BIOCHEMICAL ANALYSIS OF RBBP6 PROTEINS AND THEIR IMPACT ON TUMOUR SUPPRESSORS

Brent Oosthuysen
0510727H

A dissertation submitted to the Faculty of Science, University of the Witwatersrand, in fulfilment of the requirements for the degree of Master of Science.

Johannesburg, 2014

DECLARATION

I declare that this dissertation is my own, unaided work. It is being submitted for the degree of Master of Science at the University of the Witwatersrand, Johannesburg. It has not been submitted before for any degree or examination at any other university.

Brent Oosthuysen

30 September 2014

RESEARCH OUTPUT

CONFERENCE POSTERS:

Brent Oosthuysen, Umar-Faruq Cajee, Lehlogonolo Mokgohloa and Monde Ntwasa. The characterization of *Drosophila melanogaster* RBBP6 protein, SNAMA. SASBMB / FASBMB Congress, 29 January – 1 February 2012. Champagne Sports Resort, Drakensberg, South Africa.

Brent Oosthuysen and Monde Ntwasa. Characterization of the *Drosophila melanogaster* RBBP6 protein, SNAMA. 4th Cross-faculty Postgraduate Symposium, 19-22 October 2012. Professional Development Hub, University of the Witwatersrand, Johannesburg, South Africa.

Brent Oosthuysen and Monde Ntwasa. Characterization of the *Drosophila melanogaster* RBBP6 protein, SNAMA. Molecular Biosciences Research Thrust Research Day, 5 December 2012. Professional Development Hub, University of the Witwatersrand, Johannesburg, South Africa.

ABSTRACT

Snama is the *Drosophila melanogaster* homologue of the human *Retinoblastoma Binding Protein 6 (RBBP6)* gene. While mammalian RBBP6 proteins have been shown to interact with tumour suppressor proteins such as RETINOBLASTOMA (Rb) and p53, this has not been determined in *Drosophila melanogaster*. Moreover, SNAMA proteins have yet to be experimentally verified. To date, studies have shown *Snama* to be involved in cellular proliferation, DNA metabolism and cell cycle control. In this study, we sought to identify and characterize SNAMA proteins through immunological and biochemical approaches. Poly-clonal chicken antibodies were raised against an amino-terminal fragment common to both putative SNAMA proteins and purified via pre-adsorption. The antibodies were utilized in immunoblots – to investigate the spatio-temporal regulation, as well as in co-immunoprecipitation assays – to discover potential interaction partners of SNAMA proteins. Putative SNAMA-PA was not detected in this study. Acetylation of protein extracts, which has been shown to overcome the inability to visualize RBBP6 homologues, did not result in the detection of SNAMA-PA. Up-regulation of *Snama-RA* through camptothecin-treatment also did not result in the detection of SNAMA-PA; neither did a selective precipitation procedure aimed at enriching extracts for RS-domain-containing proteins, such as SNAMA-PA. As putative SNAMA-PB lacks the C-terminal extension of putative SNAMA-PA, it could not be manipulated via these biochemical procedures. Regardless, a protein of 55 kDa – which is the expected mass of SNAMA-PB – was detected. Furthermore, the spatio-temporal regulation of this 55 kDa protein was consistent with that expected of putative SNAMA-PB; thus, this study has possibly verified the existence of SNAMA-PB. Furthermore, we have identified potential interaction partners – most significantly DREF – which is known to function in a similar capacity to the expected roles of SNAMA proteins, namely: cell cycle control and cellular proliferation.

ACKNOWLEDGMENTS

- Prof. Theresa Coetzer and Jacky Foster, Department of Biochemistry at the University of KwaZulu-Natal, for antibody production and advice on purification of antibodies.
- Dr. Susan Chemaly, School of Chemistry at the University of the Witwatersrand, for use of her facilities.
- Dr. Stoyan Stoychev, Biosciences Division at the Council for Scientific and Industrial Research, for mass spectrometry work.
- Dr. Rodney Hull, for his friendship and willingness to always lend an ear.
- My father, Stephen Oosthuysen, for financial assistance.
- My supervisor, Dr. Monde Ntwasa, for his patience and guidance in completion of this dissertation.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

a.a.	amino acid
ATP	adenosine triphosphate
b.p.	base pairs
C-terminal	carboxyl-terminal
COIP	co-immunoprecipitation
DCM	DWNN catalytic module
ddH_2O	distilled, deionized water
Dmp53	*Drosophila melanogaster* p53
DMSO	dimethyl sulfoxide
DNABD	DNA binding domain
DNAse	deoxyribonuclease
DTT	dithiothreitol
DWNN	Domain-With-No-Name
EDTA	ethylenediaminetetraacetic acid
GSH	glutathione
GST	glutathione S-transferase
GuHCl	guanidium hydrochloride
HDM2	HUMAN-DOUBLE-MINUTE 2
HEPES	4-(2-hydroxyethyl)piperazine-1-ethanesulfonic acid sodium salt
HGNC	Human Genome Naming Convention
HSP27	HEAT SHOCK PROTEIN 27
IMAC	immobilized metal-ion affinity chromatography
IPTG	isopropyl-β-D-thiogalactoside
kDa	kilodalton
LC MS/MS	liquid chromatography tandem mass spectrometry
mA	milliamperes
MALDTI-TOF	matrix assisted laser desorption ionization – time-of-flight
Mdm2	Mouse-double-minute 2

