteine rich RING finger-like domain

A cysteine rich domain follows the zinc finger motif. Although this motif has been identified by SMART and ScanProsite as a RING finger, a closer look reveals that this sequence does not have a histidine in the fourth position that is characteristic of RING finger motifs. This histidine has been substituted by a serine residue. RING finger motifs are known to have ubiquitin ligase activity. The association of this RING finger like domain with a ubiquitin-like domain could indicate that SNAMA functions in modifying other proteins or is responsible for catalysing its own auto-ubiquitination. This is only true if we speculate that the RING finger fold of SNAMA will coordinate two zinc atoms per RING molecule with each atom ligated tetrahedrally by four cysteine residues and three cysteine molecules and a serine in the typical 'cross-brace' system (Figure 11 (A)) (Borden and Freemont, 1996). The alignment of orthologous RING finger-like domains shows that this position is occupied by various other polar amino acids such as lysine in *A. thaliana* and *S. pombe*, asparagine in *H. sapiens, C .elegans* and *M. musculus*, and arginine in *O. sativa* (Figure 11 (B)). Interestingly, the histidine residue is retained in *E. cuniculi.* It has, however, not been shown what effect substitution at this position would have on the ability of the domain to coordinate zinc and maintain protein ligase activity. It is important to note, however, that the RING finger motif has been found to adopt subtle changes in key zinc ligands without losing the overall conformation of the motif. Examples include the retinoblastoma-binding protein RBQ-1 (RBBP6) (where His1 is replaced by Asn residue), the mouse double minute protein Mdm2 (where a

Thr residue replaces Cys3) (Boddy *et al.*, 1994) and CART1 (where Cys7 is substituted for an Asp residue) (Regnier *et al.*, 1995). In all these RING finger variants the substituted amino acids are able to participate in zinc ligation and thus the three dimensional structure of the domain is conserved (Saurin *et al.*, 1996). In Mdm2 the substituted amino acid maintains ubiquitin ligase activity (Fang *et al.*, 2000). The alignment of the RING domains of proteins that interact with the ubiquitin-conjugating enzyme UbcM4 identified the UIP4 protein as having the His replaced by a Cys (Martinez-Noel *et al.*, 1999). All the other UbcM4 interacting proteins maintain the typical RING consensus sequence. It is also important to note that the affinity for zinc by the RING domain is higher for site 1 than for site 2 (Borden, 2000). There exists a degree of flexibility in site 2 for zinc binding. It has been shown that in cases of lowered zinc concentration, it is likely that only site 1 is folded and able to associate with its protein partners (Kentsis and Borden, 2000). The observed substitutions only occur in site 2 on the RING where this flexibility is possible. The conserved architecture of site 1 in the RING domain of SNAMA suggests that the domain is able to associate with its protein partners even though site 2 has a substitution. The conservative replacement of histidine with serine in the RING domain of SNAMA is capable of ligating zinc due to ionisation of the hydroxyl group on the serine residue similar to the chemistry with threonine.

A



B

```
A. gambia     CGICKDLFTDAVMIP.CCGSSFCDECVRTALLESEDNECPD..C      41
H. sapiens    CLICKDIMTDAVVIP.CCGNSYCDECIRTALLESDEHTCPT..C      41
C. elegans    CPICQSLFKEAIVTS.CCGNSYCADCIEARILDPDNQKCPGADC      43
A. thaliana   CPLCKEVMKDAALTSKCCYKSFCDKCIRDHIISKSM..CV...C      39
S. pombe      CTLCKKLARNACRTP.CCDKLFCEECIQTALLDSDFE.CPN..C      40
O. sativa     CPLCKEVMKDAVLTSKCCFRSFCDKCIRDYIINKSM..CV...C      39
M. musculus   CLICKDIMTDAVVIP.CCGNSYCDECIRTALLESDEHTCPT..C      41
E. cuniculi   CPSCHCLLINPMKTS.CG.HTLCDQCIGMDKK......CYV..C      34
D. melanogasCGICRDIFVDAVMIP.CCGSSFCDDCVRTSLLESEDSECPD..C      41
```

Figure 11: **(A)** Schematic representation of the "cross-brace" structure adopted by the unique RING finger-like domain of SNAMA (Mather *et al*., 2005). This scheme is identical to the arrangement adopted by the typical RING finger with histidine in the serine position for the SNAMA sequence. Red spheres represent cysteine residues, green spheres represent serine residues and the blue spheres correspond to zinc atoms. **(B)** Alignment (DNAMAN) of the cysteine rich RING finger-like domain conserved in SNAMA homologues. The *E. cuniculi* protein has a true RING finger domain. D. melanogas is *Drosophila melanogaster.*

Whether or not the SNAMA cysteine rich domain has ubiquitin ligase activity is unknown. It has, however, been shown that this domain exists and is conserved in a significant group of proteins and may represent a distinct class of RING finger-like proteins that have as yet not been characterised.

3.3.3 The ubiquitin-like fold of the DWNN domain of SNAMA

In order to establish the ubiquitin-like identity and secondary structure the N-terminal 76 amino acid DWNN domain of SNAMA was analysed using FUGUE v2.s.07 (Shi *et al.*, 2001) (Figure 12). FUGUE is a programme for recognizing distant homologues by sequence-structure comparison. The programme utilises environment-specific substitution tables and structure-dependent gap penalties where scores for amino acid matching and insertions/deletions are evaluated depending on the local environment of each amino acid residue in a known structure. Given a query sequence (or a sequence alignment), FUGUE scans a database of structural profiles, calculates the sequence-structure compatibility scores and produces a list of potential homologues and alignments. The fold library and substitution tables are based on a HOMSTRAD (**HOM**ologous **STR**ucture **A**lignment **D**atabase) database (Mizuguchi *et al.*, 1998). This is a curated database of structure-based alignments for homologous protein families. All known protein structures are clustered into homologous families (i.e., common ancestry) and the sequences of representative members of each family are aligned on the basis of their 3D structures using the programmes MNYFIT, STAMP and COMPARER. The alignment generated by FUGUE included the PSI-BLAST

homologues of SNAMA aligned against all the representative structures from the HOMSTRAD family. This search produced a number of similar proteins three of which had remarkable secondary structure similarities with the DWNN domain of SNAMA. The predicted structure has a z-score of 3.71 indicating a 90% level of confidence for the prediction. SNAMA was aligned with ubiquitin and ubiquitin-like proteins, Nedd8 and Rub1. An important observation is that all the other proteins scored were ubiquitin-like proteins. Homologues were identified from a range of organisms. These proteins included the yeast Mpe1 protein and human RBBP6. Another secondary structure prediction method used was PSIPRED (Jones, 1999). The PSIPRED server (McGuffin *et al.*, 2000) allows submission of a protein sequence, performs a prediction and outputs the result has an animated figure (Figure 13). The predicted secondary structure was the same as that predicted by FUGUE (see Figure 12). We also compared the secondary structure of ubiquitin to the DWNN domain of SNAMA using the Network Protein Sequence Analysis tool (NPSA) (Combet *et al.*, 2000). NPSA uses ten different statistical prediction schemes applied to the sequence. All of the schemes agree on the approximate locations of the alpha helices (h) and beta strands (e), but they disagree considerably on the lengths and end positions of these segments. The probable position of loops (c) and turns (t) are inconsistently predicted. The sequence conservation suggests a similar arrangement of four strands and a helix (Figure 14).

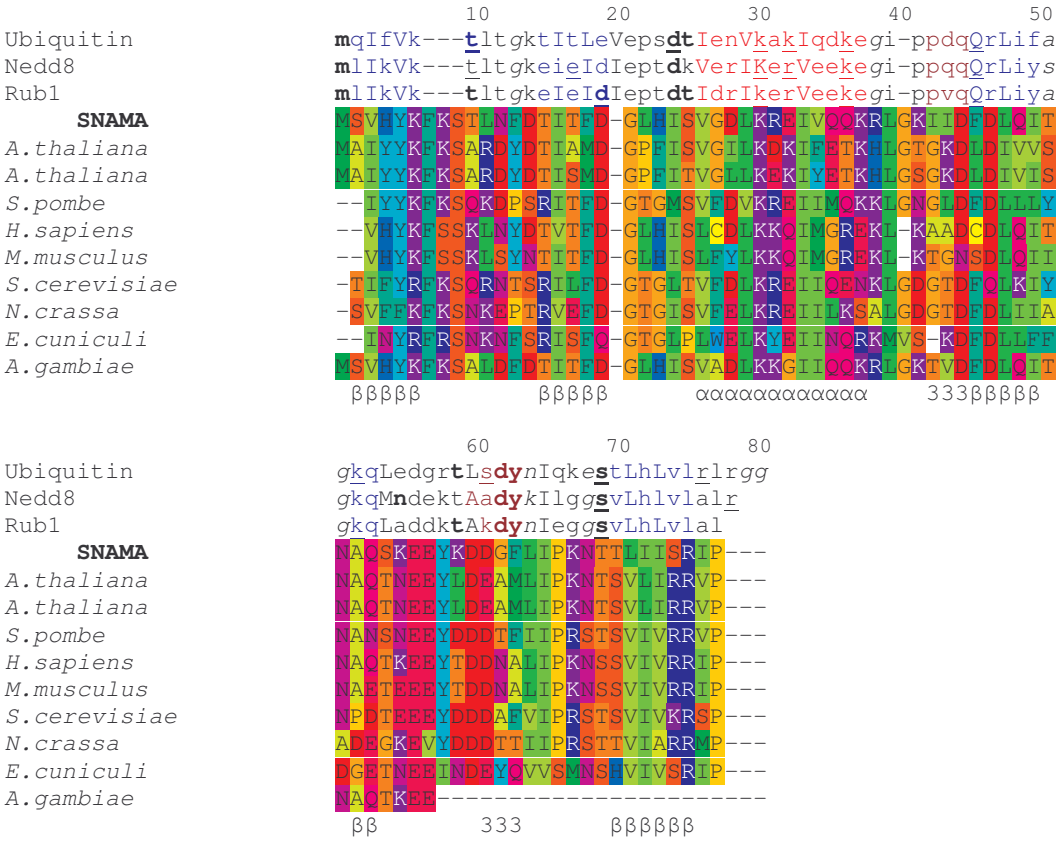


Figure 12: The alignment of SNAMA (Mather *et al.,* 2005) with ubiquitin and ubiquitin-like sequences, Nedd8 and Rub1, generated using FUGUE (Shi *et al.,* 2001). This algorithm searches sequences against a fold library using environmental-specific substitution tables and structure-dependent gap penalties. The alignment also shows PSI-BLAST homologues from a range of organisms (shaded in colour) and secondary structure prediction (β indicates β-strands, α indicates α-helices and 3 indicates 3_{10} helix). Four β-strands and a single α-helix were predicted. This secondary structure prediction is similar to that of ubiquitin.

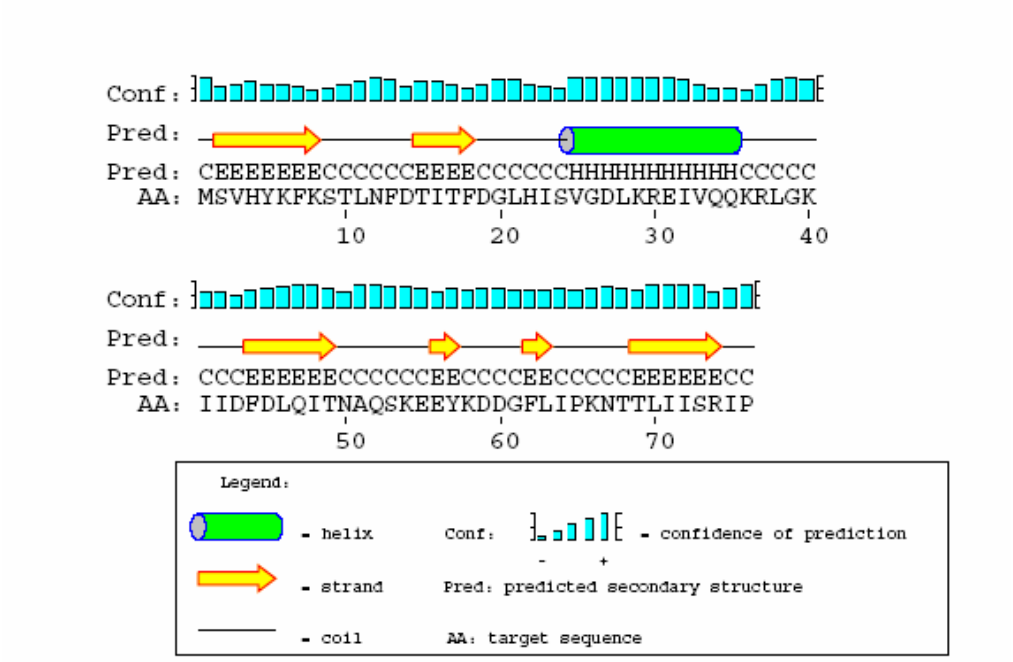
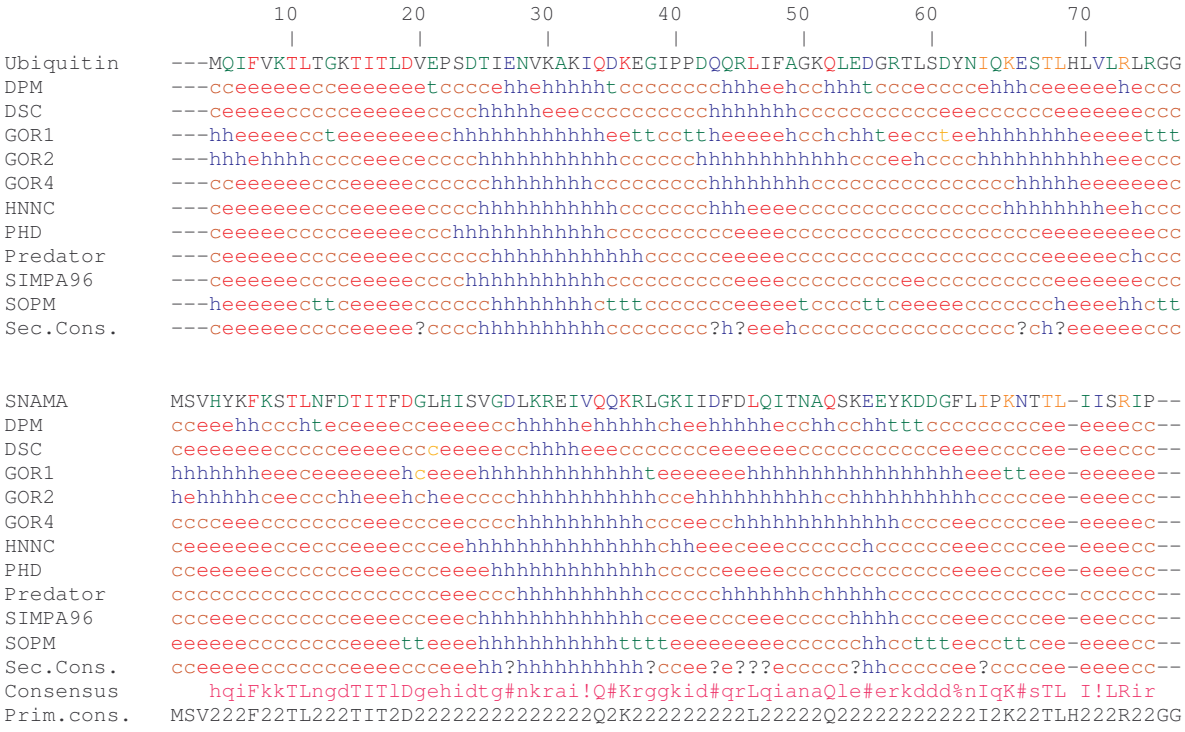


Figure 13: Secondary structure prediction of the DWNN domain of SNAMA using PSIPRED. The secondary structure compared well with the DWNN domain structure predicted by FUGUE. PSIPRED predicted four β-strands and a single α-helix similar to that of ubiquitin.


10 20 30 40 50 60 70
Ubiquitin ---MQIFVKTLTGKTITLDVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQKESTLHLVLRLRGG
DPM ---ccccccccccccccccccchhhh

Taken together these results indicate that the DWNN domain of SNAMA adopts a fold similar to that of ubiquitin. The fold is characterised by two α-helices and four β-strands ordered as β-β-α-β-α-β along the sequence. The structural fold of the DWNN domain is significantly similar to ubiquitin even though the sequence has only a 23.5% identity with ubiquitin. The predicted SNAMA structure lacks the second and smaller α-helix found in ubiquitin. This much smaller 3_{10} helix is not significant and does appear when SNAMA is modelled against ubiquitin and only disappears upon minimisation (Figure 15). Preliminary NMR structural analysis of the human DWNN-13 protein does indicate a folded structure similar to that of ubiquitin (Pugh, personal communication).

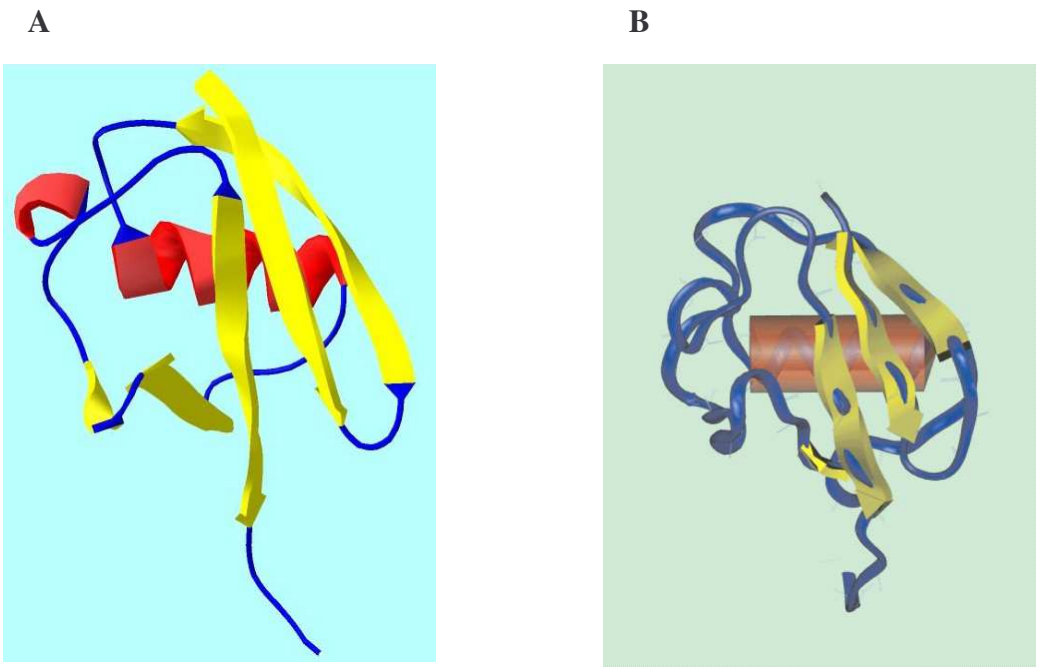


Figure 15: (A) The three-dimensional structure of ubiquitin (Cornilescu *et al.,* 1998). The figure was generated using DeepView/Swiss-PdbViewer Version 3.7. **(B)** A structural model of the DWNN domain generated by aligning the sequence with the human ubiquitin sequence (1UBI). The predicted DWNN structure was generated after a first minimisation involving first 500 steps and a second minimisation involving 500 steps of steepest descent and 500 steps conjugated gradient method using Insight II (Ntwasa, unpublished). The structure generated for DWNN compares well to the structure of ubiquitin.

CHAPTER 4

TEMPORAL AND SPATIAL DISTRIBUTION OF SNAMA IN *Drosophila melanogaster*

4.1 Introduction

4.2 *SNAMA* is a developmentally regulated gene

4.3 *SNAMA* cDNA sequence

4.4 *SNAMA* is expressed throughout development

4.5 SNAMA may be localised in the nucleus of *D. melanogaster* embryos

4.6 SNAMA interacts with Dmp53

4.1 INTRODUCTION

The localisation of SNAMA transcripts together with information about protein distribution may shed light on the translational control of SNAMA. This information can be related to gene activity at the DNA, mRNA, and protein level and may shed light on the regulation of SNAMA during embryonic development. *D. melanogaster* provides a perfect multicellular environment in which to investigate *SNAMA* localisation. The identification of SNAMA homologues by other researchers involved a functional analysis of these proteins based on their binding ability to p53 and pRb (see section 3.3). This chapter investigates the distribution of SNAMA during *D. melanogaster* development.

4.2 *SNAMA* IS A DEVELOPMENTALLY REGULATED GENE

The coupled reverse transcription and PCR amplification (RT-PCR) technique has been developed to measure gene expression in tissues and cells. Of all the available techniques, RT-PCR is the most sensitive and versatile. The technique can be used to determine the presence or absence of a transcript, to estimate expression levels and to clone cDNA products without the necessity of constructing and screening a cDNA library. To determine the levels of *SNAMA* transcription, RT-PCR was performed on total embryonic RNA at various developmental stages using primers DWNN1 and DWNN2. These primers amplify the DWNN domain of SNAMA. RNA was collected from a 0-3, 3-6, 6-12, and 12-18 hour stage of development of the embryo.

To standardise the amount of RNA amplified for each of the developmental stages, the ribosomal gene L18a (Ntwasa *et al.*, 1994) was amplified. The L18a RNA is expressed throughout the development of the embryo and the levels of L18a transcripts have been found to be the same throughout the early developmental stages (0-18hr). By comparing the levels of RNA expression of L18a in the various developmental stages we were able to establish that equal amounts of RNA were amplified for each developmental stage. The levels of SNAMA expression decreased as the embryo developed (Figure 16). Two SNAMA transcripts are clearly visible in stages 0-3hr, 3-6hr, and 6-12hr of embryonic development. The larger of these two transcripts (SNAMA-1) disappears in 12-18hr embryos. This may indicate that there are two transcripts of SNAMA in 0-12hr embryonic stages, or that the two bands are splice variants of SNAMA. Indeed, DWNN proteins do exist either as the DWNN domain by themselves or associated with other domains as observed in humans where a 13 kDa (DWNN-13) and a 200 kDa (DWNN-200) DWNN protein has been identified (Figure 9).

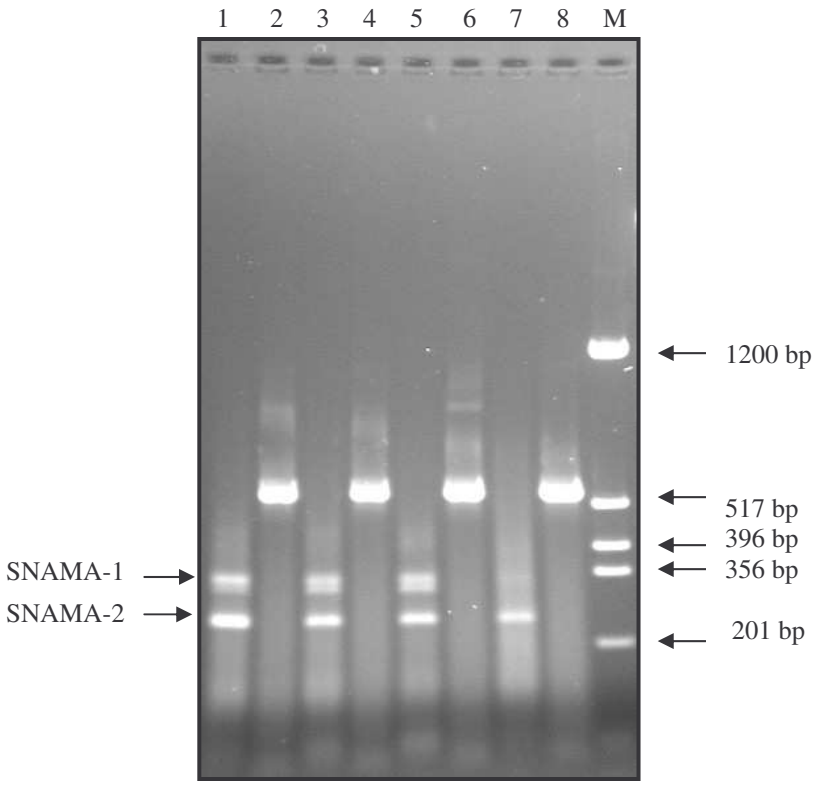


Figure 16: Reverse transcriptase PCR was performed on various developmental stages of *D. melanogaster* embryonic RNA and resolved on a 1.5% agarose gel. To standardise the amount of RNA amplified per developmental stage, a ribosomal protein L18a was amplified. Lanes 1,3,5 and 7: SNAMA transcripts and lanes 2,4,6 and 8: L18a transcripts. Lanes 1 and 2: 0-3hr embryonic development, lanes 3 and 4: 3-6hr embryonic development, lanes 5 & 6: 6-12hr embryonic development, lanes 7 and 8: 12-18hr embryonic developmental stage and M: marker *(HinfI* digested pTZ plasmid). SNAMA transcripts were detected throughout embryonic development. Two transcripts are expressed in 0-3hr, 3-6hr, and 6-12hr embryos. The level of SNAMA-2 (234 bp) decreases as development progresses. SNAMA-1 was not detected in 12-18hr embryos.