MS	mass spectrometry
MWCO	molecular weight cut-off
N-terminal	amino-terminal
$O.D_{600}$	optical density at a wavelength of 600 nm
P2P-R	PROLIFERATION POTENTIAL PROTEIN-RELATED
PBS	phosphate buffered saline
PDI	PROTEIN DISULFIDE ISOMERASE
PHAX	PHOSPHORYLATED ADAPTER FOR RNA EXPORT
PMSF	phenylmethylsulfonyl fluoride
PVDF	Polyvinylidene fluoride
Rb	Retinoblastoma
RBBP6	Retinoblastoma Binding Protein 6
RING	Really Interesting New Gene
RRM	RNA recognition motif
RS	Arginine/Serine dipeptide
SDS	sodium dodecyl sulphate
SDS-PAGE	sodium dodecyl sulphate – polyacrylamide gel electrophoresis
SR	Serine/Arginine dipeptide
SRPK1	Serine/Arginine-rich splicing factor protein kinase 1
SRSF	Serine/Arginine-rich splicing factor
TEMED	N,N,N’,N’-tetramethylethylenediamine
ULM	ubiquitin-like modification
UBD	ubiquitin binding domain
UPP	ubiquitin-proteasomal pathway
YB-1	Y-box binding protein 1

CHAPTER ONE

1. INTRODUCTION

A RING domain, found in the RETINOBLASTOMA BINDING PROTEIN 6 (RBBP6) family, has been shown to impart these proteins with ubiquitin ligase activity (Antunes, 2008; Chibi *et al.*, 2008). RBBP6 proteins also contain an amino-terminal domain called Domain-With-No-Name (DWNN), which is found exclusively in RBBP6 proteins and is highly similar to ubiquitin in terms of its tertiary structure (Mather, 2005; Mather *et al.*, 2005; Pugh *et al.*, 2006). Furthermore, the DWNN domain is also expressed as a stand-alone domain-protein in mammals and may be representative of a putative form of ubiquitin-like modification (Pugh *et al.*, 2006). These facts, taken together, suggest that the RBBP6 family of proteins are involved in protein modification by ubiquitination or ubiquitin-like mechanisms.

RBBP6 proteins have also been implicated in other cellular functions, such as mRNA processing (Vo *et al.*, 2001) and cell cycle control (Mather, 2005; Mather *et al.*, 2005; Jones *et al.*, 2006). Furthermore, *Mus musculus* RBBP6 has been demonstrated to facilitate MDM2-mediated ubiquitination of p53 (Li *et al.*, 2007), thus implicating RBBP6 with cell cycle control and apoptosis. Despite involvement in all these diverse, yet crucial, cellular processes, little research has been carried out on this important family of proteins. For instance, only two crystal structures pertaining to the RBBP6 family of proteins exist: those of the DWNN domain (Pugh *et al.*, 2006) and RING domain (Kappo *et al.*, 2012) of *Homo sapiens* RBBP6. Furthermore, while two *Drosophila melanogaster* RBBP6 transcripts are known to exist (Mather, 2005; Mather *et al.*, 2005; Jones *et al.*, 2006; Daines *et al.*, 2011), the resultant proteins have not been experimentally verified.

This chapter begins by providing a very brief overview of the ubiquitin system. Thereafter, RBBP6 proteins in human and prominent model organisms, as well as

structural features of these proteins, are discussed. Particular emphasis is given to *D. melanogaster*, which forms the basis of this study. Lastly, this dissertation presents a contemporary account of known binding partners of RBBP6 proteins, as well as their characterized roles in these interactions.

1.1 THE UBIQUITIN SYSTEM

Ubiquitin is a highly evolutionarily conserved, 76 amino-acid (a.a.) protein, which is covalently attached onto substrates as a form of post-translational modification. The process of ubiquitin conjugation is well characterized and has been extensively reviewed by Hershko & Ciechanover (1998) – two of three joint-winners of the 2004 Nobel Prize in Chemistry, for their work on ubiquitin. Briefly, the attachment of ubiquitin to a substrate (Figure 1.1) involves three key enzymes and is initiated by the conjugation of an ubiquitin molecule, at its Gly-76 residue, to a Cys residue on an E1 ubiquitin-activating enzyme, with the input of ATP. This results in the formation of an activated ubiquitin-E1 complex. The activated ubiquitin is then transferred to an E2 ubiquitin-conjugating enzyme, also through thioester bond formation between a Cys on the E2 enzyme and Gly-76 of ubiquitin. Lastly, an E3 ubiquitin protein ligase – having bound to the substrate – then interacts with the E2-ubiquitin complex and catalyzes an amide bond between Gly-76 of ubiquitin and an ε-amino group of a specified Lys residue on the substrate. As the E3 ubiquitin ligase binds to the target of interest, it confers the greatest level of specificity of the reaction (Passmore & Barford, 2004). Indeed, this also manifests in the number of known E1, E2 and E3 enzymes: E1s are the least numerous, while E3s are the most abundant as well as substrate-specific enzymes involved in ubiquitin conjugation.

A substrate may be modified with a single ubiquitin protein or even a chain of multiple, linked ubiquitin molecules. In the case of the latter, a fourth enzyme – the E4 ubiquitin-chain elongation factor – is required for poly-ubiquitin chain formation (Koegl *et al.*, 1999; Kuhlbrodt *et al.*, 2005). It is still unknown whether E4s are absolutely essential or whether substrates can also be poly-ubiquitinated

through multiple rounds of the E1-E2-E3 cascade (Jung *et al.*, 2009). The number of ubiquitin molecules attached, as well as the lysine residue of ubiquitin proteins involved in a poly-ubiquitin chain, determines the substrate's fate. Mono-ubiquitination of a substrate can impart a wide range of effects, such as altering the protein's cellular localization, or affecting processes such as transcription and signal transduction (Hochstrasser, 2009). Poly-ubiquitination of a substrate involving Lys-76 to Lys-63 linked ubiquitin molecules has also been shown to impart a wide range of effects. However, the majority of poly-ubiquitin chains are formed between an ubiquitin protein conjugated via its Lys-76 onto Lys-48 of the previous ubiquitin molecule (Hershko & Ciechanover, 1998). A chain of four such linked ubiquitin molecules is known to direct a substrate to the 26S proteasome, for destruction.

The 26S proteasome consists of a 20S inner "core" and two outer 19S "caps". The 20S core is made up of four rings, each in turn made up of seven different subunits. The two inner rings, known as the beta rings, are responsible for degradation of a substrate. The two outer rings, known as the alpha rings, bind the two 19S caps. A variety of caps exist (reviewed in Jung *et al.*, 2009), which can impart different functionalities for the proteasome. In the canonical 26S proteasome, the two 19S caps are responsible for regulating entry into the core, after first removing the poly-ubiquitin chain from a substrate and recycling the chain to monomers, for re-use. The caps are, therefore, also responsible for binding the poly-ubiquitin chain, consisting of four Lys-76 to Lys-48 linkages, attached to the substrate. Specificity of binding occurs in the form of ubiquitin-binding domains (UBDs), which are present not only on the 19S caps but on all ubiquitin receptors. The UBDs are specific for a particular linkage involved in a poly-ubiquitin chain, as well as the length of the chain. This specificity occurs, in part, due to the different linkages having different spatial orientations (see Figure 1.1); for e.g. Lys-76 to Lys-63 linkages form linear chains, whereas Lys-76 to Lys-48 linkages form kinked chains (Broemer & Meier, 2009).