The presence of two transcripts in early embryonic development and evidence that the DWNN domain of SNAMA assumes a folded structure similar to ubiquitin (section 3.3.3) together with the presence of the cysteine rich RING finger-like domain suggest that SNAMA functions in a role as an ubiquitin-like protein. SNAMA could function similarly to ubiquitin to modify proteins by covalent attachment of this ubiquitin-like domain to other proteins. But the absence of the C-terminal di-glycine residues on SNAMA suggest that some other enzymatic mechanism is involved in this attachment of the ubiquitin-like domain, since the known ubiquitin ligase requires this di-glycine residues.

To determine whether these two amplicons are splice variants of each other, the bands containing the cDNA were excised and purified from the agarose gel as indicated in section 2.7.4. The purified DNA was sequenced using the automated sequencer at the School of Molecular and Cell Biology, University of the Witwatersrand. Alignment of the sequences of these DNA fragments with SNAMA genomic DNA indicated that SNAMA-1 included sequence that aligned to intron 1 (Figure 17). It is therefore clear that genomic DNA had been amplified in the RNA samples collected since the primer pair used flanked the region including intron 1. This would, however, not explain the additional nucleotide at position 193 resulting in a frame shift. It is likely that an additional nucleotide was introduced during the RT-PCR.

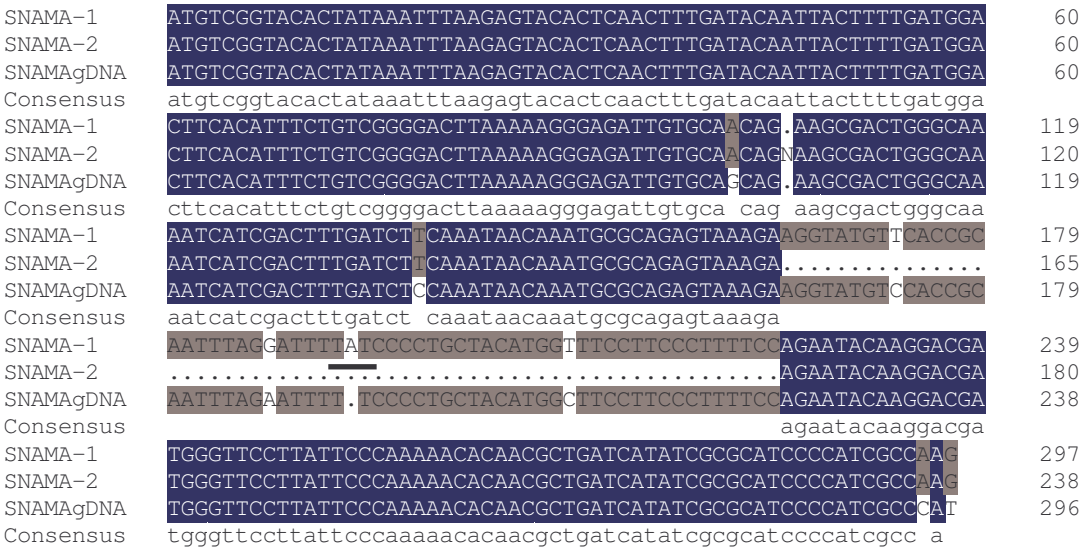


Figure 17: Sequence alignment of the two RT-PCR products SNAMA-1 and SNAMA-2 with genomic DNA sequence (SNAMAgDNA). SNAMA-1 includes sequence that aligns to intron 1. An additional nucleotide resulting in a frame shift is observed at position 193 (underlined). RT-PCR amplified genomic DNA in the RNA samples.

This result suggests that there exist a single SNAMA transcript in developing embryos. The RT-PCR results indicated *SNAMA* to be a developmentally controlled gene with levels decreasing as development progresses. *Drosophila melanogaster* cell death begins at stage 11 (~7 hr) of embryogenesis marking the beginning of the patterning of the embryo (Abrams *et al.*, 1993). Based on initial screening experiments used to identify DWNN (Pretorius, 1999), we propose that SNAMA plays a key role in the apoptosis pathway. This is consistent with evidence that CHO cells lacking the DWNN protein are resistant to staurosporine, the apoptosis-causing agent. The RT-PCR result indicates that the levels of SNAMA transcripts are increased during early embryonic developmental stages. After sequencing SNAMA-1 and SNAMA-2 transcripts it is clear that there is a single transcript. Northern blot analysis will verify the observations above.

4.3 *SNAMA* cDNA SEQUENCE

The cDNA for *SNAMA* was obtained from Berkley Drosophila Genome Project (BDGP), University of California Berkley, Berkley, CA 94720. The BDGP is a consortium of the Drosophila Genome Centre. The *SNAMA* gene was identified during the 1999 EST Project of the BDGP in the LD library made by Ling Hong (Rubin *et al.,* 2000). *SNAMA* was submitted to the EST database on 02/14/2000, as clone number BcDNA: LD21643. The sequence has been submitted to GenBank by BDGP on the 01/03/1999, accession number AF132177. The cDNA for *SNAMA* was cloned into the 5′ EcoRI site and the 3′ XhoI site of the pOT2 vector

[SP6Promoter.XhoI.***SNAMA***.EcoRI.BstXI.EcoRV.BglIII.T7 Promoter] (Figure 18). The pOT2 vector was constructed by Oroon Hubbard, and was designed to give optimal results for transposon mapping and sequencing. The inserts were then sequenced to construct the cDNA library for the BDGP EST sequencing project. We obtained the *SNAMA* cDNA from BDGP. The *SNAMA* insert was released after restriction digest with *EcoRI* and *XhoI*. This results in two fragments because of an internal EcoRI site within the cDNA of SNAMA downstream to the DWNN domain. The digest was electrophoressed and the 914 bp insert purified from the gel. The cDNA fragment includes the conserved DWNN domain. This DNA was used as a probe in the Northern blot analysis and in *in situ* hybridisation experiments.

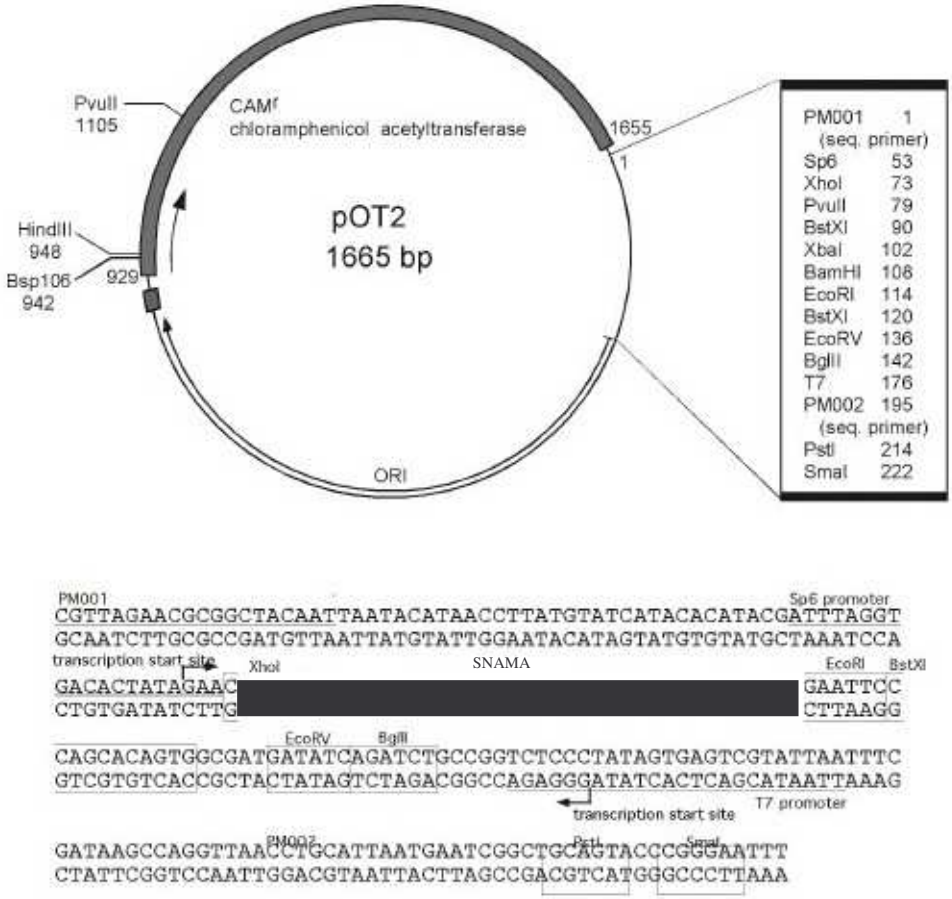


Figure 18: pOT2 vector with the *SNAMA* insert. *SNAMA* DNA probes were obtained after a restriction enzyme digest with *EcoRI* and *XhoI*. This insert was labeled either for use in the Northern blotting experiment (section 2.8.2) or for probes in the *In situ* hybridisation experiments (2.8.1).

4.4 SNAMA IS EXPRESSED THROUGHOUT DEVELOPMENT

To establish the temporal distribution and the size of the *SNAMA* mRNA, developmental Northern blots were prepared as indicated in the materials and methods section 2.8.2. The blot was probed with a radioactive ^{32}P-labeled DNA probe prepared using random prime labelling of a 914 bp *SNAMA* fragment (*EcoRI* and *XhoI* restricted). The 914 bp N-terminal *SNAMA* fragment included sequence for the conserved DWNN domain. A single 4.6 kbp transcript is clearly visible throughout embryogenesis (Figure 19(A)). The transcript is present at higher levels in early embryogenesis and is significantly reduced in the late stages of embryonic development. Few genes are transcribed in the first 3 hours after fertilisation (Edgar and Schubiger, 1986). Thus, the high levels of expression of *SNAMA* in 0-3hr embryos indicate its vital role in development. This is consistent with a similar observation in *l(2)rQ313* flies. Homozygous offspring of the P-element flies *l(2)rQ313* die during embryogenesis since they lack *SNAMA* transcription. The SNAMA transcript is not easily detected by the Northern blot in adult flies, but was detected by RT-PCR (Figure 20(B)). This indicates that SNAMA expression is reduced in both male and female flies. RT- PCR is a sensitive technique for detection of very low levels of RNA. The Northern blot result is not consistent with results obtained from RT-PCR (section 4.2).

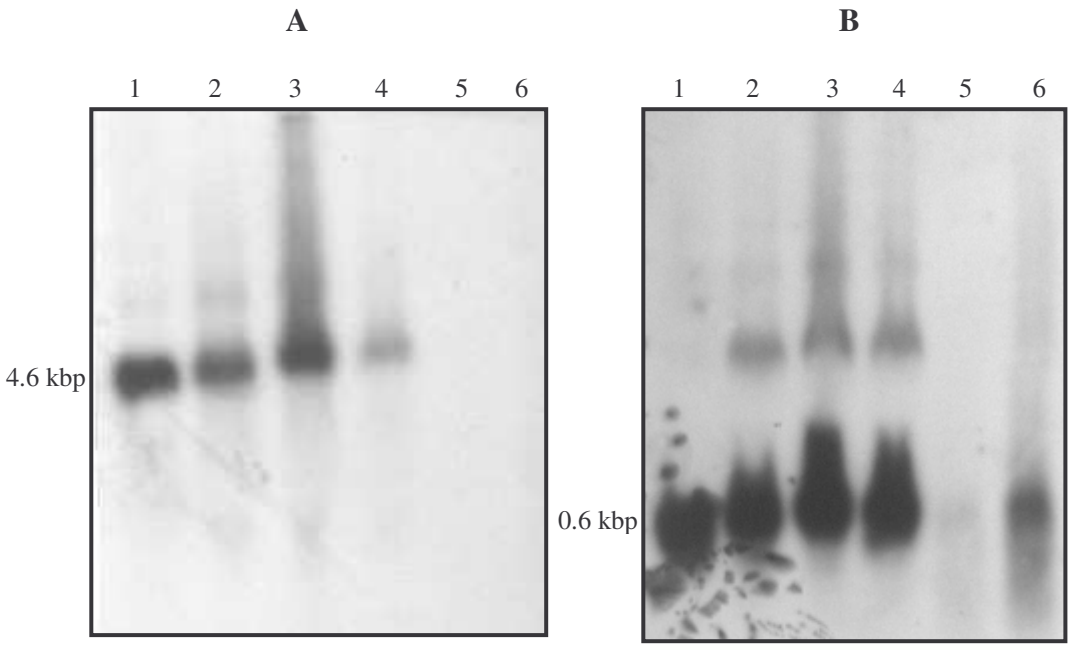


Figure 19: Developmental Northern blot of *D. melanogaster* RNA from various embryonic developmental stages and adult flies. Approximately 50 µg of total RNA was loaded on each lane. Lane 1: 0-3hr, lane 2: 3-6hr, lane 3: 6-12hr, lane 4: 12-18hr, lane 5: adult males and lane 6: adult females. The transcripts were detected using radiolabeled DNA probes. **(A)** The blots were first probed with a 914 bp *SNAMA* probe, then stripped and **(B)** reprobed with L18a as a control. A single 4.6 kbp *SNAMA* transcript was detected in embryos but not in adults. The 0.6 kbp L18a was detected on the same blot after stripping and reprobing. L18a was detected in adult male and female flies.

It is now obvious that there was possible genomic DNA amplification in the RT-PCR (Figure 17) resulting in the two transcripts. After repeating the RT-PCR using the same RNA samples collected for the Northern blot, a single transcript was detected in the 0-3 hr, 6-12 hr and 12-18 hr embryos. A decreased level of *SNAMA* transcription was observed in later stages of embryogenesis (Figure 20(A)), consistent with Northern blot results. SNAMA is expressed at very low levels in adult male and female flies indicating the importance of the protein in early development. It is now evident that there is only a single *SNAMA* transcript in all stages of development.

To confirm that an equal amount of RNA was loaded into each lane, the blot was stripped and reprobed using a probe complementary to the *D. melanogaster* ribosomal protein L18a. L18a is expressed at equal levels throughout embryonic development (Ntwasa *et al.*, 1994). L18a is only weakly expressed in adult male embryos and also appears at lower levels in female adults. The larger band on the L18a blot was *SNAMA* transcripts that were not removed after stripping of the blot before reprobing.

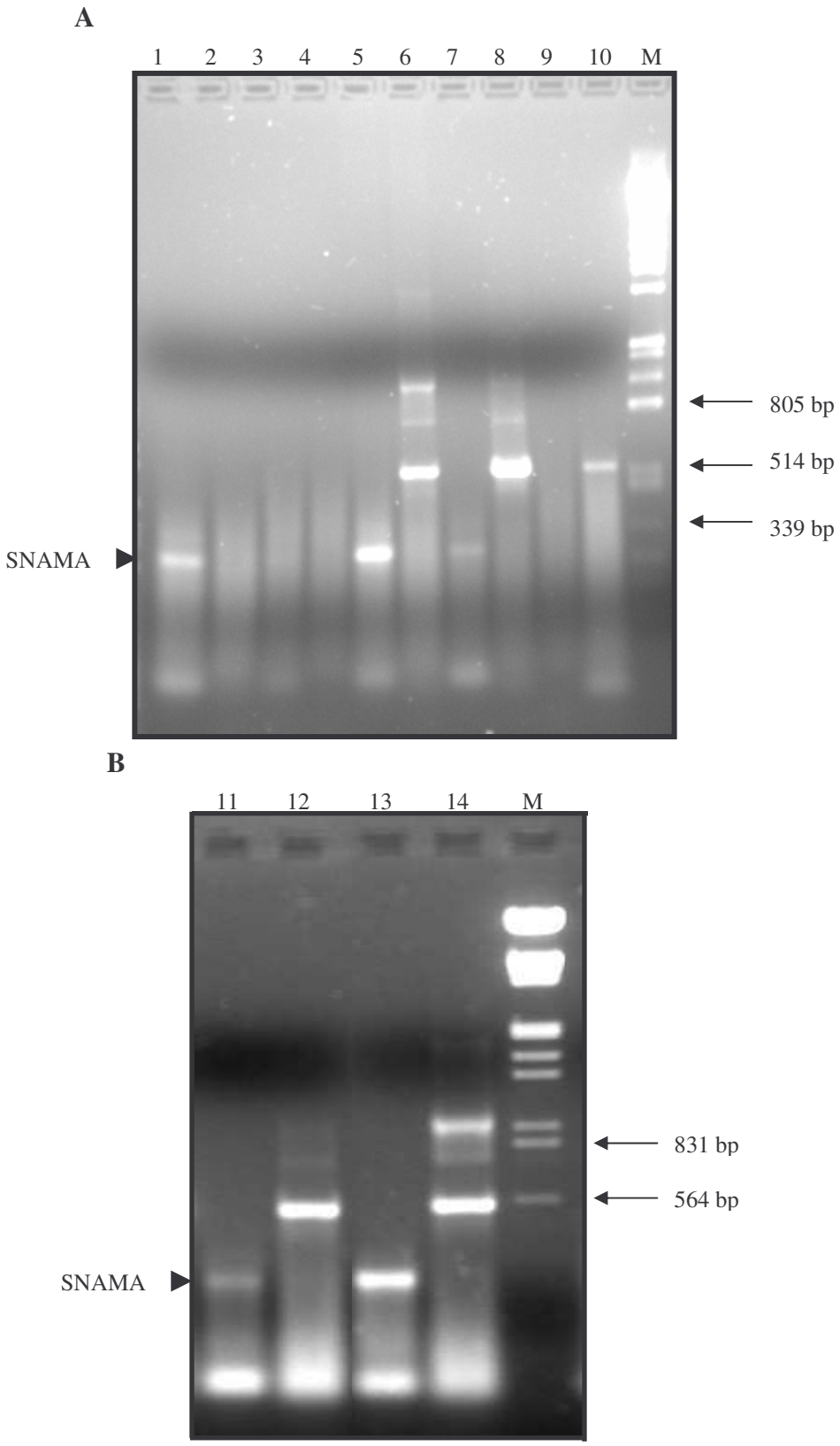
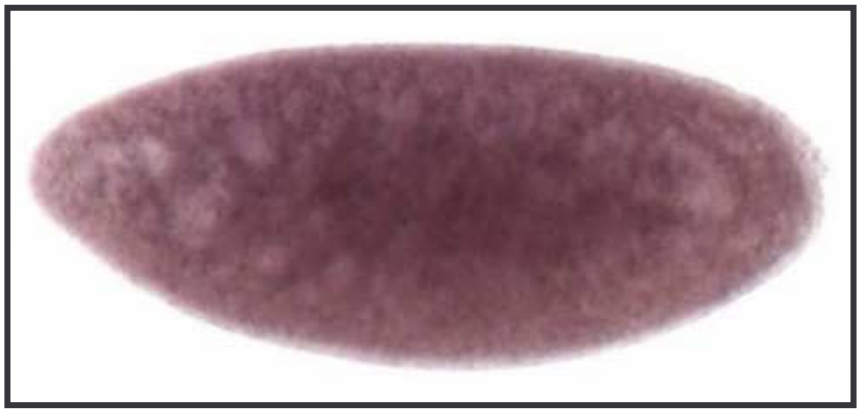

A
1 2 3 4 5 6 7 8 9 10 M
805 bp
514 bp
339 bp
SNAMA
B
11 12 13 14 M
831 bp
564 bp
SNAMA

Figure 20: Reverse transcriptase PCR was performed on various developmental stages of *D. melanogaster* RNA and resolved on a 2% agarose gel **(A)** or on a 1.5% gel **(B)**. To standardize the amount of RNA amplified per developmental stage, a ribosomal protein L18a was amplified. Lanes 1,3,5,7,9,11 and 13: SNAMA transcripts, lanes 2,4,6,8,10,12 and 14: L18a transcripts. Lanes 1 and 2: 0-3hr embryonic development, lanes 3 and 4: 3-6hr embryonic development, lanes 5 and 6: 6-12hr embryonic development, lanes 7 and 8: 12-18hr embryonic developmental stage, lanes 9,10,11 and 12: adult males, lanes 13 and 14: adult females and M: marker (Roche Marker III). Single transcripts at 234 bp are noticed in 0-3hr, 6-12hr, 12-18hr embryos and in adult males and females.

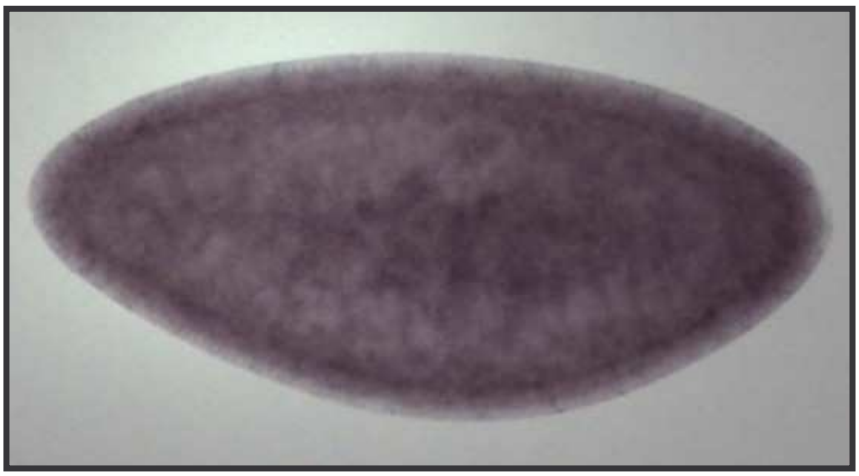
4.5 SNAMA MAY BE LOCALISED IN THE NUCLEUS OF *D. melanogaster* EMBRYOS

The technique for the localisation of mRNA *In situ* was first developed in the late 1970s. This technique reveals patterns of gene expression, both temporally and spatially, at a cellular resolution. Whole mount embryo *in situ* hybridisation was used to investigate the tissue expression of *SNAMA* transcripts. Embryos from various stages of development ranging from stages 1-15 (0-12hr development) (Campos-Ortega and Hartenstein, 1997) were collected in the same tube. DNA probes specific to *SNAMA* were prepared as indicated in the materials and methods section 2.8.1. *SNAMA* transcripts were detected throughout the precellular blastoderm distributed uniformly seemingly expressed throughout the embryo (Figure 21). Stronger staining is observed in stage 5, cellular blastoderm stage embryos as compared to stage 11 embryos. Stage 11 embryos stain along the germ band with no staining near the head region. Further investigation of the staining pattern is necessary to confirm the authenticity of the staining reactions.