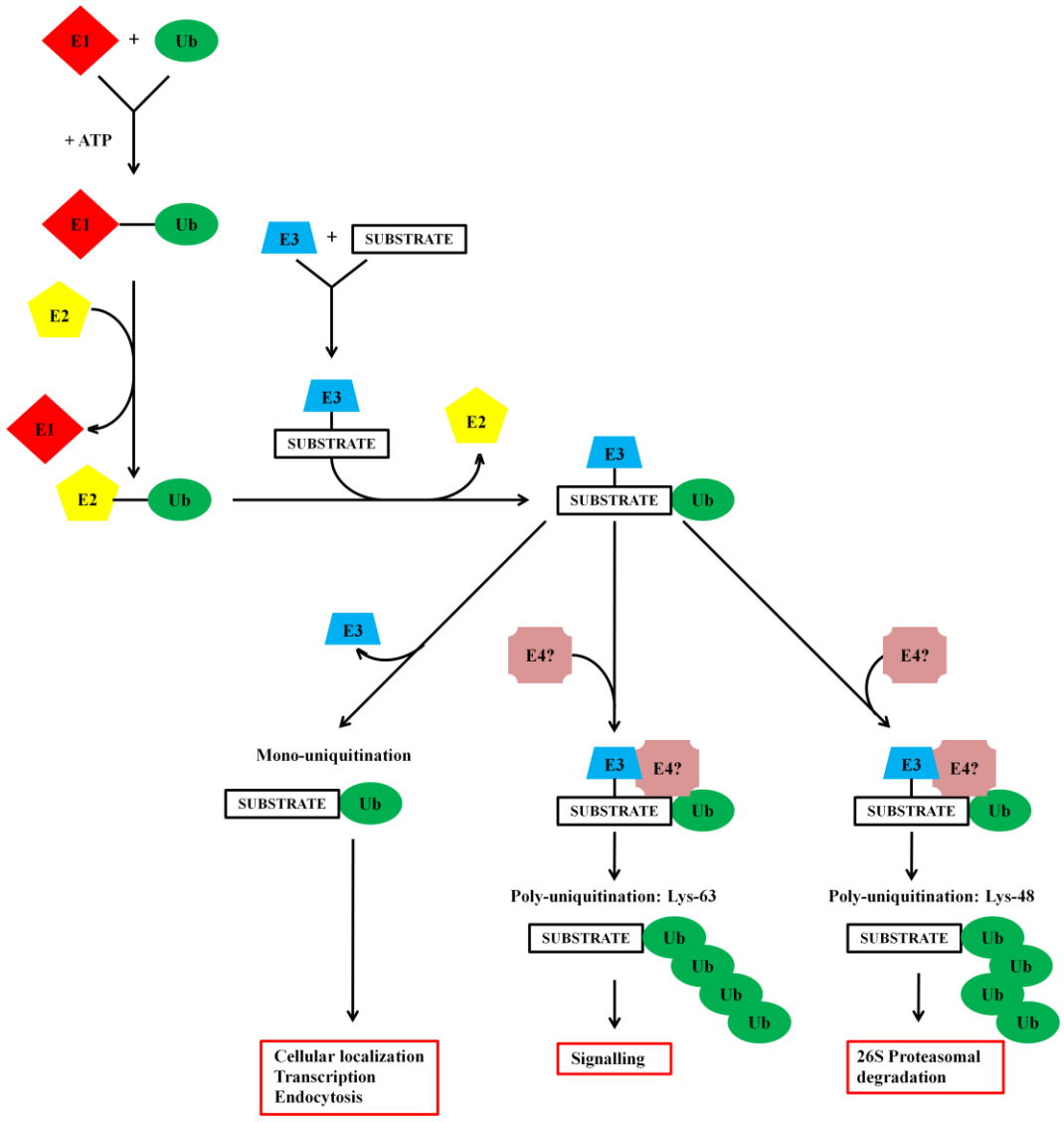

Figure 1.1: The ubiquitin conjugation pathway and fate of ubiquitinated proteins. Ubiquitin is conjugated onto a substrate via a cascade involving three enzymes: E1, E2 and E3, with a fourth enzyme, E4, necessary for poly-ubiquitin chain formation. Ubiquitination of a substrate can have a wide range of effects. In the case of poly-ubiquitin chains, the linkages connecting successive ubiquitin molecules together affect the chains topology, which enables selectivity on ubiquitin binding domains.

A form of post-translational modification closely-related to ubiquitination exists, which has been called ubiquitin-like modification (ULM). As with ubiquitin, ubiquitin-like proteins contain the same, conserved ubiquitin superfold (Welchman *et al.*, 2005). Ubiquitin-like proteins are also conjugated onto lysine

residues of a substrate – often in competition with ubiquitin and each other (Bossis & Melchior, 2006). Presently, ubiquitin-like modifications are assigned to 9 major classes (reviewed in Cajee *et al.*, 2012). ULM also occurs in the same sequential manner of an E1-E2-E3 cascade, although ubiquitin-like modifiers have their own specific E1, E2 and E3 enzymes. Interestingly some variations exist – for e.g. SUMOylation (small ubiquitin-related modifier) has been shown to occur either with or without an E3 enzyme. *In vitro*, in the absence of an E3 ligase, both conjugation and ligation of SUMO can be performed by the E2 conjugating enzyme Ubc9 (Desterro *et al.*, 1999; Okuma *et al.*, 1999), with E3 enzymes often enhancing the reaction. Like ubiquitination, ULM is also involved in diverse cellular processes. Furthermore, modification of a substrate with an ubiquitin-like protein can alter various characteristics of the substrate, such as its functionality and localization. The study of ULM is still relatively new compared to ubiquitination. As such, not only is our current knowledge of known components of ULM expected to expand, but also, new forms of ULM are likely to be discovered. One such putative form of ULM, DWNNylation, is discussed in Section 1.2.

1.2 *RBBP6* HOMOLOGUES AND PROTEIN PRODUCTS

In a study aimed at discovering novel binding partners of the *Homo sapiens* RETINOBLASTOMA (RB) tumour suppressor protein, one of the isolated proteins was named RETINOBLASTOMA-BINDING Q-PROTEIN 1 (RBQ-1) (Saijo *et al.*, 1995). This partial protein was the first recorded product produced from what would become known as the *retinoblastoma binding protein 6 (RBBP6)* gene (Sakai *et al.*, 1995).