A



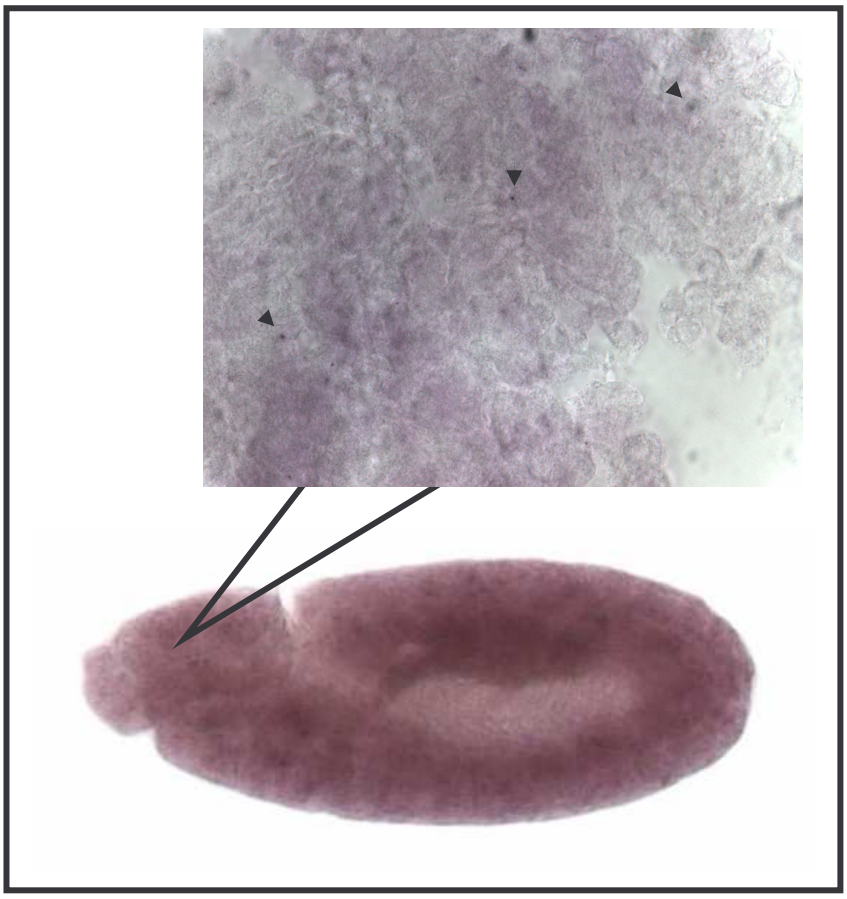
B



C



D



E



F

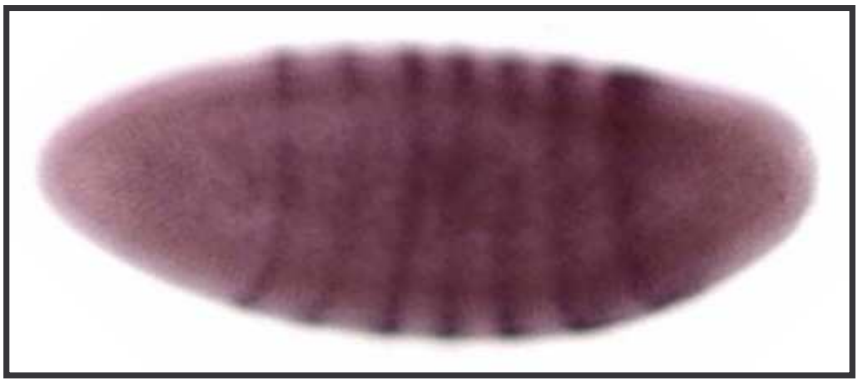
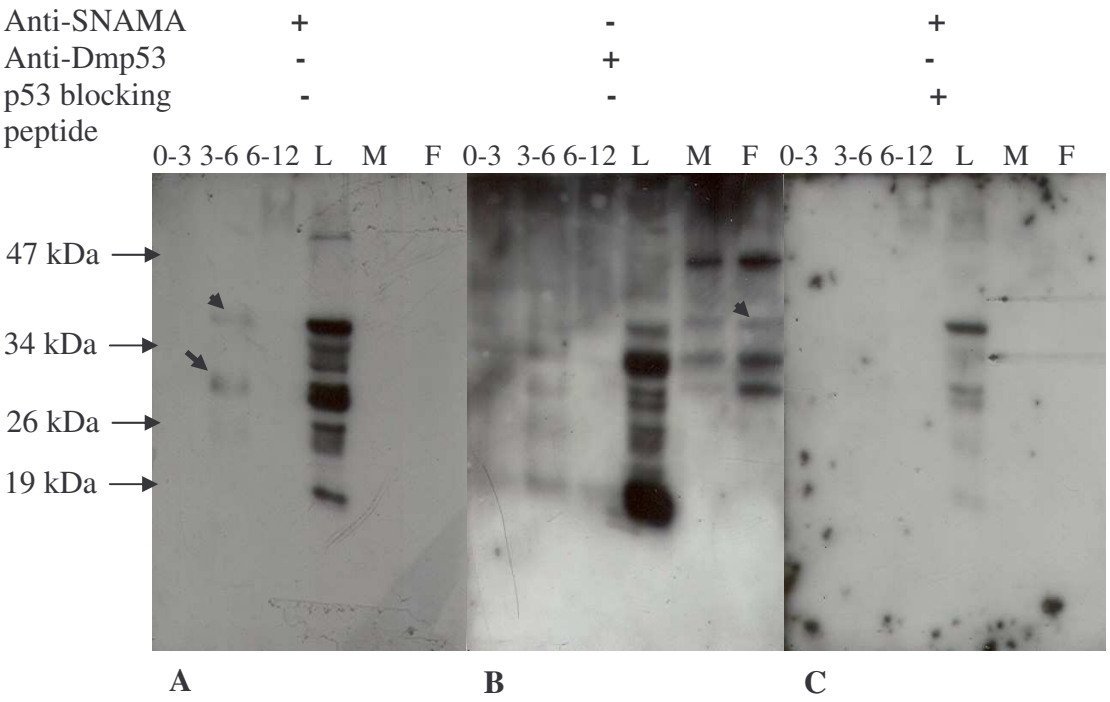


Figure 21: *In situ* staining of developmental stages of *D. melanogaster* embryos to detect *SNAMA* transcription. All embryos are orientated with anterior to the left. **(A)** Expression of *SNAMA* transcripts at syncitial blastoderm stage. **(B)** Stage 5, cellular blastoderm stage. **(C)** Stage 11, lateral view. **(D)** Stage 14, lateral view. Inset is a 1000X magnification of the head region showing nuclear localisation of *SNAMA* transcripts. Transcripts are distributed randomly throughout the embryo in all stages of development. **(F)** and **(E)** Controls done in parallel with the above experiments. **(E)** Blastoderm embryos stained with *evenskipped* DNA probe. **(F)** Stage 11 embryos stained with *Dpp* DNA probe.

4.6 SNAMA INTERACTS WITH Dmp53

The homologues of SNAMA, PACT and RBBP6 in humans and mouse respectively, have been shown to bind and inactivate p53 and Rb. This suggests that SNAMA also has p53 and pRB binding properties and may inactivate p53. To investigate this interaction, Western blots were prepared as described in section 2.8.5. Protein extracts were run on a 12% SDS-polyacrylamide gel. The blot was first probed with anti-Dmp53 (Figure 22(A)), stripped and then reprobed with anti-SNAMA antibody (Figure 22(B)), stripped again and then blocked with p53 blocking peptide and probed with anti-SNAMA (Figure 22(C)). A 42 kDa band was identified as Dmp53 on a Western blot using cell lysates from wild-type homozygous larvae (Lee *et al.*, 2003). This anti-Dmp53 antibody cross-reacted with other proteins as indicated by multiple bands ranging from 19 kDa to about 118 kDa on the Western blot (Figure 22(B)). This is not unusual and a Western blot using this anti-Dmp53 antibody has been shown to result in a banding pattern (Lee *et al.*, 2003). It is important to note that a major band at 20 kDa was also observed in addition to the band at 42 kDa and that most of the banding is found at the larval stage. In a personal communication, the suppliers of the antibody agreed that multiple banding is observed with the antibody. Smaller bands probably represent degradation products of Dmp53. The pattern of multiple bands is more dramatic at larval stages when significant apoptosis takes place (Figure 22(D)). All the bands disappear when the blot is first blocked with Dmp53 blocking peptide and then probed with Dmp53. When the blot is blocked with Dmp53 blocking peptide and probed with anti-SNAMA some of the bands remain. In

this way a set of bands have been isolated that are recognised by both antibodies indicating a tight association between SNAMA and Dmp53. It is also possible that SNAMA interacts with other proteins in addition to Dmp53 because the anti-SNAMA antibodies cross-reacts with these proteins. A possible reason for this is that these are degradation products of SNAMA. A similar result has been observed on Western blots of the human DWNN homologue (Mbita Z., personal communication). Western blots on the human SNAMA homologue probed with anti-DWNN human antibody also identified multiple bands. It is also likely that SNAMA is attached to other proteins in a manner similar to that of ubiquitin. This attachment may be either non-covalent as seen in the Hub1 protein, or a covalent attachment requiring a yet unidentified set of ligases. This would explain the multiple banding patterns in the larval developmental stage on the anti-SNAMA blot. Yet another possibility is that the multiple bands observed in the larval stages for SNAMA may be due to the very basic charge centred at the C′-terminus of the protein interfering with the migration of SNAMA on the gel. This would result in antibodies reacting with SNAMA protein through charge recognition and not through conformational recognition entirely. This point needs to be clarified by an *in vitro* acetylation reaction. This chemical modification will neutralise the positive charge at the C′-terminus, and throughout the protein, and ensure antibody interaction with SNAMA protein through conformational recognition and also ensure that the protein enters the SDS-PAGE gel.


Anti-SNAMA + - +
Anti-Dmp53 - + -
p53 blocking peptide - - +
0-3 3-6 6-12 L M F 0-3 3-6 6-12 L M F 0-3 3-6 6-12 L M F
47 kDa
34 kDa
26 kDa
19 kDa
A
B
C

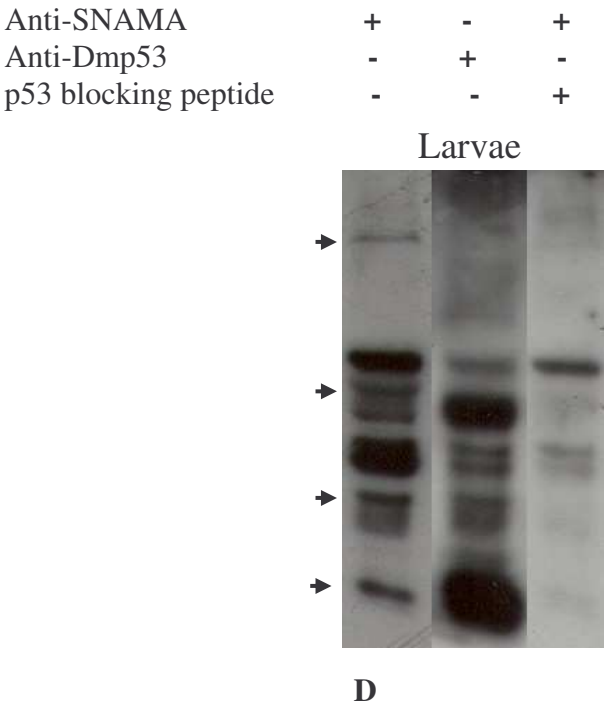

Anti-SNAMA + - +
Anti-Dmp53 - + -
p53 blocking peptide - - +
Larvae
D

Figure 22: Western blot analysis of proteins from embryonic, larval and adult stages of wild-type *D. melanogaster* development. Protein extracts were separated on a 12% SDS PAGE gel. Lane 1: 0-3hr, lane 2: 3-6hr, lane 3: 0-12hr, lane 4: lavae (L), lane 5: male adults (M) and lane 6: female adults (F). Marker position is indicated by distance migrated on the polyacrylamide gel. The blot was first probed with **(A)** anti-SNAMA antibody, stripped and then reprobed with **(B)** anti-Dmp53 antibody and **(C)** with anti-SNAMA after blocking with Dmp53 blocking peptide. The lanes of the larval stages are aligned in **(D)**. A 42 kDa-Dmp53 signal was detected (arrowhead in **(B)**). The banding pattern is characteristic for this anti-Dmp53 antibody. Arrows indicate proteins that are recognised by both antibodies.

D. melanogaster larvae that were fed with the apoptosis-inducing agent camptothecin showed increased levels of SNAMA protein expression on Western blots probed with anti-SNAMA. Camptothecin is a topoisomerase inhibitor shown to affect the p53 pathway (Takada *et al.*, 2003). Protein extract from larvae fed with media supplemented with 1 mM camptothecin overnight indicated increased levels of Dmp53 expression when Western blots were probed with anti-Dmp53 (Figure 23(A)). When the same blot is stripped and probed with anti-SNAMA we noticed increased levels of SNAMA expression indicated by the arrow (Figure 23(B)). The anti-SNAMA antibody forms immunoreactive complexes with both SNAMA and Dmp53. These results indicate interactions between Dmp53 and SNAMA and also establish SNAMA as a protein involved in apoptosis via the Dmp53 pathway.

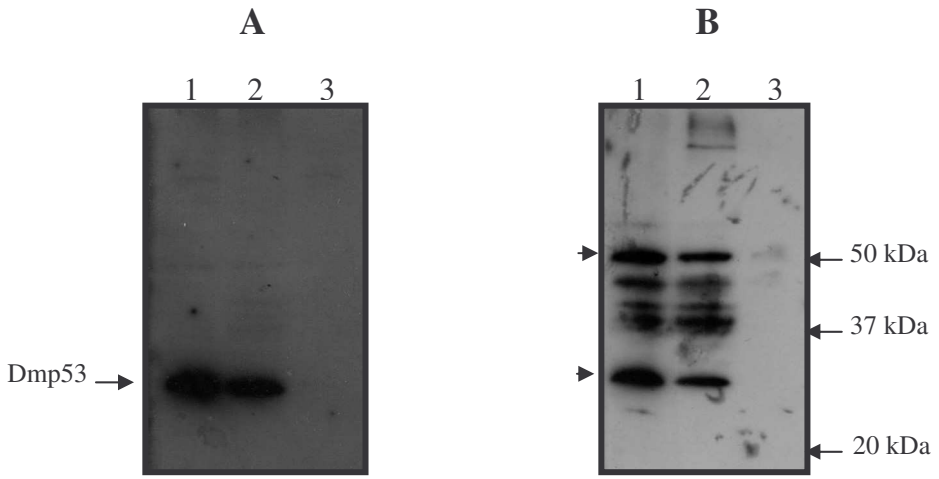


Figure 23: Western blot analyses of proteins from *D. melanogaster* larvae that have been fed with 1 mM camptothecin overnight. Marker positions are indicated. **(A)** Camptothecin activates Dmp53 protein expression. **(B)** The blot was stripped and probed with anti-SNAMA. Proteins with increased levels of expression are indicated by arrowheads. Again we noticed the anti-SNAMA antibody cross-reacting with Dmp53 suggesting association of Dmp53 with SNAMA. Lanes 1: larvae fed with 1 mM camptothecin, Lane 2: larvae media supplemented with DMSO (control) and Lane 3: Wild-type larvae.

CHAPTER 5

THE BIOLOGICAL FUNCTION OF SNAMA IN *Drosophila melanogaster*

5.1 Introduction

5.2 *SNAMA* is an essential gene required for development

5.3 *SNAMA* is characteristic of genes that are involved in apoptosis

5.1 INTRODUCTION

The functional roles of DWNN proteins have not been established. The isolation of SNAMA, the *D. melanogaster* DWNN homologous protein, offers us the opportunity to study its role in a complex biological system. The *D. melanogaster* embryo is an ideal tool to determine developmental roles of genes. The *D. melanogaster* embryo is easy to manipulate and goes through very distinct developmental stages. The embryo starts as a unicellular blastoderm that eventually develops into a multicellular complex organism. In addition, we have also obtained a *P{PZ}* insertion fly line $l(2)rQ313^{rQ313}$ (accession number AQ026141) (Spradling *et al.*, 1999) from the EP flystation by Exelixis Inc. This fly stock has a P-element single insertion at position 60B9 and was developed as part of the BDGP gene disruption project. The P-element insertion results in disruption of the SNAMA gene. This fly stock was used to investigate the role of SNAMA in developing embryos.

5.2 *SNAMA* IS AN ESSENTIAL GENE REQUIRED FOR DEVELOPMENT

5.2.1 P-ELEMENT MUTATIONS IN *Drosophila melanogaster*

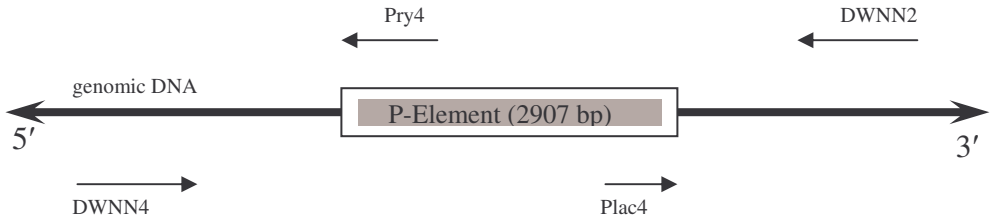
P-elements are natural transposable elements that are active in the germline of flies. They encode an enzyme called transposase that enables these elements to move to other chromosomal locations. P-element plasmid constructs have been used as genetic vehicles for the introduction of cloned DNA into the germline. The introduction of

these elements allow for the functional analysis of any piece of DNA in the context of the whole organism (Tautz, 1992). This has become an important tool for the determination of molecular interactions and gene functions. The P-element insertion EP line *P{PZ}l(2)rQ313*^{rQ313} has a P{PZ} insertion at 60B9. This insertion results in a disruption of the *SNAMA* gene resulting in a lethal mutation. It is thus evident that SNAMA protein has a vital role during the development of *D. melanogaster* embryos since embryos from *l(2)rQ313* flies do not develop to adulthood and die during late embryonic stages.

The gene disrupted by the P-element insertion is likely to be lost from the population because it is an essential gene for survival. This lethal recessive gene confers a disadvantage on the progeny of the heterozygotes. To keep the mutant P-insertion lines alive, the flies are crossed with another recessive lethal gene on the same chromosome. This is known as the balancer chromosome. The balancer carries a clearly visible dominant marker that allows us to follow the fate of the balancer chromosome in any further crosses. The balancer chromosome in *l2rQ313* flies is the *CyO* gene.

5.2.2 IDENTIFICATION OF THE POSITION OF THE P-ELEMENT IN l*(2)rQ313* FLIES

The P-element insertion line *P{PZ}l(2)rQ313*^{rQ313} has a P-element inserted in the first intron in the 5´-untranslated region of the *SNAMA* gene. This fly line was obtained from the EP flystation by Exelixis Inc., stock EP0503. The P-element insertion that resulted in the disruption of the expression of the *SNAMA* gene is lethal in homozygous offspring and embryos never develop to adulthood. These flies were used to investigate the functional role of SNAMA in developing embryos. The balancer chromosome on *l(2)rQ313* flies is *CyO*. This results in extra veinage in the wing cells and the vein V does not go to the margin resulting in the wing forking. The *CyO* phenotype has wings curled upwards. This curvature is caused by the unequal contraction of the upper and lower epithelia during the drying period following emergence from the pupa case. The *CyO* balancer chromosome is also an essential gene that results in death during late embryo or early lavae stages when not expressed (Kidwell, 1972). The sequence for the P-element insertion was recovered by inverse PCR (Spradling *et al.*, 1999) and we have identified its position on the *D. melanogaster* genome (Figure 8). PCR was used to identify the P-element insertion site. Genomic DNA was collected from wild-type *Canton-S* adults, heterozygous *l(2)rQ313* adults and homozygous *l(2)rQ313* unhatched embryos. The genomic DNA was then amplified using 5′and 3′ primers designed to complement the P-element inverted terminal repeat sequence (Pry4 and Plac4) and the flanking region of SNAMA (DWNN4 and DWNN2), respectively (Figure 24).



Primer Combination	Wild-type genomic DNA	Homozygous *l(2)rQ313* genomic DNA	Heterozygous *l(2)rQ313* genomic DNA
Plac4/ DWNN4	---	---	---
Pry4/ DWNN4	---	663bp	663bp
DWNN2/ DWNN4	1172bp	1172bp	---

Figure 24: Schematic representation of the primer positions on the genomic DNA from *l(2)rQ3131* flies. The table indicates the expected PCR product from wild-type and *l(2)rQ313* flies using the chosen primer combinations. The PCR products expected differ because w/type flies do not contain the P-element insert. The P-element insert prevents amplification across the length due to the large size of the insert in *l(2)rQ313* flies.

DNA concentrations were determined spectroscopically so that equal amounts could be amplified by PCR. For *l(2)rQ313* mutants that had a P-element insertion at the expected position, PCR products would be observed at the indicated position for the chosen primer combinations (Figure 24). As expected, a 1172 bp product was detected in lane 3 and no product was amplified for the wild-type genomic DNA in lanes 1 and 2 (Figure 25(A)). No PCR product from the heterozygous *l(2)rQ313* genomic DNA in lane 4, a 663 bp product in lane 5 and a 1172 bp product in lane 6 was detected. For the homozygous *l(2)rQ313*, there was no PCR product in lane 7, a 663 bp product in lane 8 and no amplification in lane 9.

A primer combination of Plac4 and DWNN4 results in no amplification products for the wild-type, heterozygous and homozygous *l(2)rQ313* since these primers are in the same direction. This primer combination was chosen as an internal control. Primers Pry4 and DWNN4 results in no amplification products in the wild-type DNA due to the absence of the P-element in these flies. Pry4 is complementary to sequence on the P-element. For this primer set, both the *l(2)rQ313* extractions amplified a 663 bp product. There was reduced amplification in the heterozygous extraction since only a single copy of the insertion exists in these flies. They develop to adulthood, while the homozygous embryos never hatch. Primers DWNN2 and DWNN4 amplified a 1172 bp product in wild-type flies and in heterozygous flies while no amplification occurred in homozygous flies. The P-element is large and amplification across the P-element is not possible in the homozygous mutants. A decrease in amplification

product was observed for the heterozygous flies since they only have a single copy of the P-insertion.

Homozygous *l(2)rQ313* embryos were selected from apple juice agar plates since these embryos don't hatch because of the *SNAMA* gene being disrupted. *l(2)rQ313* was balanced with *CyO* (with an effective lethal phase late embryo, early larva) (Kidwell, 1972). Thus the collection of unhatched embryos not only contained homozygous *l(2)rQ313* , but also flies that died due disruption of the balancer gene. To ensure that only homozygous *l(2)rQ313* are collected, these flies were crossed with *CyO pAct GFP*. *Cyo pAct GFP* flies express GFP (green fluorescent protein) ubiquitously. Embryos from these flies fluoresce green when placed under white light. Crossing these flies with the *l(2)rQ313* mutants would result in the homozygous SNAMA P-element flies not expressing the GFP tag and the required flies are then selected from GFP fluorescent embryos. To establish the presence of the P-element in this new cross, the *l(2)rQ313/GFP* DNA was amplified using primers specific for the P-element sequence (Figure 25(B)). Again, the same bands were observed as for the *l(2)rQ313* mutants. These PCR results indicated that a P-element has been inserted into the *l(2)rQ313* line at the expected position and that this P-element was not lost in the *l(2)rQ313/GFP* fly line.

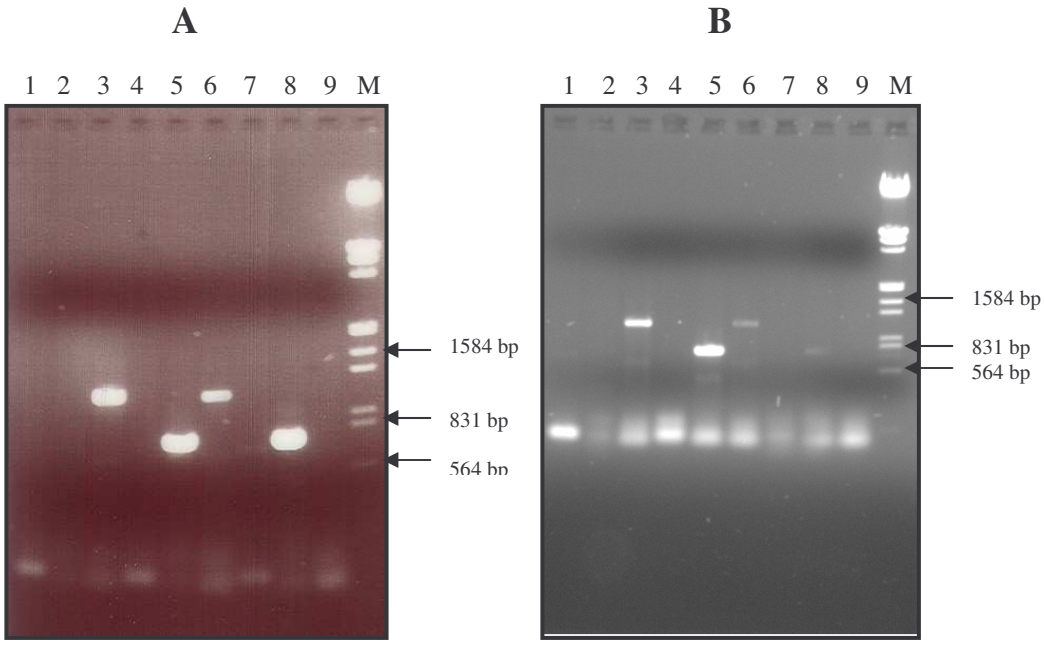


Figure 25: PCR amplification products on a 0.8% TAE agarose gel to detect P-element insertion in *l(2)rQ313* **(A)** and *l(2)rQ313/GFP* **(B)**. Lanes 1,2,3: wild-type flies (Plac4/DWNN4 primer set), lanes 4,5,6: heterozygous flies (Pry4/DWNN4) and lanes 7,8,9: homozygous embryos (DWNN2/DWNN4). M: Marker III (Roche). The bands were observed at the expected sizes respectively (see Figure 24). This confirmed the presence of the P-element in both the *l(2)rQ313* and the *l(2)rQ313/GFP* fly lines.

5.2.3 MOBILISATION OF THE P-ELEMENT FROM *l(2)rQ313* FLIES

To establish whether the P-element insertion in *l(2)rQ313/GFP* contributed to the observed phenotype, the P-element from *l(2)rQ313/GFP* was mobilised by crossing these flies with flies containing Δ2-3 transposase. To mobilise the P-element *w;Cy/GFP;Sco* female flies were crossed with *Sb [ry^{+} Δ2-3]/ TM6b* males. The male offspring were crossed with *l(2)rQ313* to generate the flies in which the P-element had been mobilised. This excision line (scored by the loss of eye colour) was analysed by a PCR method using the primers flanking the P-element in the SNAMA gene (Figure 26). A primer combination of Plac4/DWNN4 and Pry4/DWNN2 were used in lane 1 and 2 respectively. No PCR amplification was observed since primers Plac4 and Pry4 are complementary to the P-element sequence that had been excised. An amplification product was observed at 1172 bp in lane 3 using primers DWNN4 and DWNN2. This PCR product could only be amplified when the large P-element sequence was excised. This indicated mobilisation of the P-element that disrupted the SNAMA gene. These flies are viable and develop to adulthood unlike the *l(2)rQ313* flies. Some offspring show only small defects as a result of imprecise excision of the P-element. These crosses establish that SNAMA is a vital gene necessary for *D. melanogaster* development.

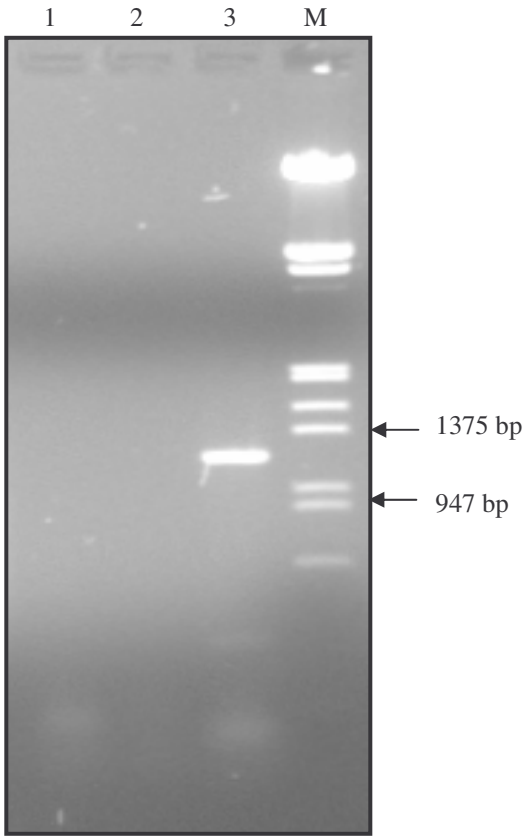
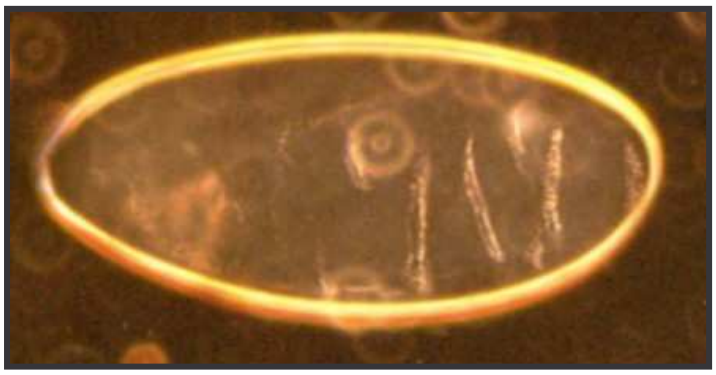


Figure 26: PCR of genomic DNA from excision cross flies to test for excisions of the P-element. Lane 1: primer combination Plac4 and DWNN4 results in no PCR product, lane 2: primers Pry4 and DWNN4 results in no PCR product and lane 3: primers DWNN2 and DWNN4 amplifies a 1172 bp product. The P-element has been removed in the cross.

5.2.4 CUTICLE ANALYSIS OF *l(2)rQ313* EMBRYOS

Analysis of the cuticles of *D. melanogaster* embryos highlighted structures such as the denticle bands. These are hair-like projections on the ventral cuticle of a larva and are indicative of development of segmental structures. *D. melanogaster* embryos homozygous for *l(2)rQ313* display a spectrum of developmental defects as indicated by the irregular and disrupted denticle bands. Figure 27(A-C) indicates missing denticle belts at the posterior end of the embryo. Embryos were observed that had bands missing at abdominal segments a5, a6, a7, and a8. Some embryos are able to form a normal number of segments, but the posterior segments are severely compressed (Figure 27(D)). This suggests a failure in posterior midgut involution. The *l(2)rQ313* embryos show severe and variable pattern defects indicative of morphological defects.

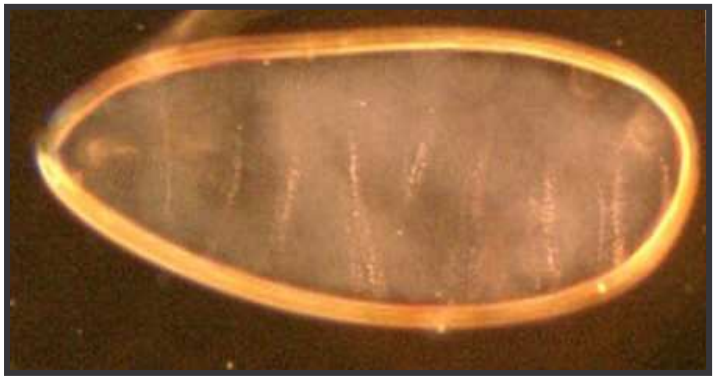
A



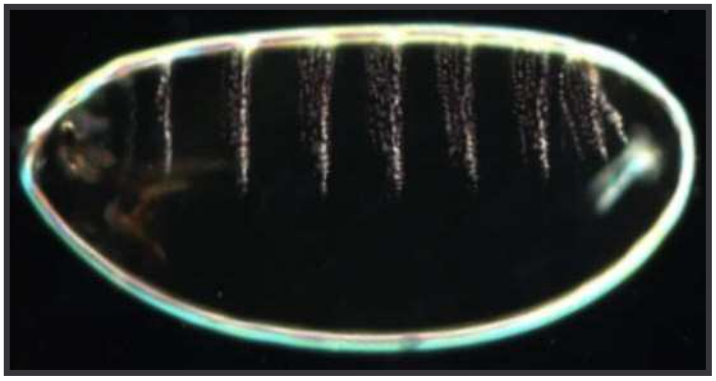
B



C



D



E

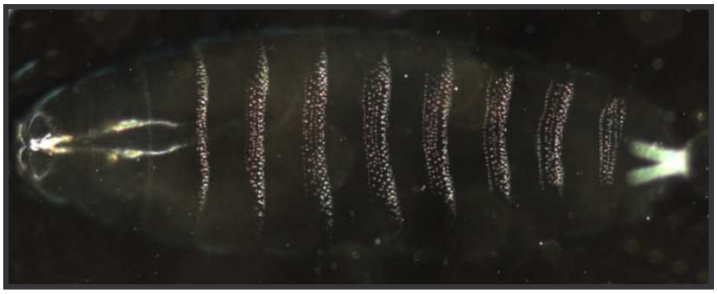


Figure 27: Cuticle pattern of *l(2)rQ313* embryos. Cuticles prepared from unhatched embryos laid by *l(2)rQ313* flies shows severe and variable pattern defects. **(A-C)** Unhatched mutant embryos showing an unusual complement of denticle bands with a very disordered arrangement. **(D)** Mutant embryos with a normal complement of denticle belts, but bands at the posterior end are compressed. **(E)** Wild-type embryo has a very regular and ordered denticle banding pattern.

5.3 *SNAMA* IS CHARACTERISTIC OF GENES INVOLVED IN APOPTOSIS

A large number of cells die during normal embryonic development in *D. melanogaster*. Apoptosis can be visualised in living wild-type and *l(2)rQ313* embryos using the vital stain acridine orange (Abrams *et al.*, 1993). Acridine orange selectively stains apoptotic cells while necrotic cells are not labelled. *D. melanogaster* cell death begins at stage 11 (~7 hours) of embryogenesis becoming widespread throughout the embryo. The initial screen that identified the DWNN proteins was a screen for the resistance of mutated cells to apoptotic inducing agents. This indicated a possible role of the DWNN proteins and the SNAMA protein in apoptosis. To test whether SNAMA plays a role in apoptosis the P-element flies in which the SNAMA gene was disrupted were examined. This would indicate the levels of apoptosis in *l(2)rQ313* flies where levels of SNAMA have been decreased.

The insertion of a P-element into the first intron of *SNAMA* disrupts the SNAMA gene creating recessive lethal mutants with abnormal patterns of apoptosis (Figure 28). Embryos were stained with the vital stain, acridine orange, which is specific for cells undergoing apoptotic cell death (Abrams *et al.*, 1993). These P-element mutants show abnormal levels of apoptosis as compared to levels in wild-type flies. This is observed as increased levels of staining in these embryos. The acridine orange stained embryos show ectopic apoptosis in late stage embryos (Figure 28(B)) and early occurrence of apoptosis in the blastoderm (Figure 28(A)). The blastoderm is a

developmental stage characterised by rapid nuclear divisions with only low levels of apoptosis. This pattern of acridine orange staining was not present in wild-type embryos or after the P-element was completely removed from the *l(2)rQ313* flies. This showed a direct correlation between the presence of SNAMA and apoptosis. An increase in the levels of apoptosis occurred in embryos with relatively low levels of SNAMA expression. SNAMA thus acts as an apoptosis suppressor. The mode of this action is either direct, or via other proteins that are involved in the apoptotic pathway.

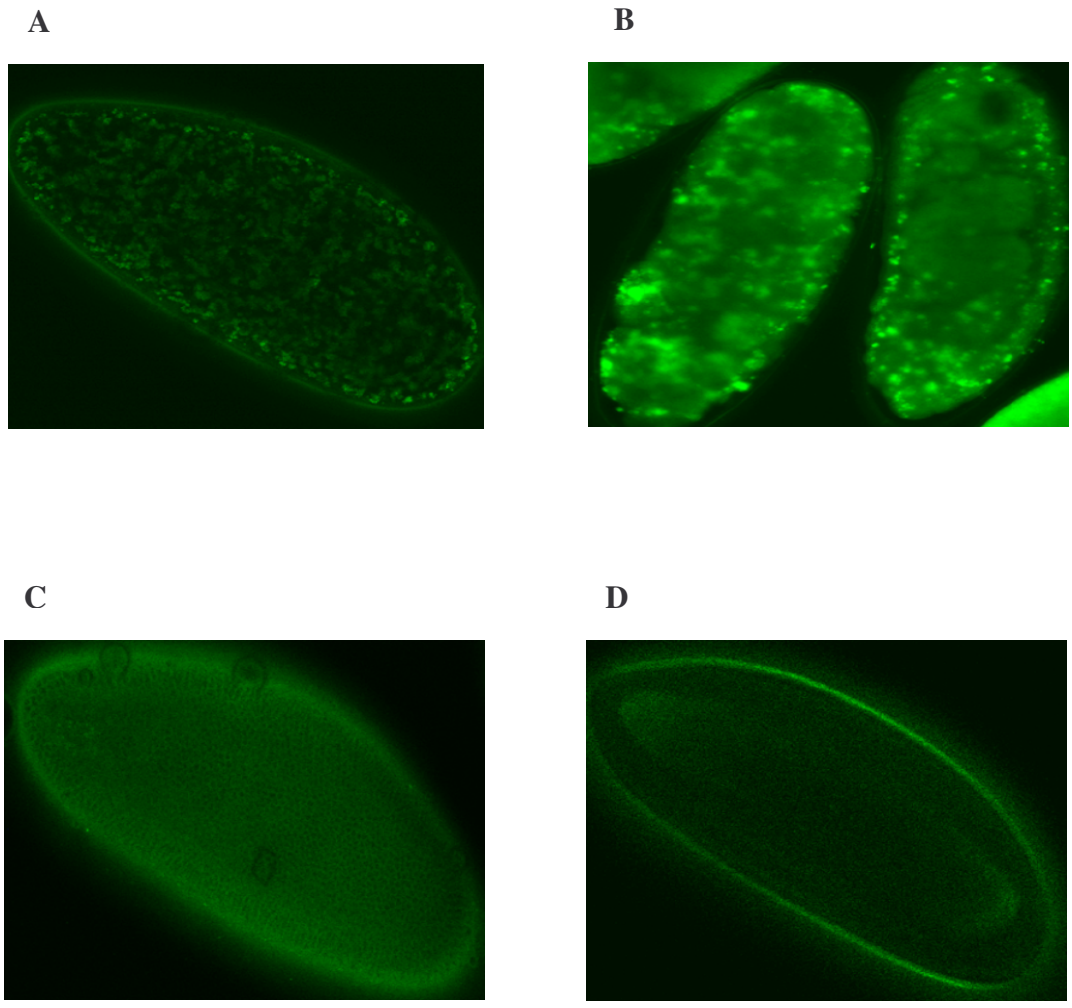


Figure 28: Acridine orange staining of embryos viewed using confocal microscopy. **(A)** *l(2)rQ313* cellular blastoderm. **(B)** Stage 10 *l(2)rQ313* embryos and **(C and D)** Wild-type embryos. The embryos were stained as described in section 2.8.6. Increased levels of apoptosis are observed in embryos lacking SNAMA. This is observed as increased levels of acridine orange staining.

CHAPTER 6

GENERAL DISCUSSION AND CONCLUSION

6.1 The domain organisation of SNAMA protein and its functional implications

6.2 The sequence analysis of *SNAMA*

6.3 The cytological position of *SNAMA*

6.4 The biological function of SNAMA

6.5 Conclusions

6.1 THE DOMAIN ORGANISATION OF SNAMA PROTEIN AND ITS FUNCTIONAL IMPLICATIONS

The DWNN proteins have been identified in a large number of organisms, but their exact function remains unknown. This study characterised the *Drosophila melanogaster* DWNN homologue, SNAMA, and attempted to identify its possible role. These proteins have been identified by virtue of their ability to prevent apoptosis in mammalian cells lacking expression of the gene. They are a novel family of proteins characterised by a highly conserved amino terminal DWNN domain which is found in many eukaryotic species but absent in prokaryotes. The domain was identified through mutagenesis experiments in Chinese hamster ovary (CHO) cell lines screening for genes involved in the antigen presentation pathway and in apoptosis (George, 1995). SNAMA was identified through homology searches using the DWNN domain sequence. Its DWNN domain is found associated with a zinc finger, a cysteine rich modified RING finger domain and a C-terminal region that aligns to the PACT protein. All these domains are well characterised except for the N-terminal DWNN domain.

Homologues of the SNAMA protein have been identified and characterised in a few organisms. These are the human RBBP6 protein, mouse PACT protein, mouse P2P-R protein, and the yeast Mpe1 protein. These proteins were identified based on their functional roles in these organisms rather than identified as DWNN-domain proteins.

The researchers that identified these proteins failed to identify the unique DWNN domain and overlooked its role in these proteins.

The SNAMA protein consists of the 76 residue DWNN domain (amino acid 1-76), a zinc binding domain (CCHC) (amino acids 151-168), a unique cysteine rich RING finger-like domain (amino acids 217-258) (Mather *et al.*, 2005), two basic lysine rich regions (amino acids 805-848 and 1188), and a p53-associated region (Figure 29).

Although very little is known about the function of the DWNN proteins, the high level of evolutionary conservation in eukaryotes suggests that they play an important functional role. There are two human DWNN proteins (Dlamini, unpublished). DWNN-13 is a very short 13 kDa protein of 118 amino acids and DWNN-200 is a 200 kDa protein associated with several characterised domains such a the zinc finger and cysteine rich RING finger-like domains (Mbita, 2004). DWNN-13 is the DWNN domain with a glycine dipeptide together with a C-terminal extension of 39 amino acids. The DWNN domain of SNAMA lacks this glycine dipeptide but retains a conserved proline. Because of the presence of DWNN-13 with a glycine dipeptide it is speculated that the DWNN domain may function similar to ubiquitin by attachment to other target proteins. This possible role of the human DWNN-13 protein suggests that SNAMA may also play a role in the ubiquitin pathway with a folded DWNN domain structure similar to ubiquitin. This thesis investigated the ability of SNAMA to target proteins.

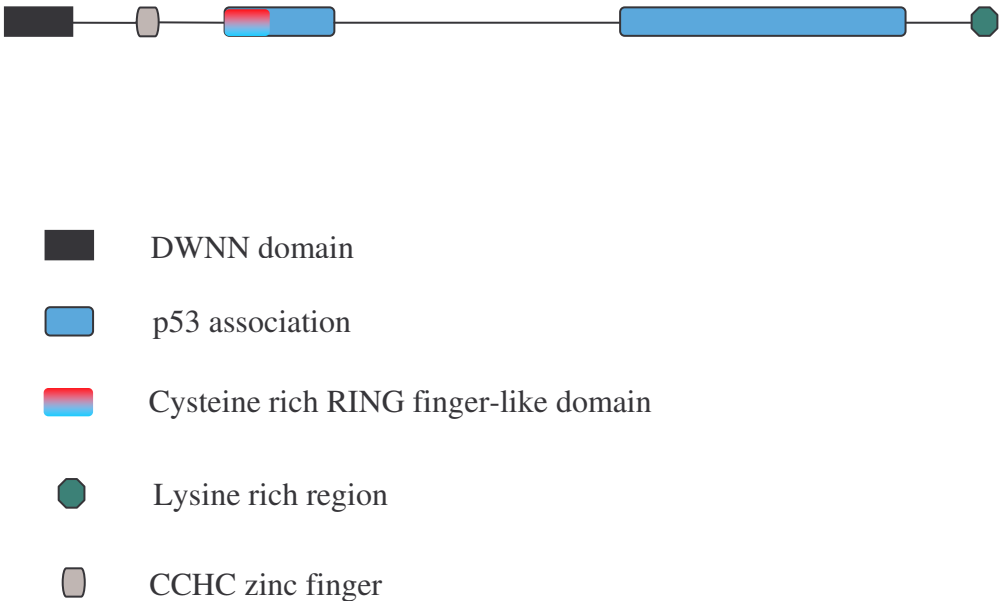


Figure 29: Schematic representation of the domain arrangement of SNAMA protein identified using the SMART programme. The protein has the N-terminal DWNN domain, a zinc finger domain (CCHC) and a cysteine rich RING finger-like domain. It also has a region with high homology to the mouse PACT protein (p53 associated protein) and a lysine-rich C-terminal region.

A closer look at the function of the mammalian SNAMA homologous proteins indicates that SNAMA plays a role in the apoptotic pathway with possible p53 and or pRb binding properties. The SNAMA homologues are the RBBP6, PACT, P2P-R and Mpe1 proteins. RBBP6 is the truncated human homologue of PACT and was identified by virtue of its binding to Rb proteins. PACT binds to wild type p53 and effects p53 specific DNA binding. PACT also interacts with Rb. P2P-R, the alternatively spliced product of the PACT gene (Gao and Scott, 2003), induces mitotic arrest and mitotic apoptosis, indicating that P2P-R has the potential to serve an important role in apoptosis. The mammalian SNAMA homologues have been shown to have p53 and Rb binding ability and are nuclear proteins. The Rb protein is involved in cell differentiation and is localized to centres of mRNA processing in the nucleus. Mpe1, the yeast homologue of SNAMA is involved in 3′ end processing of pre-mRNA forming part of the endonucleolytic complex at the cleavage and polyadenylation site. It is a putative RNA-binding protein because of the zinc finger between amino acids 181 and 192 (Vo *et al.*, 2001). This is supported by evidence that the zinc finger domain present in the CPSF 30K subunit of the mammalian cleavage and polyadenylation specific factor is involved in the interactions of this protein with the pre-mRNA (Barabino *et al.*, 1997). SNAMA shares the RNA binding zinc finger with Mpe1 protein suggesting that SNAMA may have a role in 3′ end processing as well. Taken together these homologues function in transcriptional regulation and the apoptotic pathways by binding p53 and Rb. Sequence homology suggests that SNAMA may have similar functions to its homologues. It is also important to note that p53 and Rb binding proteins are few and that Mdm2 also has

p53 and Rb binding ability (Lohrum *et al.*, 2000). Since *Drosophila melanogaster* lacks an Mdm2 homologue, it is plausible that SNAMA is the functional homologue. The role of SNAMA in the apoptosis pathway was investigated by looking at the interaction of SNAMA with the apoptosis suppressor Dmp53.