RBBP6 is a single-copy gene purportedly present in all eukaryotes, although it has not been found in prokaryotes (Pugh *et al.*, 2006). Homologues of *RBBP6* have been described in various model organisms. In *Mus musculus*, the gene was discovered independently by two research groups and named *p53 associated cellular protein-testes derived* (*PACT*) (Simons *et al.*, 1997) and *proliferation*

potential protein-related (*P2P-R*) (Witte & Scott, 1997). The *Drosophila melanogaster RBBP6* homologue was characterized by our group, followed shortly afterwards by an unrelated research group. As a result, it too was given two different names: *Snama* (Mather, 2005; Mather *et al.*, 2005) and *mini-me* (Jones *et al.*, 2006). The *Saccharomyces cerevisiae* homologue was named *mutant pcf11 extragenic suppressor 1* (*MPE1*) (Vo *et al.*, 2001), whereas, more recently, the *Caenorhabditis elegans* homologue *retinoblastoma-binding-protein-like-1* (*RBPL-1*) has been described (Huang *et al.*, 2013). Given the various names accorded in these model organisms, the HGNC-approved (Seal *et al.*, 2011) name of *RBBP6* will be used generically for all homologues of this gene and resultant protein products.

With the exception of *S. cerevisiae*, all of the afore-mentioned species produce multiple RBBP6 proteins through alternative splicing. For example, in *H. sapiens*, three RefSeq (Priutt *et al.*, 2005) *RBBP6* transcripts exist: *RBBP6-1* (NM_006910.4) consists of all 18 exons, *RBBP6-2* (NM_018703.3) lacks exon 16, whereas *RBBP6-3* (NM_032626.5) contains the first three exons only (Figure 1.2). These three transcripts encode proteins of 1792 amino acids (a.a.) (NP_008841.2), 1758 a.a. (NP_061173.1) and 118 a.a. (NP_116015.2), respectively.

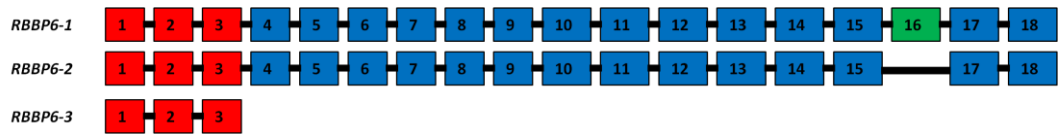

Figure 1.2: Exons constituting *H. sapiens RBBP6* transcripts. Exons 1-3 are common to all three *RBBP6* isoforms.

In *M. musculus*, only two RefSeq (Priutt *et al.*, 2005) transcripts exist. NM_011247.2 is comprised of all 18 exons and is the equivalent of *H. sapiens RBBP6-1*. The second *M. musculus RBBP6* transcript, NM_175023.3, is comprised of the first 3 exons only, thus it is equivalent to *H. sapiens RBBP6*-3.

Due to incomplete RefSeq (Priutt *et al.*, 2005) records, a third *M. musculus RBBP6* transcript (XM_006507470.1) is still officially predicted, even though Scott *et al.* (2005) demonstrated it to be the predominant splice version between NM_011247.2 and XM_006507470.1. Unsurprisingly, this predicted transcript is the equivalent of *H. sapiens RBBP6-2*, as it only lacks exon 16. Thus the first three exons are common to all *RBBP6* transcripts in both *H. sapiens* and *M. musculus*. In fact, these first three exons encode a domain which is found at the amino-terminus of all species' RBBP6 proteins. Moreover, this domain – called "Domain-With-No-Name" (DWNN) (Mather, 2005; Mather *et al*., 2005; Pugh *et al*., 2006) is found exclusively in RBBP6 proteins.

Analysis of RefSeq (Priutt *et al.*, 2005) RBBP6 proteins of the afore-mentioned species in ScanProsite (prosite.expasy.org/scanprosite/), reveals three highly conserved, structured domains present in most RBBP6 proteins (Table 1.1), which are described below.

(i) DWNN DOMAIN

As mentioned previously, the DWNN domain is the identifying feature of RBBP6 proteins. While the DWNN domain has little sequence similarity to ubiquitin – for example, *H. sapiens* DWNN and ubiquitin share only 18 % sequence homology (Mather, 2005; Mather *et al*., 2005; Pugh *et al*., 2006) – the tertiary structure of the DWNN domain is highly similar to that of ubiquitin (Figure 1.3).

A di-glycine motif, found at residues 75 and 76 in ubiquitin, serves in the covalent attachment of ubiquitin onto substrates. Interestingly, this di-glycine motif is also present in *H. sapiens* and *M. musculus* RBBP6 proteins, at residues 78 and 79, but is absent in RBBP6 orthologues. This has led to speculation that the DWNN domain may also be conjugated onto substrates in vertebrate species (Mather, 2005; Mather *et al.*, 2005; Pugh *et al.*, 2006). This possible form of post-translational modification has been termed DWNNylation (Pugh *et al.*, 2006) although it remains, as yet, unproven.