The CCHC type zinc finger domain of SNAMA suggests possible nucleic acid binding properties or a role in protein-protein interactions. Zinc finger proteins have been implicated in DNA or RNA binding, protein-protein interactions and membrane association. The C2H2 zinc finger is the first member of the zinc binding family of proteins. The structure of the CCHC finger of the *D. melanogaster* U-shaped protein has been reported as being similar to that of the classical C2H2 motif (Laity *et al.*, 2001). The CCHC zinc finger of *D. melanogaster* U-shaped protein mediates the functional interaction between the GATA-1 transcriptional activator and the Friend Of GATA (FOG) protein (Fox *et al.*, 1999). They have subsequently also been found in proteins involved in DNA binding.

SMART and ScanProsite identified the cysteine rich RING finger-like domain of SNAMA as a RING finger. A closer look at this domain indicates that the RING finger does not have a histidine in the fourth position that is typical of RING finger domains, but has a serine residue substitution at this position (amino acid 235) (Mather *et al.*, 2005). The SNAMA RING finger-like domain is unique and may form a new group of the variant RING fingers. These types of RING fingers have a histidine substituted with a functionally conserved amino acid. In the case of

SNAMA this is a histidine to serine substitution. This histidine substitution is observed in all of the SNAMA homologues with the exception of *E. cuniculi*. Here the histidine has been conserved. The RING finger-like domain of SNAMA extends from amino acids 214 to 272. It is likely that this domain functions as a RING finger with ubiquitin ligase ability and would enable SNAMA to modify other proteins or to catalyse auto attachment of the SNAMA domain to itself. However, it has not been experimentally determined what effect the histidine to serine substitution at position four of the RING finger (amino acid position 235) will have on its ability to coordinate zinc and maintain ubiquitin ligase activity. It is not uncommon to find the appropriation of a new set of residues to stabilise this domain in proteins, along with the loss of the original metal chelating residues. The U-box is one such example of a modified version of the RING finger domain. The U-box lacks the hallmark metal chelating cysteine and histidine residues of RING fingers. They are structurally similar to RING domains and are likely to function similar to the RING fingers (Aravind and Koonin, 2000). The U-box scaffold is stabilised by salt-bridges and hydrogen bonding rather than by zinc ion coordination, and it is possible that these interactions stabilise the RING of SNAMA. The metal chelating residues are replaced by other amino acids in U-box proteins. In the case of the U-box of the human UIP5 protein, a serine residue replaces the cysteine of the RING finger domain (Pringa *et al.*, 2001). The U-box of UIP5 is able to form physical interactions with E2 enzymes in the same way that RING fingers function. This indicates that a serine substitution is capable of maintaining the three-dimensional structure of the U-box domain and would be able to maintain the three-dimensional structure of the

RING finger domain when the histidine is replaced by serine in the RING of SNAMA. Alignment of the SNAMA homologues shows that this position is occupied by various other polar amino acids such as lysine, asparagine and arginine. The suggestion that the cysteine rich region of SNAMA functions as a RING finger is supported by the fact that that substitution in the RING domain has been observed in proteins such as Mdm2, RBQ-1 and CART1. These substitutions do not affect the three dimensional structure of the RING finger and in the case of Mdm2 the protein maintains ubiquitin ligase activity. The serine substitution is in the second zinc coordination site. This site shows a degree of flexibility for zinc binding, and since site one in SNAMA remains intact it is possible that only site one binds zinc to maintain the integrity of the RING domain. The serine substitution is also a conservative replacement of histidine and could also possibly ligate zinc due to ionisation of the hydroxyl group on the serine residue similar to the chemistry with threonine in Mdm2. Other examples of substitutions in the RING domain include the RBR family of proteins. A phylogenetic analysis of these proteins identified a substitution of the characteristic histidine in the C-terminal RING finger domain by a range of amino acids in 28 proteins (Marin and Ferrus, 2002). The RBR proteins play a role in ubiquitination. The protein dTopors is an E3 ligase with a RING finger (Secombe and Parkhurst, 2004). The first cysteine of the RING finger in dTopors is essential for E3 activity. RING fingers activate E2-dependent ubiquitination, which suggest a general role of these proteins in the ubiquitin pathway (Lorick *et al.*, 1999). The E3 proteins containing RING finger domains do not form thiol conjugates with ubiquitin but rather allosterically activate E2 to transfer ubiquitin to the substrate

lysine (Seol *et al.*, 1999). Several nuclear RING finger proteins are involved in the regulation of chromatin structure and RNA polymerase II-mediated transcription. The presence of a modified RING finger on SNAMA and evidence that this cysteine rich region could fold as a RING finger suggests that SNAMA has E3 ubiquitin ligase activity or may be an E4 type ubiquitin ligase. SNAMA may also be involved in transcription.

6.2 THE SEQUENCE ANALYSIS OF *SNAMA*

The cDNA for *SNAMA* was received from the Berkley *Drosophila* Genome Project. The cDNA is 3937 bp long with an open reading frame of 3783 bp coding for 1231 amino acids. The cDNA of *SNAMA* was compared to the *Drosophila* genomic sequence to identify the exon-intron organisation. *SNAMA* consists of nine exons and eight introns (Figure 30). The Northern blot detected a single transcript at approximately 4.6 kbp in all the developmental stages. The difference of almost 0.8 kbp between the mRNA and the observed ORF can be accounted for by the 5′ and 3′ untranslated regions. The high levels of expression during early developmental stages of the embryo indicates the vital role this gene plays in development since very few genes are expressed in the first three hours after fertilisation. Transcripts are also markedly reduced in the adult flies when the organism's body pattern is established. Again we notice a requirement for the *SNAMA* gene during stages of development when apoptosis is widely spread to mould the organism into its final form.

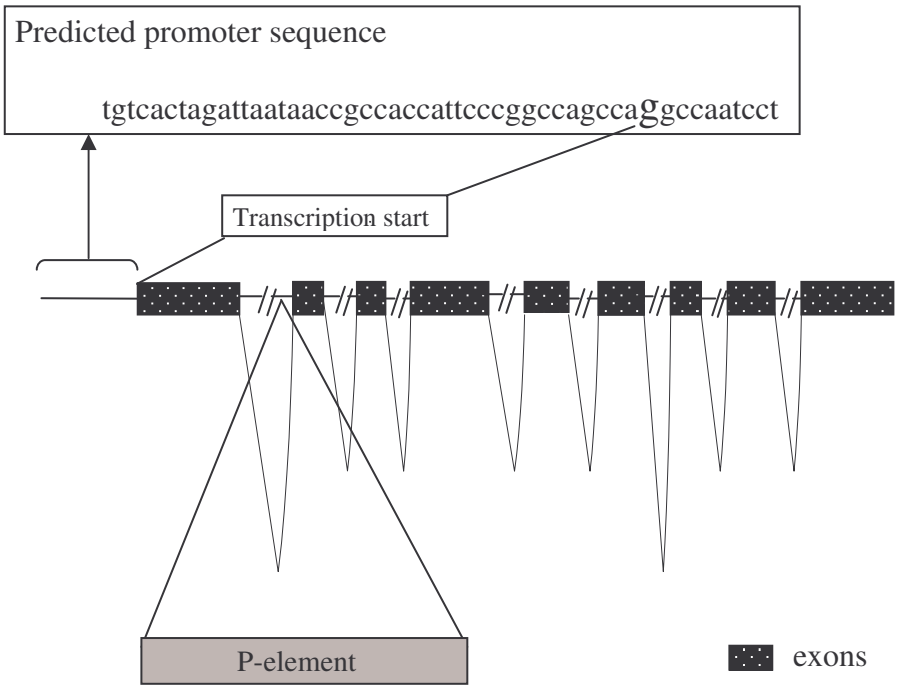


Figure 30: A schematic representation of the genomic organisation of the *SNAMA* gene (Mather *et al.,* 2005). The exons are indicated by filled boxes interrupted by the introns. Introns 1 and 6 are much larger than the other introns. The first intron is in the 5′ UTR. The position of the P-element insertion in *l(2)rQ313* is also indicated. The P-element was inserted into the first intron.

These transcripts were also observed after amplification with primers complementary to the DWNN domain. The RT-PCR reaction is a more sensitive technique than Northern blot analysis and was able to detect the low levels of expression of *SNAMA* in adult flies. In addition, initial screening experiments detected two transcripts after PCR. This evidence seemed convincing since two transcripts exist for the SNAMA mammalian homologues. Sequence alignments of these two transcripts with *SNAMA* genomic DNA identified the larger transcript as having sequence that aligns to intron one. The sequence of the 234 bp fragment aligned perfectly with the cDNA. The larger amplicon included intron sequence with an in frame internal stop codon that could not code for a protein of the larger size. Repeating the RT-PCR then identified only a single transcript in embryos consistent with what was observed on Northern blots. The larger amplicon was amplification of genomic DNA that was present in the RNA samples.

The predicted molecular weight of SNAMA is 139 kDa. This size was never observed on Western blots. Typically Western blot analysis indicated a banding pattern with proteins ranging from 19 kDa to 118 kDa. The SNAMA mammalian homologues have two transcripts, DWNN-13 and DWNN-200. The Western blot was unable to identify bands at these positions. It is possible that the banding pattern observed is a result of SNAMA degradation or DWNN of SNAMA being attached by some yet unidentified mechanism to other proteins. This attachment would require the ubiquitin ligase activity of the RING finger-like domain together with the

ubiquitin-like fold of the DWNN domain. The banding pattern on the Western blot is typical for ubiquitinated proteins suggesting that SNAMA attaches the N-terminal DWNN domain to a range of target proteins.

A putative promoter sequence and a transcriptional start site was predicted at 2.2 kbp from the translational start site (Mather *et al.*, 2005). This could not be verified experimentally, though its likelihood was examined using information of known *Drosophila melanogaster* promoters and their regulatory regions. More evidence about the nature of the 5′ region of *SNAMA* mRNA is required before the regulatory regions can be delineated conclusively. This research is continuing and involves ligating the putative promoter into a promoter-less reporter vector and observing if the promoter is capable of driving expression of the reporter gene (Hull, personal communication).

SNAMA also contains a region rich in arginine-serine/serine-arginine dipeptides. There are 33 RS or SR dipeptides with most of these repeats clustered towards a specific region of the protein (Figure 10). RS-rich proteins are involved in mRNA splicing and include proteins such as SRP75 (Zahler *et al.*, 1993) and CARS-Cyp (Mattioni *et al.*, 1992; Khan *et al.*, 1994). However, SNAMA lacks the consensus RNA-binding domain (CS-RBDs) found in the RS proteins (Fu, 1995). Other RS proteins have been found lacking the CS-RBDs. These proteins are either kinases (e.g. Clk-1) or helicases (e.g. HRH-1) and also function in the regulation of mRNA splicing (Colwill *et al.*, 1996; Ohno and Shimura, 1996). Of particular interest is the

human topoisomerase I binding protein named Topors. This protein has an arrangement of domains similar to that of SNAMA. Topors has a RING domain, RS domains and bipartite nuclear targeting domains. This suggests that Topors is involved in transcription, possibly recruiting topoisomerase I to RNA polymerase II transcriptional complexes (Haluska *et al.*, 1999). We have already established that the SNAMA yeast homologue Mpe1 is involved in transcriptional processing. The PACT protein has also been shown to include an RS region. This suggests that SNAMA could belong to a large group of RS related proteins (described by Fu, 1995) and may play a role in mRNA processing. The RS domain has been found to cause aggregation of the RS containing proteins in 20 mM $MgCl_2$ (Zahler *et al.*, 1992) and this may account for the difficulty experienced in SNAMA protein expression (Zungu, 2003). The SR domains are positively charged and may undergo functionally significant chemical modifications *in vivo*. The RS domain is known to be a site in which serine residues are phosphorylated by a kinase specific for this region (Gui *et al.*, 1994a; Gui *et al.*, 1994b). Phosphorylation of the RS region of RS family splicing factors regulates splice site recognition. The *Drosophila* Topors homologue has been named dTopors (Secombe and Parkhurst, 2004). It does not align with SNAMA although they share a RS region and a RING finger domain. dTopors is similar to human Topors, the p53 binding protein p53BP3 (Zhou *et al.*, 1999) and to LUN (Chu *et al.*, 2001). This protein binds specifically to the basic region of Hairy; a pair rule gene involved in transcriptional repression and also has ubiquitin-protein isopeptide ligase activity *in vitro*. There is a reduced density of RS repeats in dTopors and it lacks the consensus RNA-binding domain suggesting that although the RS region of

dTopors may be involved in protein-protein interactions or phosphorylation, it is unlikely to be a direct regulator of mRNA splicing. The RS region of dTopors has been suggested to act as a regulatory phosphor domain to affect the E3 activity of dTopor. It is likely that SNAMA, too, is not directly involved in mRNA splicing but does play a role in splicing and that these RS regions act as regulatory phosphor domains that regulates the possible E3 activity of the protein. There exist two reported connections between the tumor suppressor proteins and the pre-mRNA splicing machinery. This connection would then be possible in SNAMA, which has been shown to play a role in apoptosis by interacting with Dmp53. In addition the presence of the RS region and related evidence establishes a possible function in mRNA processing.

The lysine rich regions of SNAMA are also found in PACT. These regions may undergo chemical modification *in vivo*. We have no data to support this, however, this domain is reminiscent of the basic N-termini of several histones. Core histones are known to undergo cell cycle dependent acetylation and deacetylation. These modifications may cause a differential affinity of the nucleosomes to DNA, allowing differential chromatin packaging throughout the cell cycle. The positively charged lysine rich region in SNAMA is adjacent to the C-terminal and has relatively low secondary structure (DNAMAN programme). Upon acetylation of this area in PACT the protein develops a neutral charge overall and may alter the functional capabilities of the protein (Simons *et al.*, 1997). Acetylation of the SNAMA protein needs to be investigated.

The presence of a potential nuclear localisation signal suggests that SNAMA may be a nuclear protein. To investigate this we looked at *In situ* staining of embryos. The SNAMA protein localises non-homogeneously in the nucleus and is marked by an area of staining. This nuclear staining is similar to the nuclear localisation patterns exhibited by RS-splicing proteins (Misteli *et al.*, 1997) and RING finger proteins such as the viral protein family ICP0 (Everett and Maul, 1994) that are implicated in transcriptional control.

6.3 THE CYTOLOGICAL POSITION OF SNAMA

Polytene chromosomes are found in the salivary glands of *Drosophila.* They consist of a hundred or more copies of DNA arranged side by side. The high local DNA concentration of aligned DNA sequences in polytene chromosomes makes them ideal targets for *in situ* hybridisation with specific sequence probes. The banding pattern that is observed is an indicator of genomic organisation and reflects the position of specific DNA sequences. The bands also provide a tool to assess major chromosomal rearrangements that accompany species divergence (Pardue, 1986). The SNAMA gene has been localised by *in situ* hybridisation to the polytene chromosomes and correspond to position 60B9 on the right arm of chromosome 2 (Rubin *et al.*, 2000).

6.3.1 GENES OCCUPYING POSITION 60B

The SNAMA homologous proteins have DNA and RNA binding ability and are involved in protein-protein interactions with transcription factor p53 and with the Rb protein. Transcriptional factors and chromosomal proteins are found localised to various bands on the chromosome. Localisation of these proteins to specific regions is thought to have a functional significance and to reflect the activity of these proteins in gene regulation. It is for that reason that it was necessary to identify genes in the region of *SNAMA*. Several genes occupy position 60B8 to 60B12; the region flanking *SNAMA* on the chromosome. These genes include *Mlp60A*, *gek*, *enok*, *tamo*, and *Zfrp8*. We will discuss and assess the relationship of these genes to *SNAMA*.

6.3.1.1 Muscle LIM protein at 60A (Mlp60A)

Muscle LIM proteins (Mlp) is a conserved LIM finger protein found in striated muscle. They are essential promoters of myogenesis and play a role in cell differentiation (Arber *et al.*, 1994). They control muscle differentiation by binding to the myogenic regulatory factors myocyte enhancer-binding 2 (MEF2) protein (Stronach *et al.*, 1999). The muscle LIM protein at 60A(Mlp60A) protein is involved in the regulation of myogenesis in *Drosophila*. Mlp60A contains the LIM domain, which is a zinc-binding domain, found in a number of eukaryotic proteins that regulate cell growth and differentiation during development. *Mlp60A* encodes a 96 amino acid protein with a single LIM domain linked to a glycine-rich region.

Expression is spatially restricted to larval somatic muscle, larval visceral muscle, pharyngeal muscle, somatic mesoderm and visceral mesoderm. *Mlp60A* transcripts are found throughout the muscle fibres. Gene expression is developmentally regulated with a biphasic pattern over the course of the *Drosophila* life cycle (Stronach *et al.*, 1996). It has been mapped cytologically to position 60B8.

6.3.1.2 Genghis Khan (Gek)

The Rho family of small GTPases consists of the Rho, Rac and Cdc42 subfamilies. This family of proteins controls morphogenetic processes that involve the remodeling of the actin cytoskeleton by acting as molecular switches (Mizuno *et al.*, 1999). Cdc42 regulates a variety of biological processes including actin polymerisation, cell proliferation, and JNK/mitogen-activated protein kinase activation via specific effectors. Genghis Khan (Gek), a multidomain protein kinase, is a downstream effector of *D. melanogaster* Cdc42 and a regulator of actin polymerisation (Luo *et al.*, 1997). Gek binds to Dcdc42 in a GTP-dependent and effector domain-dependent manner. Gek is a protein of 1613 amino acids which contains a serine/threonine kinase catalytic domain, a coiled-coil domain, a cysteine-rich domain similar to the phorbol ester binding domain, a pleckstrin homology domain which is found in signalling molecules, and the Cdc42 binding consensus sequence. *Gek* has been mapped cytologically to position 60B9-60B10.

6.3.1.3 Enoki mushroom (Enok)

Enok encodes a product with histone acetyltransferase (HAT) activity of the MYST family of proteins involved in neuroblast proliferation. This family of proteins is responsible for the acetylation of histones. The significance of this modification is to regulate processes such as gene repression, gene activation and replication. Lysines at the N-terminal ends of the histones are the sites of acetylation (Pandey *et al.*, 2002). Enok is 2291 amino acids long containing a zinc finger C2H2 type, a histone H1 and H5 family and a PHD finger. It has been mapped cytologically to position 60B10-60B11. Enok is required for mushroom body neuroblast proliferation.

6.3.1.4 Tamo

Tamo is a cytoplasmic protein that binds to *Drosophila* Rel protein Dorsal. Tamo functions during oogenesis and controls the levels of the nuclear protein Dorsal in early embryos. It also regulates the accumulation of Dorsal in the nucleus after immune challenge. Tamo negatively regulates Dorsal nuclear targeting. Tamo interacts with Dorsal via the nuclear localisation signal (NLS). The NLSs are 6 amino acids that bind adapter proteins such as importins (Nakielny and Dreyfuss, 1999). *Tamo* encodes three different transcripts; a 4.1 kbp transcript expressed throughout development, a 3.6 kbp transcript present mainly in ovaries and early embryos and a 1.2 kbp transcript which is highly expressed in third instar larvae and weakly in adults. The 3.6 kbp and 4.1 kbp mRNAs encode the same polypeptide of 797 amino

acids with a predicted molecular weight of 89 kDa (Minakhina *et al.*, 2003). The protein has been mapped cytological to 60B11. Tamo has a zinc finger domain at the C-terminus of the protein. This domain interacts with Ran on Ran GDP and is thus likely to function in nuclear import, and may be regulated by the Ran-GDP-GTP gradient.

6.3.1.5 Zinc finger RP-8 (Zfrp8)

Zfrp8 encodes a product that has DNA binding ability and is a component of the nucleus. *Zfrp8* encodes a 347 amino acid protein that is an apoptotic gene. The gene has been mapped cytologically to position 60B12 (Owens *et al.*, 1991).

None of these genes and their translational products matched the nucleotide and protein sequence, respectively, of *SNAMA* during searches of the databases. The only significant similarity to SNAMA is the presence of the zinc finger in *Mlp60A, Enok, Tamo* and *Zfrp-8*. Zinc finger domains occur in a variety of eukaryotic transcription factors, in *Drosophila* developmental regulators, and certain proto-oncogene proteins. It is evident that the genes occupy position 60B8 - 60B12, the region flanking *SNAMA* on the chromosome codes for proteins that are all regulatory proteins involved in the development of the organisms. This suggests that SNAMA plays a role in embryo development acting as a regulatory protein in an unknown pathway.

6.4 THE BIOLOGICAL FUNCTION OF SNAMA

6.4.1 THE DWNN DOMAIN OF SNAMA HAS A UBIQUITIN-LIKE FOLD

It has been predicted that the N-terminal 76 amino acid residue of the DWNN domain of SNAMA has a ubiquitin-like fold (chapter 3.3.3). Preliminary NMR structural analysis of the human DWNN domain also indicates a folded structure similar to that of ubiquitin (Pugh, personal communication). This identifies SNAMA has a protein possibly functioning in the ubiquitin pathway. It should however be noted that the DWNN domain lacks the double glycine motif that is required both for ubiquitin precursor processing and for ubiquitin conjugation. The ubiquitin molecule consists of a five-stranded β-sheet and a α-helix and a flexible tail formed by four protruding residues. The predicted fold of the N-terminal region of SNAMA has this typical fold, but lacks some important surface residues essential for ubiquitin conjugation and deubiquitination (Figure 31(A)). The ubiquitin tail consists of the essential residues Leu73, Arg74, Gly75, and Gly76 (Sloper-Mould *et al.*, 2001). Gly75 and Gly76 are essential for ubiquitin conjugation and Arg74 is essential but not important for E1 interaction or ubiquitin conjugation. Arg74 and Leu73 play minor roles in endocytosis. Isoleucine and proline residues replace the Gly75 and Gly76, respectively, in the DWNN domain of SNAMA. The proline imparts restrictions on the three-dimensional structure of the protein. Pro76 of SNAMA is observed in most of the homologues and is observed in a yeast ubiquitin-like protein DSK2 where it prevents cleavage of the ubiquitin-like domain by specific hydrolases (Vucic *et al.,*

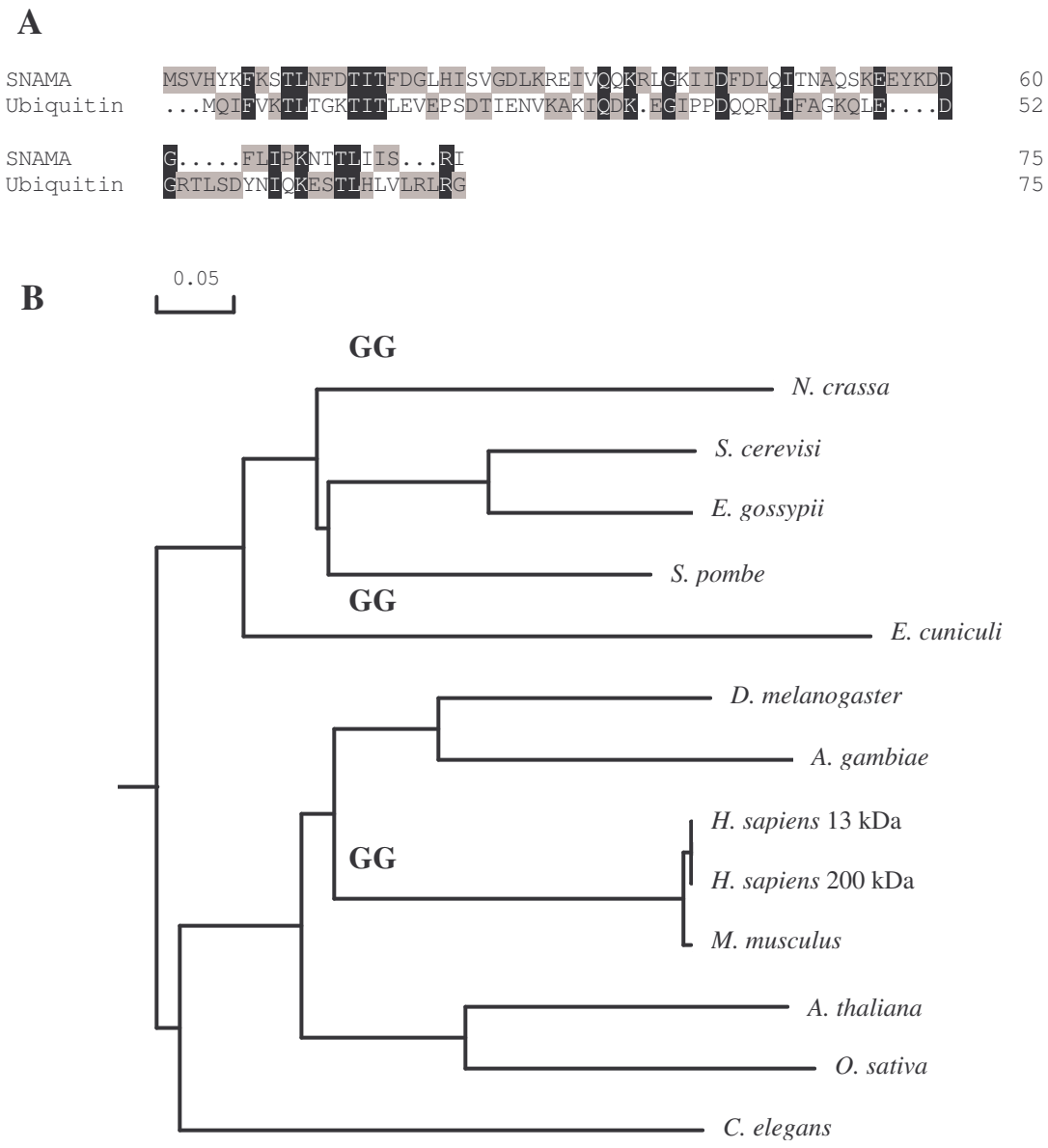


Figure 31: (A) Ubiquitin amino acid sequence and a comparison with the DWNN domain of SNAMA. Conserved residues are indicated in black. SNAMA lacks the di-glycine residues at the C-terminal end, which has been replaced by an isoleucine and a proline. **(B)** Neighbour-joining (NJ) phylogenetic tree of the ubiquitin-like domain of the DWNN family of proteins. Proteins with the C-terminal di-glycine residues are indicated by GG.

1998). The Arg74 of ubiquitin is conserved in SNAMA. Although many of the SNAMA homologues have a proline, there are some homologues that have a di-glycine near the proline and this appears to be an evolutionary event (Figure 31(B)). The DWNN proteins from animals, plants and fungi form groups. None of the plant proteins have a di-glycine. The fungi diversified independently with two examples that have a di-glycine (the microsporidea and a ascomycota pezizomycotina). In animals, metazoa chordata have the di-glycine while the arthropoda do not. It appears that the more complex metazoan have diverged to maintain the di-glycine essential for conjugation (Table 11). The presence of the Pro76 of SNAMA suggests that the ubiquitin-like domain does not modify other proteins, at least not in a manner that requires cleavage by the known hydrolases. Alternatively, it could mean that it is cleaved by a different and as yet uncharacterised set of enzymes.

Lysine residues at positions 29, 48 and 63 of ubiquitin have been implicated in the formation of polyubiquitin chains in *S. cerevisiae*. The DWNN domain of SNAMA has eight lysine residues at positions 6, 8, 29, 36, 40, 54, 58 and 66 capable of forming polyubiquitin chains. Lys48 and Lys63 are also important for targeting proteins for 26S proteasome degradation and for mediating cellular functions such as DNA repair and intracellular cell sorting. The E3 enzymes conjugate the C-terminal glycine of one ubiquitin molecule to the internal Lys48 of the adjacent ubiquitin molecule. A Lys48→Arg mutation is unable to mediate ubiquitination by E3 enzymes (Hatakeyama *et al.*, 2001) but dependence on this lysine residue appears to differ amongst U-box proteins.

Table 11: Presence of the di-glycine residues at the C-terminal end of the DWNN domain in SNAMA homologous proteins. Organisms have been arranged in order of complexity to see if any evolutionary pattern exists.

ORGANISM	PHYLA	PRESENCE OF DI-GLYCINE
Saccharomyces cerevisiae	Fungi; Ascomycota; Saccharomycotina	NO
Eremothecium gossypii	Fungi; Ascomycota; Saccharomycotina	NO
Schizosaccharomyces pombe	Fungi; Ascomycota; Schizosaccharomycetes	NO
Neurospora crassa	Fungi; Ascomycota; Pezizomycotina	YES
Encephalitozoon cuniculi	Fungi; Microsporidia	YES
Arabidopsis thaliana	Viridiplantae	NO
Oryza sativa	Viridiplantae	NO
Caenorhabditis elegans	Metazoa; Nematoda	NO
Drosophila melanogaster	Metazoa; Arthropoda	NO
Anopheles gambiae	Metazoa; Arthropoda	NO
Mus musculus	Metazoa; Chordata	YES
Homo sapiens	Metazoa; Chordata	YES

Polyubiquitin chains that are assembled through lysine residues of ubiquitin other than Lys48 may function in distinct biological processes and not only in proteolysis (Hatakeyama and Nakayama, 2003). SNAMA lacks the conserved Lys48 but retains the Lys63 required for poly-ubiquitin chain formation and targeting modified proteins for 26S proteasome degradation. The Leu8, Ile44 and Val70 amino acids in ubiquitin are collectively known as the hydrophobic patch and are critical for proteasomal degradation. Leu8 and Ile44 in SNAMA are conserved. Val70 has been substituted by a hydrophobic isoleucine conserving the overall hydrophobic patch of the ubiquitin-like DWNN domain essential for proteasomal binding and degradation. The amino acids essential for endocytosis Phe4 and Thr12 are conserved in SNAMA. Taken together, this comparison supports our hypothesis that the DWNN domain of SNAMA could be involved in poly chain formation, being transferred to proteins in a manner similar to ubiquitin either targeting them for proteasomal degradation or some other biological process. The mode of transfer is not yet understood.

The ubiquitin domain proteins (UDP) have a ubiquitin-like domain within a large protein. The ubiquitin-like domains of these proteins do not always end in a di-glycine and this domain is not processed by proteases. The domain is not released as in ubiquitin but these proteins are linked to the ubiquitin pathway. An interesting example of a UDP is UIP28. UIP28 does not have a di-glycine at the C-terminal end of the ubiquitin-like domain. The domain organisation is similar to that of SNAMA with a ubiquitin-like domain, a zinc finger domain and a RING finger domain. UIP 28 interacts with the ubiquitin-conjugating enzyme UbcM4. The domain arrangement

of SNAMA, as in UIP28, may indicate a role for SNAMA as a UDP interacting with a ubiquitin-conjugating enzyme.

6.4.2 SNAMA IS INVOLVED IN APOPTOSIS

Initial experiments that led to the identification of the DWNN proteins were based on the screening of a mutated CHO cell line for resistance to apoptosis inducing agents. The resistance developed because of disruption of DWNN expression. Subsequently the mammalian SNAMA homologues have been shown to bind Rb and p53. A C-terminal region of SNAMA aligns with the mammalian PACT protein suggesting that SNAMA could interact with proapoptotic molecules such as Dmp53 and RBF. These are the *D. melanogaster* homologues of p53 and Rb, respectively. The RBF protein acts downstream of dE2F; a transcription factor that regulates Dmp53 (Du and Dyson, 1999). To establish whether SNAMA plays a role in apoptosis, I took advantage of a P-element insertion fly line *l(2)rQ313*. Homozygous offspring develop to late embryonic stages and then die because of the P-element insertion in the first intron of the SNAMA gene. This gene codes for protein that is vital for embryonic development. *l(2)rQ313* embryos were stained with the vital dye acridine orange. The staining pattern indicated abnormal ectopic apoptosis in late developmental stage embryos and early occurrence of apoptosis in the blastoderm as compared to wild-type flies. The blastoderm is characterised by rapid nuclear division with very limited apoptosis. When the P-element was excised, the staining pattern resembles that of wild-type embryos. Further evidence of a link between apoptosis

and development is observed in Northern blots where there are reduced levels of SNAMA mRNA expression in late stages of development (12-18hrs) when normal apoptosis in the embryo is taking place.

Western analysis of embryonic, larval and adult protein extracts identified bands recognised by both anti-Dmp53 and anti-SNAMA. This banding pattern is more dramatic at larval stages when significant apoptosis takes place. This indicated a possible association between Dmp53 and SNAMA. The Western blots also identified bands not recognised by the anti-Dmp53 antibody, which suggest that SNAMA interacts with other proteins in addition to Dmp53 because of the immunoreactive complexes formed with the anti-SNAMA antibody. This is consistent with the suggested ubiquitin-like function of SNAMAs DWNN domain. These bands could also possibly be degradation products of SNAMA given the presence of proteasomal recognition patches on the DWNN domain. But there is no C-terminal di-glycine in the DWNN domain necessary for conjugation. It is however possible that there exists some other, yet unidentified, alternative mechanism that allows cleavage of the DWNN domain at the conserved Pro76 freeing the domain to associate with Dmp53. An alternative to this mechanism would be that SNAMA associates with Dmp53 forming a bond that is resistant to boiling in SDS. Such associations have been observed in the UBL (ubiquitin-like protein) Hub1 (Luders *et al.*, 2003). Although Hub1 lacks a C-terminal tail bearing the glycine residue it associates tightly with other proteins.

Given that no Mdm2 homologue has been identified in *D. melanogaster*, SNAMA is a good candidate regulator of Dmp53 though the mechanism of this regulation has not been established. Mdm2 in mammals represses p53 or promotes p53 degradation via the proteasomal pathway. The importance of negative regulation of p53 by Mdm2 was demonstrated in two similar but independent experiments using homozygotes for Mdm2 and p53 null alleles (Jones *et al.*, 1995; Montes *et al.*, 1995). Mice deficient for both p53 and Mdm2 develop normally and resemble wild-type mice while those deficient in Mdm2 only die during embryogenesis. In *l(2)rQ313* embryos apoptosis is detected very soon after egg laying and homozygous embryos die. SNAMA may be the negative regulator of Dmp53 similar to the function of Mdm2 in mammals. Dmp53 null homozygotes have normal phenotypes similar to that seen in mammals. It should be noted that Dmp53 controls only apoptosis and not cell cycle arrest like the p53 protein (Sogame *et al.*, 2003). Interaction of SNAMA with Dmp53 is also observed in Western blots on protein extracts from lavae that were fed with Camptothecin, an activator of the Dmp53 protein. When the Dmp53 pathway is activated by the topoisomerase inhibitor, Camptothecin, the levels of SNAMA expression increase and the anti-SNAMA antibody forms immunoreactive complexes with both the Dmp53 and SNAMA indicating possible interactions between these proteins. Immunoprecipitation experiments using anti-Dmp53 and anti-SNAMA antibody also strongly suggest that SNAMA interacts with Dmp53 (Rakgotho, personal communication).

6.5 CONCLUSIONS

SNAMA is the *D. melanogaster* homologue of a group of proteins with interesting sequence features including a conserved N-terminal DWNN domain that has a ubiquitin-like fold, a zinc finger, a RING finger-like domain, a probable p53 binding region, and glutamic acid-rich and lysine-rich regions. The domains that associate with SNAMA indicate that this protein plays an important regulatory role in the cell and may function in RNA processing and in apoptosis. SNAMA may interact with Dmp53 probably as a suppressor of apoptosis or a negative regulator of an activator of apoptosis. These results are supported by abnormal patterns of apoptosis in *l(2)rQ313* embryos, the domain configuration, interactions with Dmp53, and increased levels of protein when Dmp53 is activated. The presence of the ubiquitin-like DWNN domain and the cysteine rich RING finger-like region also suggests that SNAMA could be a ubiquitin domain protein with ligase activity. This would enable the domain to be cleaved from SNAMA and attached to a subset of cellular proteins posttranslationally. The exact function of the SNAMA system has got to be determined. The identification of possible substrates should identify its possible physiological roles.

7: REFERNCES

Abrams, J. M. (1999). An emerging blueprint for apoptosis in *Drosophila*. *Trends Cell Biol.* **9**, 435-440.

Abrams, J. M., White, K., Fessler, L. I. and Steller, H. (1993). Programmed cell death during *Drosophila* embryogenesis. *Development* **117**, 29-43.

Adams, J. M. and Cory, S. (1998). The Bcl-2 protein family: arbiters of cell survival. *Science* **281**, 1322-1326.

Aravind, L. and Koonin, E. V. (2000). The U box is a modified RING finger - a common domain in ubiquitination. *Curr. Biol.* **10**, R132-R134.

Aravind, L., Dixit, V. M. and Koonin, E. V. (1999). The domains of death: evolution of the apoptosis machinery. *Trends Biochem Sci* **24**, 47-53.

Aravind, L., Dixit, V. M. and Koonin, E. V. (2001). Apoptotic molecular machinery: vastly increased complexity in vertebrates revealed by genome comparisons. *Science* **291**, 1279-1284.

Arber, S., Halder, G. and Caroni, P. (1994). Muscle LIM protein, a novel essential regulator of myogenesis, promotes myogenic differentiation. *Cell* **79**, 221-231.

Arnason, T. and Ellison, M. J. (1994). Stress resistance in Saccharomyces cerevisiae is strongly correlated with assembly of a novel type of multiubiquitin chain. *Mol. Cell. Biol.* **14**, 7876-7883.

Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A. and Struhl, K. (1995). *Short Protocols in Molecular Biology*. John Wiley & Sons, Inc.

Bangs, P. and White, K. (2000). Regulation and execution of apoptosis during *Drosophila* development. *Dev Dyn* **218**, 68-79.

Barabino, S. M., Hubner, W., Jenny, A., Minvielle-Sebastia, L. and Keller, W. (1997). The 30-kD subunit of mammalian cleavage and polyadenylation specificity factor and its yeast homolog are RNA-binding zinc finger proteins. *Genes & Dev.* **11**, 1703-1716.

Barnes, W. M. (1994). PCR Amplification of up to 35-kb DNA with High Fidelity and High Yield from {lambda} Bacteriophage Templates. *PNAS* **91**, 2216-2220.

Bateman, A., Coin, L., Durbin, R., Finn, R. D., Hollich, V., Griffiths-Jones, S., Khanna, A., Marshall, M., Moxon, S., Sonnhammer, E. L. L. et al. (2004). The Pfam protein families database. *Nucleic Acids Res.* **32**, D138-D141.

Bender, W., Spierer, P. and Hogness, D. S. (1983). Chromosomal walking and jumping to isolate DNA from the Ace and rosy loci and the bithorax complex in *Drosophila melanogaster*. *J. Mol. Biol* **168**, 17-33.

Berg, J. M. and Shi, Y. (1996). The Galvanization of Biology: A Growing Appreciation for the Roles of Zinc. *Science* **271**, 1081-1085.

Bergmann, A., Agapite, J., McCall, K. and Steller, H. (1998). The *Drosophila* gene *hid* is a direct molecular target of Ras-dependent survival signaling. *Cell* **95**, 331-341.

Bernstein, S. I., Hansen, C. J., Becker, K. D., Wassenberg 2nd, D. R., Roche, E. S., Donady, J. J. and Emerson Jr, C. P. (1986) Alternative RNA splicing generates transcripts encoding a thorax-specific isoform of *Drosophila melanogaster* myosin heavy chain. *Mol Cell Biol* **6**, 2511-2519.

Biggins, S., Ivanovska, I. and Rose, M. D. (1996). Yeast ubiquitin-like genes are involved in duplication of the microtubule organizing center. *J Cell Biol* **133**, 1331-1346.

Birnboim, H. C. and Doly, J. (1979). A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic Acids Res.* **7**, 1513-1523.

Boddy, M. N., Freemont, P. S. and Borden, K. L. (1994). The p53-associated protein MDM2 contains a newly characterized zinc-binding domain called the RING finger. *Trends Biochem Sci* **19**, 198-199.

Borden, K. L. (2000). RING domains: master builders of molecular scaffolds? *J. Mol. Biol.* **295**, 1103-1112.

Borden, K. L. and Freemont, P. S. (1996). The RING finger domain: a recent example of a sequence-structure family. *Curr. Opin. Struct. Biol* **6**, 395-401.

Bradford, M. M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* **72**, 248-254.

Brodsky, M. H., Nordstrom, W., Tsang, G., Kwan, E., Rubin, G. M. and Abrams, J. M. (2000). *Drosophila* p53 binds a damage response element at the reaper locus. *Cell* **101**, 103-113.

Buschmann, T., Fuchs, S. Y., Lee, C. G., Pan, Z. Q. and Ronai, Z. (2000). SUMO-1 modification of Mdm2 prevents its self-ubiquitination and increases Mdm2 ability to ubiquitinate p53. *Cell* **101**, 753-762.

Campos-Ortega, J. A. and Hartenstein, V. (1997). *The embryonic development of Drosophila melanogaster*. Springer-Verlag, Germany, Berlin.

Cavener, D. R. (1987). Comparison of the consensus sequence flanking translational start sites in *Drosophila* and vertebrates. *Nucleic Acids Res.* **15**, 1353-1361.

Cenci, G., Rawson, R. B., Belloni, G., Castrillon, D. H., Tudor, M., Petrucci, R., Goldberg, M. L., Wasserman, S. A. and Gatti, M. (1997). UbcD1, a Drosophila ubiquitin-conjugating enzyme required for proper telomere behavior. *Genes & Dev.* **11**, 863-875.

Chang, W., Hubbard, S. C., Friedel, C. and Ruley, H. E. (1993). Enrichment of insertional mutants following retrovirus gene trap selection. *Virology* **193**, 737-747.

Chen, J., Marechal, V. and Levine, A. J. (1993). Mapping of the p53 and mdm-2 interaction domains. *Mol. Cell. Biol.* **13**, 4107-4114.

Chen, P. and Abrams, J. M. (2000). *Drosophila* apoptosis and Bcl-2 genes: outliers fly in. *J. Cell Biol.* **148**, 625-627.

Chen, P. L., Scully, P., Shew, J. Y., Wang, J. Y. and Lee, W. H. (1989). Phosphorylation of the retinoblastoma gene product is modulated during the cell cycle and cellular differentiation. *Cell* **58**, 1193-1198.

Chen, P., Lee, P., Otto, L. and Abrams, J. (1996a). Apoptotic activity of REAPER is distinct from signaling by the tumor necrosis factor receptor 1 death domain. *J. Biol. Chem.* **271**, 25735-25737.

Chen, P., Nordstrom, W., Gish, B. and Abrams, J. M. (1996b). *grim*, a novel cell death gene in *Drosophila*. *Genes Dev* **10**, 1773-1782.

Chen, P., Rodriguez, A., Erskine, R., Thach, T. and Abrams, J. M. (1998). Dredd, a novel effector of the apoptosis activators reaper, grim, and hid in Drosophila. *Dev Biol.* **201**, 202-216.

Chomczynski, P. (1993). A reagent for the single-step simultaneous isolation of RNA, DNA and proteins from cell and tissue samples. *Biotechniques* **15**, 532-537.

Chu, D., Kakazu, N., Gorrin-Rivas, M. J., Lu, H. P., Kawata, M., Abe, T., Ueda, K. and Adachi, Y. (2001). Cloning and Characterization of LUN, a Novel RING Finger Protein That Is Highly Expressed in Lung and Specifically Binds to a Palindromic Sequence. *J. Biol. Chem.* **276**, 14004-14013.

Chung, C. H. and Baek, S. H. (1999). Deubiquitinating enzymes: their diversity and emerging roles. *Biochem Biophys Res Commun* **266**, 633-640.

Ciechanover, A., Elias, S., Heller, H. and Hershko, A. (1982). "Covalent affinity" purification of ubiquitin-activating enzyme. *J. Biol. Chem.* **257**, 2537.

Ciechanover, A., Hod, Y. and Hershko, A. (1978). A heat-stable polypeptide component of an ATP-dependent proteolytic system from reticulocytes. *Biochem. Biophys. Res. Commun.* **81**, 1100-1105.

Clem, R. and Duckett, C. (1997). The *iap* genes: unique arbitrators of cell death. *Trends Cell Biol* **7**, 337-339.

Colussi, P. A., Quinn, L. M., Huang, D. C., Coombe, M., Read, S. H., Richardson, H. and Kumar, S. (2000). Debcl, a proapoptotic Bcl-2 homologue, is a component of the Drosophila melanogaster cell death machinery. *J. Cell Biol.* **148**, 703-714.

Colwill, K., Pawson, T., Andrews, B., Prasad, J., Manley, J. L., Bell, J. C. and Duncan, P. I. (1996). The Clk/Sty protein kinase phosphorylates SR splicing factors and regulates their intranuclear distribution. *EMBO J.* **15**, 265-275.

Combet, C., Blanchet, C., Geourjon, C. and Deleage, G. (2000). NPS@: network protein sequence analysis. *Trends Biochem Sci* **25**, 147-150.

Cornilescu, D., Marquardt, J. L., Ottiger, M. and Bax, A. (1998). Validation of Protein Structure from Anisotropic Carbonyl Chemical Shifts in a Dilute Liquid Crystalline Phase. *J. Am. Chem. Soc.* **120**, 6836-6837.

D'Andrea, A. and Pellman, D. (1998). Deubiquitinating enzymes: a new class of biological regulators. *Crit Rev Biochem Mol Biol* **33**, 337-352.

De, L., V and Melino, G. (2000). Apoptosis. The little devil of death. *Nature* **406**, 135-136.

DeCaprio, J. A., Ludlow, J. W., Lynch, D., Furukawa, Y., Griffin, J., Piwnica-Worms, H., Huang, C. M. and Livingston, D. M. (1989). The product of the retinoblastoma susceptibility gene has properties of a cell cycle regulatory element. *Cell* **58** , 1085-1095.

Deshaies, R. J. (1999). SCF and Cullin/Ring H2-based ubiquitin ligases. *Annu Rev Cell Dev Biol* **15:435-67.**, 435-467.

Desterro, J. M., Rodriguez, M. S. and Hay, R. T. (1998). SUMO-1 modification of IkappaBalpha inhibits NF-kappaB activation. *Mol Cell* **2**, 233-239.

Dittmar, G. A. G., Wilkinson, C. R. M., Jedrzejewski, P. T. and Finley, D. (2002). Role of a Ubiquitin-Like Modification in Polarized Morphogenesis. *Science* **295**, 2442-2446.

Dorstyn, L., Colussi, P. A., Quinn, L. M., Richardson, H. and Kumar, S. (1999). DRONC, an ecdysone-inducible *Drosophila* caspase. *Proc. Natl. Acad. Sci. U. S. A* **96**, 4307-4312.

Doumanis, J., Quinn, L., Richardson, H. and Kumar, S. (2001). STRICA, a novel Drosophila melanogaster caspase with an unusual serine/threonine-rich prodomain, interacts with DIAP1 and DIAP2. *Cell Death. Differ.* **8**, 387-394.

Du, C., Fang, M., Li, Y., Li, L. and Wang, X. (2000). Smac, a mitochondrial protein that promotes cytochrome c-dependent caspase activation by eliminating IAP inhibition. *Cell* **102**, 33-42.

Du, W. and Dyson, N. (1999). The role of RBF in the introduction of G1 regulation during *Drosophila* embryogenesis. *EMBO J.* **18**, 916-925.

Edgar, B. A. and Schubiger, G. (1986). Parameters controlling transcriptional activation during early *Drosophila* development. *Cell* **44**, 871-877.

Epps, J. L. and Tanda, S. (1998). The *Drosophila semushi* mutation blocks nuclear import of bicoid during embryogenesis. *Curr Biol* **8**, 1277-1280.