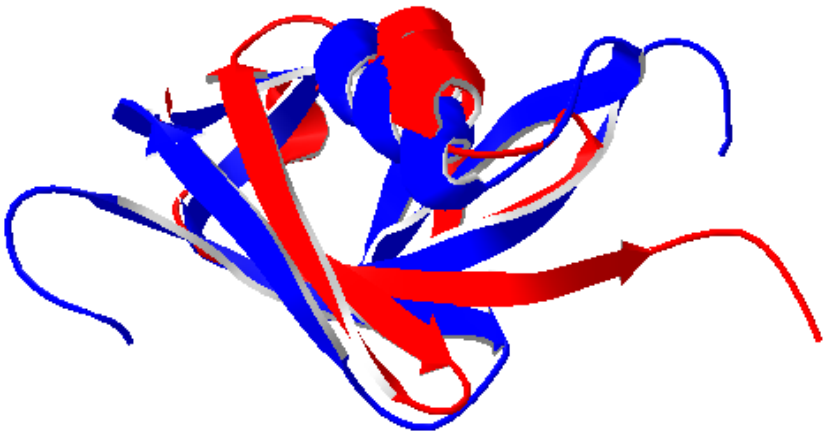

Figure 1.3: Super-imposed 3-D structures of *H. sapiens* DWNN domain and ubiquitin. In both polypeptides, beginning at the amino-terminus, the secondary structural arrangement occurs as ββαβ. While ubiquitin contains a final, fourth beta-sheet, in DWNN it is split into two shorter beta-sheets. Overall, these two polypeptides possess highly similar secondary and tertiary structures. DWNN domain (PDB: 2C7H) in blue and ubiquitin (PDB: 1UBQ) in red. Image generated using SPDBV version 3.7 with secondary structure analysis performed using PsiPred version 3.3 (bioinf.cs.ucl.ac.uk/psipred/).

(ii) ZINC FINGER DOMAIN

The second structured domain, common to all RBBP6 proteins except for the two DWNN domain proteins (*H. sapiens* RBBP6-3 and *M. musculus* RBBP6-2), is a CCHC-type zinc finger. Zinc fingers are typically involved in nucleic acid binding; interaction partners and functions of RBBP6 proteins will be discussed in Section 1.4.

(iii) ZINC FINGER-LIKE DOMAIN

The last, structured domain common to most RBBP6 proteins is a cysteine-rich, zinc finger-like domain. ScanProsite analysis identifies this domain as a RING domain in all RBBP6 proteins listed in Table 1.1 which contain this third domain, except for NP_001032975.1 in which it is classified as a U-box. RING domains

were originally characterized by a sequence of eight conserved cysteine and histidine residues, which adopt a cross-brace arrangement and bind two zinc-ions (Borden, 2000; Freemont, 2000). The sequence of conserved residues in the canonical RING domain, which are responsible for binding zinc, occurs in the order C3HC4 (Lovering *et al.*, 1993), in which the first zinc binding site is formed between C_1,C_2, C_5 & C_6, with C_3, H_4, C_7 and C_8 forming the second zinc binding site (Borden, 2000). Despite also adopting a cross-brace fold, the U-box domain does not contain these conserved cysteine and histidine residues and does not bind zinc-ions; instead, the U-box is stabilized by hydrogen bonds and salt bridges (Aravind & Koonin, 2000; Ohi *et al.*, 2003). While both the RING domain and U-box domains were found to be involved in the ubiquitin proteasome pathway, originally it was thought that only RING-containing proteins could perform E3 ubiquitin ligase activity; U-boxes were considered E4s, or accessory proteins to E3s which are involved in poly-ubiquitin chain formation (Koegl *et al.*, 1999). This view was later changed when it was demonstrated by Hatakeyama *et al.* (2001) that some U-boxes were also able to mediate E3 ubiquitin ligase activity. Furthermore, RINGs have been shown to reversibly bind divalent metal-ions other than Zn^{2+}, such as Cu^{2+} (Lovering *et al.*, 1993), Co^{2+} (Roehm & Berg, 1997) and Cd^{2+} (Lovering *et al.*, 1993; Kappo *et al.*, 2012). Moreover, RINGs were discovered in which a single, divalent metal-ion – chelating cysteine or histidine residue was replaced with a lesser-conserved residue, such as glycine or aspartate, often with the effect that only one metal-ion co-ordinating site would form. Taken together, the ability to bind different divalent metal-ions, as well as the occurrence of RING domains with backbone residues other than cysteine or histidine, is thought to impart binding and functional specificity for RING-containing proteins (Roehm & Berg, 1997; Borden, 2000).

The three, structured domains, described above, have collectively been named the DWNN Catalytic Module (DCM) (Antunes, 2008). As the DCM has been shown to mediate ubiquitin ligase activity (Antunes, 2008; Chibi *et al.*, 2008), it is thought that the DCM implicates RBBP6 proteins in protein turnover (Kappo *et al.*, 2012). Most RBBP6 proteins also contain multiple, unstructured domains –

typically tracts enriched in particular amino acids, which are present after the DCM. Of particular importance in RBBP6 proteins are serine and arginine tracts – the regions of which overlap (Table 1.1). These two amino acids are mostly present as SR and RS dipeptides and form a domain called the RS-domain.