Evans, E. K., Kuwana, T., Strum, S., Smith, J. J., Newmeyer, D. and Kornbluth, S. (1997). Reaper-induced apoptosis in a vertebrate system. *EMBO J.* **16**, 7372-7381.

Everett, R. D. and Maul, G. G. (1994). HSV-1 IE protein Vmw110 causes redistribution of PML. *EMBO J.* **13**, 5062-5069.

Eytan, E., Armon, T., Heller, H., Beck, S. and Hershko, A. (1993). Ubiquitin C-terminal hydrolase activity associated with the 26 S protease complex. *J. Biol. Chem.* **268**, 4668-4674.

Fang, S., Jensen, J. P., Ludwig, R. L., Vousden, K. H. and Weissman, A. M. (2000). Mdm2 Is a RING Finger-dependent Ubiquitin Protein Ligase for Itself and p53. *J. Biol. Chem.* **275**, 8945-8951.

Figge, J. and Smith, T. F. (1988). Cell-division sequence motif. *Nature* **334**, 109.

Finley, D. and Chau, V. (1991). Ubiquitination. *Annu Rev Cell Biol* **7**, 25-69.

Fisk, H. A. and Yaffe, M. P. (1999). A Role for Ubiquitination in Mitochondrial Inheritance in *Saccharomyces cerevisiae*. *J. Cell Biol.* **145**, 1199-1208.

Fox, A. H., Liew, C., Holmes, M., Kowalski, K., Mackay, J. and Crossley, M. (1999). Transcriptional cofactors of the FOG family interact with GATA proteins by means of multiple zinc fingers. *EMBO J.* **18**, 2812-2822.

Fraser, A. G. (1999). Programmed cell death in C. elegans. *Cancer Metastasis Rev* **18**, 285-294.

Freedman, D. A. and Levine, A. J. (1998). Nuclear Export Is Required for Degradation of Endogenous p53 by MDM2 and Human Papillomavirus E6. *Mol. Cell. Biol.* **18**, 7288-7293.

Fu, X. D. (1995). The superfamily of arginine/serine-rich splicing factors. *RNA* **1**, 663-680.

Furukawa, K., Mizushima, N., Noda, T. and Ohsumi, Y. (2000). A Protein Conjugation System in Yeast with Homology to Biosynthetic Enzyme Reaction of Prokaryotes. *J. Biol. Chem.* **275**, 7462-7465.

Gao, S. and Scott, R. E. (2002). P2P-R protein overexpression restricts mitotic progression at prometaphase and promotes mitotic apoptosis. *J. Cell Physiol* **193**, 199-207.

Gao, S. and Scott, R. E. (2003). Stable overexpression of specific segments of the P2P-R protein in human MCF-7 cells promotes camptothecin-induced apoptosis. *J Cell Physiol* **197**, 445-452.

Gao, S., Witte, M. M. and Scott, R. E. (2002). P2P-R protein localizes to the nucleolus of interphase cells and the periphery of chromosomes in mitotic cells which show maximum P2P-R immunoreactivity. *J. Cell Physiol* **191**, 145-154.

George, A. E. (1995). A new method for Isolating Genes involved in the Processing and Presentation of Antigens to Cytotoxic T Cells. University of Oxford.

Gil, A. and Proudfoot, N. J. (1987). Position-dependent sequence elements downstream of AAUAAA are required for efficient rabbit beta-globin mRNA 3' end formation. *Cell* **49**, 399-406.

Glickman, M. H. and Ciechanover, A. (2002). The ubiquitin-proteasome proteolytic pathway: destruction for the sake of construction. *Physiol Rev* **82**, 373-428.

Glotzer, M., Murray, A. W. and Kirschner, M. W. (1991). Cyclin is degraded by the ubiquitin pathway. *Nature* **349**, 132-138.

Goyal, L. (2001). Cell death inhibition: keeping caspases in check. *Cell* **104**, 805-808.

Goyal, L., McCall, K., Agapite, J., Hartwieg, E. and Steller, H. (2000). Induction of apoptosis by *Drosophila* reaper, hid and grim through inhibition of IAP function. *EMBO J* **19**, 589-597.

Green, D. R. and Reed, J. C. (1998). Mitochondria and apoptosis. *Science* **281**, 1309-1312.

Grether, M. E., Abrams, J. M., Agapite, J., White, K. and Steller, H. (1995). The head involution defective gene of *Drosophila melanogaster* functions in programmed cell death. *Genes Dev* **9**, 1694-1708.

Groll, M., Ditzel, L., Lowe, J., Stock, D., Bochtler, M., Bartunik, H. D. and Huber, R. (1997). Structure of 20S proteasome from yeast at 2.4 A resolution. *Nature* **386**, 463-471.

Gui, J. F., Lane, W. S. and Fu, X. D. (1994b). A serine kinase regulates intracellular localization of splicing factors in the cell cycle. *Nature* **369**, 678-682.

Gui, J., Tronchere, H., Chandler, S. D. and Fu, X. (1994a). Purification and Characterization of a Kinase Specific for the Serine-and Arginine-Rich Pre-mRNA Splicing Factors. *PNAS* **91**, 10824-10828.

Haining, W. N., Carboy-Newcomb, C., Wei, C. L. and Steller, H. (1999). The proapoptotic function of Drosophila Hid is conserved in mammalian cells. *Proc. Natl. Acad. Sci. U. S. A* **96**, 4936-4941.

Haluska, P., Saleem, A., Rasheed, Z., Ahmed, F., Su, E. W., Liu, L. F. and Rubin, E. H. (1999). Interaction between human topoisomerase I and a novel RING finger/arginine-serine protein. *Nucleic Acids Res.* **27**, 2538-2544.

Hamilton, B. A. and Zinn, K. (1994). *Drosophila melanogaster: Practical Uses in Cell and Molecular Biology*. Acedemic Press, Inc. pp. 81-94.

Hatakeyama, S. and Nakayama, K. I. (2003). U-box proteins as a new family of ubiquitin ligases. *Biochem Biophys Res Commun* **302**, 635-645.

Hatakeyama, S., Yada, M., Matsumoto, M., Ishida, N. and Nakayama, K. I. (2001). U box proteins as a new family of ubiquitin-protein ligases. *J Biol Chem.* **276**, 33111-33120.

Hay, B. A. (2000). Understanding IAP function and regulation: a view from Drosophila. *Cell Death. Differ.* **7**, 1045-1056.

Hershko, A. and Ciechanover, A. (1998). The ubiquitin system. *Annu Rev Biochem* **67**, 425-479.

Hochstrasser, M. (1996). Ubiquitin-dependent protein degradation. *Annu Rev Genet* **30**, 405-439.

Holzl, H., Kapelari, B., Kellermann, J., Seemuller, E., Sumegi, M., Udvardy, A., Medalia, O., Sperling, J., Muller, S. A., Engel, A. et al. (2000). The regulatory complex of *Drosophila melanogaster* 26S proteasomes. Subunit composition and localization of a deubiquitylating enzyme. *J Cell Biol* **150**, 119-130.

Honda, R., Tanaka, H. and Yasuda, H. (1997). Oncoprotein MDM2 is a ubiquitin ligase E3 for tumor suppressor p53. *FEBS Lett* **420**, 25-27.

Hori, T., Osaka, F., Chiba, T., Miyamoto, C., Okabayashi, K., Shimbara, N., Kato, S. and Tanaka, K. (1999). Covalent modification of all members of human cullin family proteins by NEDD8. *Oncogene* **18**, 6829-6834.

Horwich, A. L., Weber-Ban, E. U. and Finley, D. (1999). Chaperone rings in protein folding and degradation. *PNAS* **96**, 11033-11040.

Hough, R., Pratt, G. and Rechsteiner, M. (1986). Ubiquitin-lysozyme conjugates. Identification and characterization of an ATP-dependent protease from rabbit reticulocyte lysates. *J. Biol. Chem.* **261**, 2400-2408.

Hu, S. and Yang, X. (2000). dFADD, a novel death domain-containing adapter protein for the *Drosophila* caspase DREDD. *J. Biol. Chem.* **275**, 30761-30764.

Hubbard, S. C., Walls, L., Ruley, H. E. and Muchmore, E. A. (1994). Generation of Chinese hamster ovary cell glycosylation mutants by retroviral insertional mutagenesis. Integration into a discrete locus generates mutants expressing high levels of N-glycolylneuraminic acid. *J. Biol Chem.* **269**, 3717-3724.

Huibregtse, J. M., Scheffner, M. and Howley, P. M. (1993). Cloning and expression of the cDNA for E6-AP, a protein that mediates the interaction of the human papillomavirus E6 oncoprotein with p53. *Mol. Cell. Biol.* **13**, 775-784.

Huibregtse, J. M., Scheffner, M., Beaudenon, S. and Howley, P. M. (1995). A Family of Proteins Structurally and Functionally Related to the E6-AP Ubiquitin-Protein Ligase. *PNAS* **92**, 2563-2567.

Huibregtse, J. M., Yang, J. C. and Beaudenon, S. L. (1997). The large subunit of RNA polymerase II is a substrate of the Rsp5 ubiquitin-proteináligase. *PNAS* **94**, 3656-3661.

Hulo, N., Sigrist, C. J. A., Le Saux, V., Langendijk-Genevaux, P. S., Bordoli, L., Gattiker, A., De Castro, E., Bucher, P. and Bairoch, A. (2004). Recent improvements to the PROSITE database. *Nucleic Acids Res.* **32**, D134-D137.

Igaki, T., Kanuka, H., Inohara, N., Sawamoto, K., Nunez, G., Okano, H. and Miura, M. (2000). Drob-1, a *Drosophila* member of the Bcl-2/CED-9 family that promotes cell death. *Proc. Natl. Acad. Sci. U. S. A.* **97**, 662-667.

Inohara, N., Koseki, T., Hu, Y., Chen, S. and Nunez, G. (1997). CLARP, a death effector domain-containing protein interacts with caspase-8 and regulates apoptosis. *Proc. Natl. Acad. Sci. U. S. A.* **94**, 10717-10722.

Ish-Horowicz, D. and Burke, J. F. (1981). Rapid and efficient cosmid cloning. *Nucleic Acids Res.* **9**, 2989-2998.

Jackson, P. K., Eldridge, A. G., Freed, E., Furstenthal, L., Hsu, J. Y., Kaiser, B. K. and Reimann, J. D. (2000). The lore of the RINGs: substrate recognition and catalysis by ubiquitin ligases. *Trends Cell Biol* **10**, 429-439.

Jentsch, S. (1992). The ubiquitin-conjugation system. *Annu Rev Genet* **26**, 179-207.

Jentsch, S. and Pyrowolakis, G. (2000). Ubiquitin and its kin: how close are the family ties? *Trends Cell Biol* **10**, 335-342.

Jiang, C., Lamblin, A. F., Steller, H. and Thummel, C. S. (2000). A steroid-triggered transcriptional hierarchy controls salivary gland cell death during *Drosophila* metamorphosis. *Mol. Cell* **5**, 445-455.

Jin, S. and Levine, A. J. (2001). The p53 functional circuit. *J Cell Sci* **114**, 4139-4140.

Jin, S., Martinek, S., Joo, W. S., Wortman, J. R., Mirkovic, N., Sali, A., Yandell, M. D., Pavletich, N. P., Young, M. W. and Levine, A. J. (2000). Identification and characterization of a p53 homologue in *Drosophila melanogaster*. *Proc. Natl. Acad. Sci. U. S. A.* **97**, 7301-7306.

Joazeiro, C. A. and Weissman, A. M. (2000). RING finger proteins: mediators of ubiquitin ligase activity. *Cell* **102**, 549-552.

Joazeiro, C. A., Wing, S. S., Huang, H. k., Leverson, J. D., Hunter, T. and Liu, Y. C. (1999). The Tyrosine Kinase Negative Regulator c-Cbl as a RING-Type, E2-Dependent Ubiquitin-Protein Ligase. *Science* **286**, 309-312.

Johnson, E. S. and Gupta, A. A. (2001). An E3-like factor that promotes SUMO conjugation to the yeast septins. *Cell* **106**, 735-744.

Johnson, E. S., Schwienhorst, I., Dohmen, R. J. and Blobel, G. (1997). The ubiquitin-like protein Smt3p is activated for conjugation to other proteins by an Aos1p/Uba2p heterodimer. *EMBO J.* **16**, 5509-5519.

Johnson, P. R. and Hochstrasser, M. (1997). SUMO-1: ubiquitin gains weight. *Trends Cell Biol* **7**, 408-413.

Jones, D. T. (1999). Protein secondary structure prediction based on position-specific scoring matrices. *J Mol Biol* **292**, 195-202.

Jones, S. N., Roe, A. E., Donehower, L. A. and Bradley, A. (1995). Rescue of embryonic lethality in Mdm2-deficient mice by absence of p53. *Nature* **378**, 206-208.

Kamitani, T., Kito, K., Nguyen, H. P., Wada, H., Fukuda-Kamitani, T. and Yeh, E. T. (1998). Identification of Three Major Sentrinization Sites in PML. *J. Biol. Chem.* **273**, 26675-26682.

Kanuka, H., Sawamoto, K., Inohara, N., Matsuno, K., Okano, H. and Miura, M. (1999). Control of the cell death pathway by Dapaf-1, a *Drosophila* Apaf-1/CED-4-related caspase activator. *Mol. Cell* **4**, 757-769.

Kawakami, T., Chiba, T., Suzuki, T., Iwai, K., Yamanaka, K., Minato, N., Suzuki, H., Shimbara, N., Hidaka, Y., Osaka, F. et al. (2001). NEDD8 recruits E2-ubiquitin to SCF E3 ligase. *EMBO J.* **20**, 4003-4012.

Keller, E. B. and Noon, W. A. (1985). Intron splicing: a conserved internal signal in introns of Drosophila pre-mRNAs. *Nucleic Acids Res.* **13**, 4971-4981.

Kentsis, A. and Borden, K. L. (2000). Construction of macromolecular assemblages in eukaryotic processes and their role in human disease: linking RINGs together. *Curr Protein Pept Sci* **1**, 49-73.

Khan, I. M., Fisher, R. A., Johnson, K. J., Bailey, M. E., Siciliano, M. J., Kessling, A. M., Farrer, M., Carritt, B., Kamalati, T. and Buluwela, L. (1994). The SON gene encodes a conserved DNA binding protein mapping to human chromosome 21. *Ann Hum Genet* **58 (Pt 1)**, 25-34.

Kidwell, J. F. (1972). The effective lethal phase of the curly mutant in *Drosophila melanogaster*. *J. Hered.* **63**, 100.

King, R. W., Peters, J. M., Tugendreich, S., Rolfe, M., Hieter, P. and Kirschner, M. W. (1995). A 20S complex containing CDC27 and CDC16 catalyzes the mitosis-specific conjugation of ubiquitin to cyclin B. *Cell* **81**, 279-288.

Kleijnen, M. F., Shih, A. H., Zhou, P., Kumar, S., Soccio, R. E., Kedersha, N. L., Gill, G. and Howley, P. M. (2000). The hPLIC proteins may provide a link between the ubiquitination machinery and the proteasome. *Mol Cell* **6**, 409-419.

Klionsky, D. J. and Ohsumi, Y. (1999). Vacuolar import of proteins and organelles from the cytoplasm. *Annu Rev Cell Dev Biol* **15**, 1-32.

Kloetzel, P. M. (2001). Antigen processing by the proteasome. *Nat Rev Mol Cell Biol* **2**, 179-187.

Koegl, M., Hoppe, T., Schlenker, S., Ulrich, H. D., Mayer, T. U. and Jentsch, S. (1999). A novel ubiquitination factor, E4, is involved in multiubiquitin chain assembly. *Cell* **96**, 635-644.

Kohler, A., Cascio, P., Leggett, D. S., Woo, K. M., Goldberg, A. L. and Finley, D. (2001). The axial channel of the proteasome core particle is gated by the Rpt2 ATPase and controls both substrate entry and product release. *Mol Cell* **7**, 1143-1152.

Kumar, S. (1999). Mechanisms mediating caspase activation in cell death. *Cell Death Differ* **6**, 1060-1066.

Kumar, S. and Colussi, P. A. (1999). Prodomains--adaptors--oligomerization: the pursuit of caspase activation in apoptosis. *Trends Biochem Sci* **24**, 1-4.

Kumar, S. and Doumanis, J. (2000). The fly caspases. *Cell Death. Differ.* **7**, 1039-1044.

Kurada, P. and White, K. (1998). Ras promotes cell survival in *Drosophila* by downregulating hid expression. *Cell* **95**, 319-329.

Laity, J. H., Lee, B. M. and Wright, P. E. (2001). Zinc finger proteins: new insights into structural and functional diversity. *Curr Opin Struct Biol* **11**, 39-46.

Lam, Y. A., DeMartino, G. N., Pickart, C. M. and Cohen, R. E. (1997a). Specificity of the Ubiquitin Isopeptidase in the PA700 Regulatory Complex of 26áS Proteasomes. *J. Biol. Chem.* **272**, 28438-28446.

Lam, Y. A., Xu, W., DeMartino, G. N. and Cohen, R. E. (1997b). Editing of ubiquitin conjugates by an isopeptidase in the 26S proteasome. *Nature* **385**, 737-740.

Lambertson, D., Chen, L. and Madura, K. (1999). Pleiotropic Defects Caused by Loss of the Proteasome-Interacting Factors Rad23 and Rpn10 of *Saccharomyces cerevisiae*. *Genetics* **153**, 69-79.

Lee, J. H., Lee, E., Park, J., Kim, E., Kim, J. and Chung, J. (2003). In vivo p53 function is indispensable for DNA damage-induced apoptotic signaling in Drosophila. *FEBS Lett* **550**, 5-10.

Leverson, J. D., Joazeiro, C. A., Page, A. M., Huang, H., Hieter, P. and Hunter, T. (2000). The APC11 RING-H2 finger mediates E2-dependent ubiquitination. *Mol Biol Cell* **11**, 2315-2325.

Levine, A. J. (1997). p53, the cellular gatekeeper for growth and division. *Cell* **88**, 323-331.

Li, P., Nijhawan, D., Budihardjo, I., Srinivasula, S. M., Ahmad, M., Alnemri, E. S. and Wang, X. (1997). Cytochrome c and dATP-dependent formation of Apaf-1/caspase-9 complex initiates an apoptotic protease cascade. *Cell* **91**, 479-489.

Liakopoulos, D., Doenges, G., Matuschewski, K. and Jentsch, S. (1998). A novel protein modification pathway related to the ubiquitin system. *EMBO J.* **17**, 2208-2214.

Liu, W. and Zhang, R. (1998). Upregulation of p21WAF1/CIP1 in human breast cancer cell lines MCF-7 and MDA-MB-468 undergoing apoptosis induced by natural product anticancer drugs 10-hydroxycamptothecin and camptothecin through p53-dependent and independent pathways. *Int J Oncol* **12**, 793-804.

Lohrum, M. A., Ashcroft, M., Kubbutat, M. H. and Vousden, K. H. (2000). Identification of a cryptic nucleolar-localization signal in MDM2. *Nat Cell Biol* **2**, 179-181.

Lorick, K. L., Jensen, J. P., Fang, S., Ong, A. M., Hatakeyama, S. and Weissman, A. M. (1999). RING fingers mediate ubiquitin-conjugating enzyme (E2)-dependent ubiquitination. *Proc. Natl. Acad. Sci. U. S. A.* **96**, 11364-11369.

Lowe, J., Stock, D., Jap, B., Zwickl, P., Baumeister, W. and Huber, R. (1995). Crystal structure of the 20S proteasome from the archaeon *T. acidophilum* at 3.4 A resolution. *Science* **268**, 533-539.

Luders, J., Demand, J. and Hohfeld, J. (2000). The ubiquitin-related BAG-1 provides a link between the molecular chaperones Hsc70/Hsp70 and the proteasome. *J Biol Chem.* **275**, 4613-4617.

Luders, J., Pyrowolakis, G. and Jentsch, S. (2003). The ubiquitin-like protein HUB1 forms SDS-resistant complexes with cellular proteins in the absence of ATP. *EMBO Rep.* **4**, 1169-1174.

Luo, L., Lee, T., Tsai, L., Tang, G., Jan, L. Y. and Jan, Y. N. (1997). Genghis Khan (Gek) as a putative effector for *Drosophila* Cdc42 and regulator of actin polymerization. *Proc. Natl. Acad. Sci U. S. A* **94**, 12963-12968.

Mahajan, R., Delphin, C., Guan, T., Gerace, L. and Melchior, F. (1997). A small ubiquitin-related polypeptide involved in targeting RanGAP1 to nuclear pore complex protein RanBP2. *Cell* **88**, 97-107.

Maniatis, T., Fritsch, E. and Sambrook, J., *Molecular cloning: a lab manual*. Cold Spring Harbour Laboratory Press, New York, 1989.

Maniatis, T., Goodbourn, S. and Fischer, J. A. (1987). Regulation of inducible and tissue-specific gene expression. *Science* **236**, 1237-1245.

Marin, I. and Ferrus, A. (2002). Comparative Genomics of the RBR Family, Including the Parkinson's Disease-Related Gene Parkin and the Genes of the Ariadne Subfamily. *Mol. Biol. Evol.* **19**, 2039-2050.

Martinez-Noel, G., Niedenthal, R., Tamura, T. and Harbers, K. (1999). A family of structurally related RING finger proteins interacts specifically with the ubiquitin-conjugating enzyme UbcM4. *FEBS Lett* **454**, 257-261.

Mather, A., Rakgotho, M. and Ntwasa, M. (2005). SNAMA, a novel protein with a DWNN domain and a RING finger-like motif: A possible role in apoptosis. *Biochim Biophys Acta* **1727**, 169-176.

Mattioni, T., Hume, C. R., Konigorski, S., Hayes, P., Osterweil, Z. and Lee, J. S. (1992). A cDNA clone for a novel nuclear protein with DNA binding activity. *Chromosoma* **101**, 618-624.

Mayo, L. D., Turchi, J. J. and Berberich, S. J. (1997). Mdm-2 phosphorylation by DNA-dependent protein kinase prevents interaction with p53. *Cancer Res.* **57**, 5013-5016.

Mbita, Z. *Characterization of a novel cell death related gene, DWNN, in acute renal transplant rejection and parenchymal renal diseases*, Msc Thesis, University of the Witwatesrand, Johannesburg, South Africa, 2004.

McGuffin, L. J., Bryson, K. and Jones, D. T. (2000). The PSIPRED protein structure prediction server. *Bioinformatics* **16**, 404-405.

McNally, T., Huang, Q., Janis, R. S., Liu, Z., Olejniczak, E. T. and Reilly, R. M. (2003). Structural analysis of UBL5, a novel ubiquitin-like modifier. *Protein Sci.* **12**, 1562-1566.

Meier, P. and Evan, G. (1998). Dying like flies. *Cell* **95**, 295-298.

Meier, P., Finch, A. and Evan, G. (2000). Apoptosis in development. *Nature* **407**, 796-801.

Melchior, F. (2000). SUMO--nonclassical ubiquitin. *Annu Rev Cell Dev Biol* **16**, 591-626.

Merry, D. E. and Korsmeyer, S. J. (1997). Bcl-2 gene family in the nervous system. *Annu Rev Neurosci* **20**, 245-267.

Miller, D. T. and Cagan, R. L. (1998). Local induction of patterning and programmed cell death in the developing *Drosophila* retina. *Development* **125**, 2327-2335.

Minakhina, S., Yang, J. and Steward, R. (2003). Tamo selectively modulates nuclear import in *Drosophila*. *Genes Cells* **8**, 299-310.

Misteli, T., Caceres, J. F. and Spector, D. L. (1997). The dynamics of a pre-mRNA splicing factor in living cells. *Nature* **387**, 523-527.

Mizuguchi, K., DEANE, C. M., Blundell, T. L. and OVERINGTON, J. P. (1998). HOMSTRAD: A database of protein structure alignments for homologous families. *Protein Sci.* **7**, 2469-2471.

Mizuno, T., Amano, M., Kaibuchi, K. and Nishida, Y. (1999). Identification and characterization of *Drosophila* homolog of Rho-kinase. *Gene* **238**, 437-444.

Momand, J., Wu, H. H. and Dasgupta, G. (2000). MDM2-master regulator of the p53 tumor suppressor protein. *Gene* **242**, 15-29.

Montes, d. O. L., Wagner, D. S. and Lozano, G. (1995). Rescue of early embryonic lethality in mdm2-deficient mice by deletion of p53. *Nature* **378**, 203-206.

Mori, T., Li, Y., Hata, H., Ono, K. and Kochi, H. (2002). NIRF, a novel RING finger protein, is involved in cell-cycle regulation. *Biochem. Biophys. Res. Commun.* **296**, 530-536.

Mount, S. M. (1982). A catalogue of splice junction sequences. *Nucleic Acids Res* **10**, 459-472.

Mount, S. M., Burks, C., Hertz, G., Stormo, G. D., White, O. and Fields, C. (1992). Splicing signals in *Drosophila*: intron size, information content, and consensus sequences. *Nucleic Acids Res* **20**, 4255-4262.

Mueller, T. D. and Feigon, J. (2003). Structural determinants for the binding of ubiquitin-like domains to the proteasome. *EMBO J.* **22**, 4634-4645.

Muller, S. and Dejean, A. (1999). Viral Immediate-Early Proteins Abrogate the Modification by SUMO-1 of PML and Sp100 Proteins, Correlating with Nuclear Body Disruption. *J. Virol.* **73**, 5137-5143.

Muralidhar, M. G. and Thomas, J. B. (1993). The *Drosophila* bendless gene encodes a neural protein related to ubiquitin-conjugating enzymes. *Neuron* **11**, 253-266.

Nakielny, S. and Dreyfuss, G. (1999). Transport of proteins and RNAs in and out of the nucleus. *Cell* **99**, 677-690.

Nicholson, D. W. (1999). Caspase structure, proteolytic substrates, and function during apoptotic cell death. *Cell Death Differ* **6**, 1028-1042.

Nordstrom, W. and Abrams, J. M. (2000). Guardian ancestry: fly p53 and damage-inducible apoptosis. *Cell Death. Differ.* **7**, 1035-1038.

Ntwasa, M., Buchanan, S. G. and Gay, N. J. (1994). *Drosophila* ribosomal protein L18a: cDNA sequence, expression and chromosomal localization of the gene. *Biochim. Biophys. Acta* **1218**, 210-212.

Ohno, M. and Shimura, Y. (1996). A human RNA helicase-like protein, HRH1, facilitates nuclear export of spliced mRNA by releasing the RNA from the spliceosome. *Genes & Dev.* **10**, 997-1007.

Ollmann, M., Young, L. M., Di Como, C. J., Karim, F., Belvin, M., Robertson, S., Whittaker, K., Demsky, M., Fisher, W. W., Buchman, A. et al. (2000). *Drosophila* p53 is a structural and functional homolog of the tumor suppressor p53. *Cell* **101**, 91-101.

Ortolan, T. G., Tongaonkar, P., Lambertson, D., Chen, L., Schauber, C. and Madura, K. (2000). The DNA repair protein rad23 is a negative regulator of multi-ubiquitin chain assembly. *Nat Cell Biol* **2**, 601-608.

Owens, G. P., Hahn, W. E. and Cohen, J. J. (1991). Identification of mRNAs associated with programmed cell death in immature thymocytes. *Mol Cell Biol* **11**, 4177-4188.

Ozkaynak, E., Finley, D. and Varshavsky, A. (1984). The yeast ubiquitin gene: head-to-tail repeats encoding a polyubiquitin precursor protein. *Nature* **312**, 663-666.

Pandey, R., Muller, A., Napoli, C. A., Selinger, D. A., Pikaard, C. S., Richards, E. J., Bender, J., Mount, D. W. and Jorgensen, R. A. (2002). Analysis of histone acetyltransferase and histone deacetylase families of *Arabidopsis thaliana* suggests functional diversification of chromatin modification among multicellular eukaryotes. *Nucleic Acids Res* **30**, 5036-5055.

Pardue, M. L. *In situ hybridization of DNA to chromosomes and nuclei in Drosophila: a practical approach*. pp. 111-137, 1986.

Pearson, W. R., Wood, T., Zhang, Z. and Miller, W. (1997). Comparison of DNA sequences with protein sequences. *Genomics* **46**, 24-36.

Pickart, C. M., Kasperek, E. M., Beal, R. and Kim, A. (1994). Substrate properties of site-specific mutant ubiquitin protein (G76A) reveal unexpected mechanistic features of ubiquitin-activating enzyme (E1). *J. Biol. Chem.* **269**, 7115-7123.

Pretorius, A. *The Isolation of Novel components of the Antigen Presentation and Processing Pathway via MHC class I molecules using a Retroviral Promoter Trap in Chinese Hamster Ovary (CHO) Cells*, Msc Thesis, University of the Western Cape, South Africa, 1999.

Pringa, E., Martinez-Noel, G., Muller, U. and Harbers, K. (2001). Interaction of the ring finger-related U-box motif of a nuclear dot protein with ubiquitin-conjugating enzymes. *J Biol Chem.* **276**, 19617-19623.

Proudfoot, N. J. (1989). How RNA polymerase II terminates transcription in higher eukaryotes. *Trends Biochem Sci* **14**, 105-110.

Quinn, L., Coombe, M., Mills, K., Daish, T., Colussi, P., Kumar, S. and Richardson, H. (2003). Buffy, a Drosophila Bcl-2 protein, has anti-apoptotic and cell cycle inhibitory functions. *EMBO J.* **22**, 3568-3579.

Raff, M. (1998). Cell suicide for beginners. *Nature* **396**, 119-122.

Ramelot, T. A., Cort, J. R., Yee, A. A., Semesi, A., Edwards, A. M., Arrowsmith, C. H. and Kennedy, M. A. (2003). Solution structure of the yeast ubiquitin-like modifier protein Hub1. *J Struct Funct Genomics* **4**, 25-30.

Rechsteiner, M. and Rogers, S. W. (1996). PEST sequences and regulation by proteolysis. *Trends Biochem Sci* **21**, 267-271.

Rechsteiner, M., Realini, C. and Ustrell, V. (2000). The proteasome activator 11 S REG (PA28) and class I antigen presentation. *Biochem J* **345 Pt 1:1-15.**, 1-15.

Reese, M. G. (2001). Application of a time-delay neural network to promoter annotation in the *Drosophila melanogaster* genome. *Comput Chem* **26**, 51-56.

Regnier, C. H., Tomasetto, C., Moog-Lutz, C., Chenard, M. P., Wendling, C., Basset, P. and Rio, M. C. (1995). Presence of a new conserved domain in CART1, a novel member of the tumor necrosis factor receptor-associated protein family, which is expressed in breast carcinoma. *J Biol Chem* **270**, 25715-25721.

Reiss, Y., Heller, H. and Hershko, A. (1989). Binding sites of ubiquitin-protein ligase. Binding of ubiquitin-protein conjugates and of ubiquitin-carrier protein. *J. Biol. Chem.* **264**, 10378-10383.

Richardson, H. and Kumar, S. (2002). Death to flies: *Drosophila* as a model system to study programmed cell death. *J. Immunol. Methods* **265**, 21-38.

Roberts, D. L., Merrison, W., MacFarlane, M. and Cohen, G. M. (2001). The Inhibitor of Apoptosis Protein-binding Domain of Smac Is Not Essential for its Proapoptotic Activity. *J. Cell Biol.* **153**, 221-228.

Rodriguez, A., Oliver, H., Zou, H., Chen, P., Wang, X. and Abrams, J. M. (1999). Dark is a Drosophila homologue of Apaf-1/CED-4 and functions in an evolutionarily conserved death pathway. *Nat. Cell Biol.* **1**, 272-279.

Rogers, S., Wells, R. and Rechsteiner, M. (1986). Amino acid sequences common to rapidly degraded proteins: the PEST hypothesis. *Science* **234**, 364-368.

Roth, J., Dobbelstein, M., Freedman, D. A., Shenk, T. and Levine, A. J. (1998). Nucleo-cytoplasmic shuttling of the hdm2 oncoprotein regulates the levels of the p53 protein via a pathway used by the human immunodeficiency virus rev protein. *EMBO J.* **17**, 554-564.

Rubin, G. M., Hong, L., Brokstein, P., Evans-Holm, M., Frise, E., Stapleton, M. and Harvey, D. A. (2000). A *Drosophila* complementary DNA resource. *Science* **287**, 2222-2224.

Russnak, R. and Ganem, D. (1990). Sequences 5' to the polyadenylation signal mediate differential poly(A) site use in hepatitis B viruses. *Genes & Dev.* **4**, 764-776.

Saitou, N. and Nei, M. (1987). The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* **4**, 406-425.

Sakai, Y., Saijo, M., Coelho, K., Kishino, T., Niikawa, N. and Taya, Y. (1995). cDNA sequence and chromosomal localization of a novel human protein, RBQ-1 (RBBP6), that binds to the retinoblastoma gene product. *Genomics* **30**, 98-101.

Saurin, A. J., Borden, K. L., Boddy, M. N. and Freemont, P. S. (1996). Does this have a familiar RING? *Trends Biochem. Sci.* **21**, 208-214.

Sawamoto, K., Taguchi, A., Hirota, Y., Yamada, C., Jin, M. H. and Okano, H. (1998). Argos induces programmed cell death in the developing *Drosophila* eye by inhibition of the Ras pathway. *Cell Death. Differ.* **5**, 262-270.

Scheffner, M., Huibregtse, J. M., Vierstra, R. D. and Howley, P. M. (1993). The HPV-16 E6 and E6-AP complex functions as a ubiquitin-protein ligase in the ubiquitination of p53. *Cell* **75**, 495-505.

Scheffner, M., Nuber, U. and Huibregtse, J. M. (1995). Protein ubiquitination involving an E1-E2-E3 enzyme ubiquitin thioester cascade. *Nature* **373**, 81-83.

Schultz, J., Milpetz, F., Bork, P. and Ponting, C. P. (1998). SMART, a simple modular architecture research tool: Identification of signaling domains. *PNAS* **95**, 5857-5864.

Scott, R. E., Giannakouros, T., Gao, S. and Peidis, P. (2003). Functional potential of P2P-R: a role in the cell cycle and cell differentiation related to its interactions with proteins that bind to matrix associated regions of DNA? *J Cell Biochem* **90**, 6-12.

Secombe, J. and Parkhurst, S. M. (2004). Drosophila Topors Is a RING Finger-containing Protein That Functions as a Ubiquitin-protein Isopeptide Ligase for the Hairy Basic Helix-Loop-Helix Repressor Protein. *J. Biol. Chem.* **279**, 17126-17133.

Seol, J. H., Feldman, R. M. R., Zachariae, W., Shevchenko, A., Correll, C. C., Lyapina, S., Chi, Y., Galova, M., Claypool, J., Sandmeyer, S. et al. (1999). Cdc53/cullin and the essential Hrt1 RING-H2 subunit of SCF define a ubiquitin ligase module that activates the E2 enzyme Cdc34. *Genes & Dev.* **13**, 1614-1626.

Shi, J., Blundell, T. L. and Mizuguchi, K. (2001). FUGUE: sequence-structure homology recognition using environment-specific substitution tables and structure-dependent gap penalties. *J Mol Biol* **310**, 243-257.

Shieh, S. Y., Ikeda, M., Taya, Y. and Prives, C. (1997). DNA damage-induced phosphorylation of p53 alleviates inhibition by MDM2. *Cell* **91**, 325-334.

Simons, A., Melamed-Bessudo, C., Wolkowicz, R., Sperling, J., Sperling, R., Eisenbach, L. and Rotter, V. (1997). PACT: cloning and characterization of a cellular p53 binding protein that interacts with Rb. *Oncogene* **14**, 145-155.

Sloper-Mould, K. E., Jemc, J. C., Pickart, C. M. and Hicke, L. (2001). Distinct Functional Surface Regions on Ubiquitin. *J. Biol. Chem.* **276**, 30483-30489.

Soetens, O., De Craene, J. O. and Andre, B. (2001). Ubiquitin Is Required for Sorting to the Vacuole of the Yeast General Amino Acid Permease, Gap1. *J. Biol. Chem.* **276**, 43949-43957.

Sogame, N., Kim, M. and Abrams, J. M. (2003). *Drosophila* p53 preserves genomic stability by regulating cell death. *PNAS* **100**, 4696-4701.

Spence, J., Gali, R. R., Dittmar, G., Sherman, F., Karin, M. and Finley, D. (2000). Cell cycle-regulated modification of the ribosome by a variant multiubiquitin chain. *Cell* **102**, 67-76.

Spence, J., Sadis, S., Haas, A. L. and Finley, D. (1995). A ubiquitin mutant with specific defects in DNA repair and multiubiquitination. *Mol. Cell. Biol.* **15**, 1265-1273.

Spradling, A. C., Stern, D., Beaton, A., Rhem, E. J., Laverty, T., Mozden, N., Misra, S. and Rubin, G. M. (1999). The Berkeley Drosophila Genome Project gene disruption project: Single P-element insertions mutating 25% of vital *Drosophila* genes. *Genetics* **153**, 135-177.

Stapleton, M., Carlson, J., Brokstein, P., Yu, C., Champe, M., George, R., Guarin, H., Kronmiller, B., Pacleb, J., Park, S. et al. (2002a). A *Drosophila* full-length cDNA resource. *Genome Biol* **3**, RESEARCH0080.

Stapleton, M., Liao, G., Brokstein, P., Hong, L., Carninci, P., Shiraki, T., Hayashizaki, Y., Champe, M., Pacleb, J., Wan, K. et al. (2002b). The *Drosophila* Gene Collection: Identification of Putative Full-Length cDNAs for 70% of *D. melanogaster* Genes. *Genome Res.* **12**, 1294-1300.

Staub, O., Dho, S., Henry, P., Correa, J., Ishikawa, T., McGlade, J. and Rotin, D. (1996). WW domains of Nedd4 bind to the proline-rich PY motifs in the epithelial Na+ channel deleted in Liddle's syndrome. *EMBO J.* **15**, 2371-2380.

Steller, H. (2000). *Drosophila* p53: meeting the Grim Reaper. *Nat. Cell Biol.* **2**, E100-E102.

Stott, F. J., Bates, S., James, M. C., McConnell, B. B., Starborg, M., Brookes, S., Palmero, I., Ryan, K., Hara, E., Vousden, K. H. et al. (1998). The alternative product from the human CDKN2A locus, p14ARF, participates in a regulatory feedback loop with p53 and MDM2. *EMBO J.* **17**, 5001-5014.

Strickland, E., Hakala, K., Thomas, P. J. and DeMartino, G. N. (2000). Recognition of misfolding proteins by PA700, the regulatory subcomplex of the 26 S proteasome. *J Biol Chem.* **275**, 5565-5572.

Stronach, B. E., Renfranz, P. J., Lilly, B. and Beckerle, M. C. (1999). Muscle LIM proteins are associated with muscle sarcomeres and require dMEF2 for their expression during Drosophila myogenesis. *Mol Biol Cell* **10**, 2329-2342.

Stronach, B. E., Siegrist, S. E. and Beckerle, M. C. (1996). Two muscle-specific LIM proteins in *Drosophila*. *The Journal of Cell Biology* **134**, 1179-1195.

Stronach, B. E., Siegrist, S. E. and Beckerle, M. C. (1996). Two muscle-specific LIM proteins in *Drosophila. The Journal of Cell Biology* **134**, 1179-1195.

Sudakin, V., Ganoth, D., Dahan, A., Heller, H., Hershko, J., Luca, F. C., Ruderman, J. V. and Hershko, A. (1995). The cyclosome, a large complex containing cyclin-selective ubiquitin ligase activity, targets cyclins for destruction at the end of mitosis. *Mol. Biol. Cell* **6**, 185-197.

Swaminathan, S., Amerik, A. Y. and Hochstrasser, M. (1999). The Doa4 deubiquitinating enzyme is required for ubiquitin homeostasis in yeast. *Mol Biol Cell* **10**, 2583-2594.

Takada, S., Kelkar, A. and Theurkauf, W. E. (2003). *Drosophila* checkpoint kinase 2 couples centrosome function and spindle assembly to genomic integrity. *Cell* **113**, 87-99.

Tautz, D. and Pfeifle, C. (1989). A non-radioactive *in situ* hybridization method for the localization of specific RNAs in *Drosophila* embryos reveals translational control of the segmentation gene hunchback. *Chromosoma* **98**, 81-85.

Tautz,D. Genetic and Molecular Analysis of Early Pattern Formation in *Drosophila. Drosophila* as a Genetic System . In Development. The Molecular Genetic Approach, V.E.A.Russo, S.Brody, D.Cove, and S.Ottolenghi, eds. Springer-Verlag Berlin Heidelberg), pp. 308-310, 1992.

Thornberry, N. A. and Lazebnik, Y. (1998). Caspases: Enemies Within. *Science* **281**, 1312-1316.

Thress, K., Henzel, W., Shillinglaw, W. and Kornbluth, S. (1998). Scythe: a novel reaper-binding apoptotic regulator. *EMBO J.* **17**, 6135-6143.

Towbin, H., Staehelin, T. and Gordon, J. (1979). Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. *Proc. Natl. Acad. Sci U. S. A* **76**, 4350-4354.

Tyers, M. and Willems, A. R. (1999). One Ring to Rule a Superfamily of E3 Ubiquitin Ligases. *Science* **284**, 601.

Unger, T., Juven-Gershon, T., Moallem, E., Berger, M., Sionov, R., Lozano, G., Oren, M. and Haupt, Y. (1999). Critical role for Ser20 of human p53 in the negative regulation of p53 by Mdm2. *EMBO J.* **18**, 1805-1814.

Varshavsky, A. (1992). The N-end rule. *Cell* **69**, 725-735.

Vernooy, S. Y., Copeland, J., Ghaboosi, N., Griffin, E. E., Yoo, S. J. and Hay, B. A. (2000). Cell death regulation in *Drosophila*: conservation of mechanism and unique insights. *J. Cell Biol.* **150**, F69-F76.

Vo, L. T. A., Minet, M., Schmitter, J. M., Lacroute, F. and Wyers, F. (2001). Mpe1, a Zinc Knuckle Protein, Is an Essential Component of Yeast Cleavage and Polyadenylation Factor Required for the Cleavage and Polyadenylation of mRNA. *Mol. Cell. Biol.* **21**, 8346-8356.

Vogelstein, B. and Gillespie, D. (1979). Preparative and analytical purification of DNA from agarose. *Proc. Natl. Acad. Sci U. S. A.* **76**, 615-619.

Vucic, D., Kaiser, W. J. and Miller, L. K. (1998). Inhibitor of apoptosis proteins physically interact with and block apoptosis induced by Drosophila proteins HID and GRIM. *Mol. Cell Biol.* **18**, 3300-3309.

Vucic, D., Kaiser, W. J., Harvey, A. J. and Miller, L. K. (1997). Inhibition of reaper-induced apoptosis by interaction with inhibitor of apoptosis proteins (IAPs). *Proc. Natl. Acad. Sci. U. S. A.* **94**, 10183-10188.

Wahle, E. and Keller, W. (1992). The biochemistry of 3'-end cleavage and polyadenylation of messenger RNA precursors. *Annu Rev Biochem* **61**, 419-440.

Wang, S. L., Hawkins, C. J., Yoo, S. J., Muller, H. A. and Hay, B. A. (1999). The *Drosophila* caspase inhibitor DIAP1 is essential for cell survival and is negatively regulated by HID. *Cell* **98**, 453-463.

Watkins, J. F., Sung, P., Prakash, L. and Prakash, S. (1993). The *Saccharomyces cerevisiae* DNA repair gene RAD23 encodes a nuclear protein containing a ubiquitin-like domain required for biological function. *Mol. Cell. Biol.* **13**, 7757-7765.

Weinberg, R. A. (1992). The retinoblastoma gene and gene product. *Cancer Surv* **12**, 43-57.

Weinberg, R. A. (1995). The retinoblastoma protein and cell cycle control. *Cell* **81**, 323-330.

White, K., Grether, M. E., Abrams, J. M., Young, L., Farrell, K. and Steller, H. (1994). Genetic control of programmed cell death in *Drosophila*. *Science* **264**, 677-683.

Wilkinson, K. D. (2000). Ubiquitination and deubiquitination: targeting of proteins for degradation by the proteasome. *Semin. Cell Dev Biol* **11**, 141-148.

Wing, J. P., Schwartz, L. M. and Nambu, J. R. (2001). The RHG motifs of Drosophila Reaper and Grim are important for their distinct cell death-inducing abilities. *Mech Dev* **102**, 193-203.

Wing, J. P., Zhou, L., Schwartz, L. M. and Nambu, J. R. (1998). Distinct cell killing properties of the Drosophila reaper, head involution defective, and grim genes. *Cell Death Differ* **5**, 930-939.

Witte, M. and Scott, R. (1997). The proliferation potential protein-related (P2P-R) gene with domains encoding heterogeneous nuclear ribonucleoprotein association and Rb1 binding shows repressed expression during terminal?differentiation. *PNAS* **94**, 1212.

Yaglom, J., Linskens, M. H., Sadis, S., Rubin, D. M., Futcher, B. and Finley, D. (1995). p34Cdc28-mediated control of Cln3 cyclin degradation. *Mol. Cell. Biol.* **15**, 731-741.

Yamano, H., Gannon, J. and Hunt, T. (1996). The role of proteolysis in cell cycle progression in *Schizosaccharomyces pombe*. *EMBO J.* **15**, 5268-5279.

Yeh, E. T., Gong, L. and Kamitani, T. (2000). Ubiquitin-like proteins: new wines in new bottles. *Gene* **248**, 1-14.

Yewdell, J. W. and Bennink, J. R. (2001). Cut and trim: generating MHC class I peptide ligands. *Curr. Opin. Immunol* **13**, 13-18.

Zahler, A. M., Lane, W. S., Stolk, J. A. and Roth, M. B. (1992). SR proteins: a conserved family of pre-mRNA splicing factors. *Genes & Dev.* **6**, 837-847.

Zahler, A. M., Neugebauer, K. M., Stolk, J. A. and Roth, M. B. (1993). Human SR proteins and isolation of a cDNA encoding SRp75. *Mol. Cell. Biol.* **13**, 4023-4028.

Zhang, Y., Gao, J., Chung, K. K. K., Huang, H., Dawson, V. L. and Dawson, T. M. (2000). Parkin functions as an E2-dependent ubiquitin- protein ligase and promotes the degradation of the synaptic vesicle-associated protein, CDCrel-1. *PNAS* **97**, 13354-13359.

Zhou, L., Song, Z., Tittel, J. and Steller, H. (1999a). HAC-1, a Drosophila homolog of APAF-1 and CED-4 functions in developmental and radiation-induced apoptosis. *Mol. Cell* **4**, 745-755.

Zhou, R., Wen, H. and Ao, S. Z. (1999). Identification of a novel gene encoding a p53-associated protein. *Gene* **235**, 93-101.

Zou, H., Li, Y., Liu, X. and Wang, X. (1999b). An APAF-1+Cytochrome c Multimeric Complex Is a Functional Apoptosome That Activates Procaspase-9. *J. Biol. Chem.* **274**, 11549-11556.

Zungu, M. *Characterization of a a member of DWNN superfamily in Drosophila melanogaster*, University of the Witwatersrand, Johannesburg, South Africa, 2003.

APPENDIX