The RS-domain is a common feature of both SR proteins, and SR-like proteins. Initially, SR proteins were discovered as a group of proteins with a shared phosphoepitope, as detected by a monoclonal antibody: mAb 104 (Roth *et al.*, 1990). Follow-up studies utilizing this same antibody (Roth *et al.*, 1991; Zahler *et al.*, 1992) revealed an identical banding pattern (proteins with conserved sizes of 20, 30, 40, 55 and 75 kDa), in immunoblots prepared from numerous species, including *D. melanogaster* and *H. sapiens*. Furthermore, these six proteins (two proteins of 30 kDa were discovered) were all shown to be involved in splicing activities, and contain an amino-terminal RNA recognition motif (RRM) and a carboxyl-terminal RS-domain (Zahler *et al.*, 1992). Moreover, these proteins could be purified to homogeneity through a biochemical approach which involved two precipitation steps, employing ammonium sulphate and magnesium chloride (Roth *et al.*, 1991; Zahler *et al.*, 1992). Thus, these factors were the initial requirements for the classification of a protein as a SR protein (reviewed in Fu, 1995).

Table 1.1: Conserved domains and features of RBBP6 proteins from select organisms

ORGANISM	REFSEQ ACCESSION	PROTEIN	LENGTH (A.A)	MOLECULAR MASS (kDa)	ISO-ELECTRIC POINT	DWNN	ZINC FINGER (CCHC-TYPE)	ZINC FINGER-LIKE	SERINE-RICH TRACT	ARGININE-RICH TRACT	LYSINE-RICH TRACT	PROLINE-RICH TRACT	BIPARTITE NLS
Homo sapiens	NP_008841.2	RBBP6-1	1792	201.3	10.06	4-76	161-175	259-300 (RING)	682-772; 1688-1726	679-835	999-1788	337-410; 548-608	≤ 6
Homo sapiens	NP_061173.1	RBBP6-2	1758	197	10.05	4-76	161-175	259-300 (RING)	652-738; 1654-1692	672-801	965-1754	337-410; 548-608	≤ 6
Homo sapiens	NP_116015.2	RBBP6-3	118	13.2	9.13	4-76	-	-	-	-	-	-	-
Mus musculus	NP_035377.2	RBBP6-1 (PACT)[#]	1790	199.4	10.06	4-76	162-176	260-301 (RING)	683-773; 1690-1726 (376-405)	680-836	1000-1784	338-411; 549-609	≤ 6
Mus musculus	NP_778188.1	RBBP6-2	123	13.8	9.11	4-76	-	-	-	-	-	-	-
Mus musculus	XM_006507470.1	RBBP6-3* (P2P-R)[#]	1756	195.3	10.25	4-76	162-176	260-301 (RING)	653-739; 1656-1692 (376-405)	673-802	966-1750	338-411; 549-609	≤ 6
Drosophila melanogaster	NP_611884.1	SNAMA-PA	1231	139	10.16	3-76	152-167	217-258 (RING)	-	562-647; 998-1097	805-848; 1188-1231	-	≤ 5
Drosophila melanogaster	NP_001246487.1	SNAMA-PB	494	55.6	6.3	3-76	152-167	217-258 (RING)	-	-	-	-	-
Caenorhabditis elegans	NP_001032975.1	RBPL-1 isoform A	1128	130.2	6.59	4-75	158-172	236-314 (U-box)	498-517 (1030-1072)	479-609 (891-967)	765-1019	375-463	≤ 14
Caenorhabditis elegans	NP_001032976.1	RBPL-1 isoform B	363	42.6	9.94	4-75	158-172	-	(265-307)	-	169-254	-	≤ 4
Saccharomyces cerevisiae	NP_012864.1	MPE1P	441	49.6	6.15	5-78	182-196	-	-	-	-	402-425	-

The locations of conserved structured, and unstructured, domains are shown for RBBP6 proteins, along with bio-physical properties. With the exception of *H. sapiens*, the animal species represented here are typical model organisms in which *RBBP6* has been studied. * This protein is not yet officially recognized but is thought to exist. Numbers in parentheses represent low-confidence predictions. NLS = nuclear localization signal.

Proteins were also discovered which had an RS-domain but did not meet all the criteria of the previous paragraph. For example, modifications to the original precipitation technique, such as the addition of a phosphorylation step prior to precipitation (Blencowe *et al.*, 1995), resulted in additional proteins co-purifying along with the six originally discovered SR proteins. The present criteria for the classification of a protein as a SR protein require that it must contain at least one amino-terminal RRM, and a downstream RS-domain which is at least 50 residues long and comprised of at least 40 % RS/SR dipeptides (Manley & Krainer, 2010). Simultaneously, the same authors also proposed renaming the resulting twelve proteins, which conformed to the new criteria, in line with HGNC convention. Thus, the twelve *H. sapiens* SR proteins are named SRSF1-12 (SR splicing factor), with all other RS-domain-containing proteins being unaffected by the naming convention and collectively referred to as SR-like proteins.

The absence of an amino-terminal RRM therefore establishes some RBBP6 proteins as SR-like proteins. Interestingly, SRSF7 (previously 9G8) also contains a CCHC-type zinc finger domain between its amino-terminal RRM domain and carboxyl-terminal RS-domain, although these two domains are distinctly different. The presence of the RS-domain therefore suggests that some RBBP6 proteins may be involved in activities such as mRNA processing, transcription and chromatin remodelling (Boucher *et al.*, 2001). The activities of SR-like proteins are therefore similar to those of SR proteins, which are involved in all forms of mRNA processing (constitutive and alternative splicing, nuclear export, translation and degradation) (reviewed in Twyffels *et al.*, 2011). In fact, the *S. cerevisiae* RBBP6 protein, MPE1P, has been shown to be involved in mRNA processing (Vo *et al.*, 2001).

Most RBBP6 proteins also contain at least one lysine-rich tract, which manifests in the form of multiple bipartite nuclear localization signals (NLS). These NLS are found scattered throughout the RS-domain and are in addition to the nuclear-targeting sequences, comprising mostly arginine and serine residues, found to target RS-domain-containing proteins to nuclear speckles (Li & Bingham, 1991;

Hedley *et al.*, 1995). Taken together, RBBP6 proteins are thought to localize to the nucleus, particularly nuclear speckles. As a result of these arginine-rich and lysine-rich tracts, most RBBP6 proteins also have very high isoelectric points, as arginine and lysine are, respectively, the two most basic amino acids.

1.3 SNAMA

The *D. melanogaster RBBP6* homologue will be referred to by its first accorded name: *Snama*. Two *Snama* transcripts are known to exist. *Snama-RA* (FBtr0072249) was experimentally verified by both Northern blot analysis and reverse-transcription PCR (Mather, 2005; Mather *et al.*, 2005; Jones *et al.*, 2006). More recently, based on high-throughput sequencing data from Daines *et al.* (2011) which identified a splice junction in intron 7, *Snama-RB* (FBtr0307174) was predicted by FlyBase genome annotators. Interestingly, this transcript was visualized by Northern blot analysis after a longer period of exposure than was required to visualize *Snama-RA* (Jones *et al.*, 2006), although it was not further explored as a second, possible *Snama* transcript. Consequently our group has unpublished data, generated via reverse-transcription PCR, verifying the existence of *Snama-RB* (Cajee, U.-F., in preparation).

To date, no SNAMA proteins have been experimentally verified, despite immunological approaches undertaken in previous studies (Mather, 2005; Jones *et al.*, 2006; Rakgotho, 2007; Hull, 2012). *Snama-RA* is predicted to encode a "full-length" protein of 1231 amino acids (SNAMA-PA: FBpp0072158; NP_611884.1), whereas *Snama-RB* is predicted to encode a C-terminally truncated, alternatively-spliced variant of 494 amino acids (SNAMA-PB: FBpp0298003; NP_001246487.1). These two predicted proteins (Figure 1.4) are identical for the first 493 residues; the inclusion of intron 6-7 in *Snama-RB* results in the substitution of Ser^{494} with Arg^{494} which is immediately followed by a stop codon. These two putative proteins would manifest on a SDS-PAGE gel as proteins of 139 kDa (SNAMA-PA) and 55.6 kDa (SNAMA-PB).

Both predicted proteins contain a 258 residue long DCM, comprising the DWNN domain, CCHC-type zinc finger domain and zinc finger-like domain (Table 1.1 and Figure 1.4). The *D. melanogaster* DWNN domain shares 23.5 % sequence homology with *D. melanogaster* ubiquitin (Mather, 2005; Mather *et al.*, 2005), which is 5 % more than the *H. sapiens* counterparts. However, unlike in *H. sapiens* and *M. musculus* RBBP6 proteins, there is no di-glycine at the end of the DWNN domain. Moreover, given the absence of a stand-alone DWNN-domain protein in *D. melanogaster*, it seems unlikely that DWNNylation could occur in *D. melanogaster*, unless cleavage occurs on an uncharacterized recognition site and through an uncharacterized mechanism.

The zinc finger-like domain of RBBP6 proteins differs from the canonical C3HC4 sequence, as the fourth zinc-chelating residue is degenerate. In *H. sapiens*, *M. musculus* and *C. elegans*, the histidine has been replaced with an asparagine; whereas in *D. melanogaster* and *S. cerevisiae*, a serine and isoleucine are found, respectively. As the hydroxyl group of the serine is ionizable, it is thought that serine, in the *D. melanogaster* RING, can substitute for the sulfhydryl group of cysteines ordinarily involved in chelating zinc-ions (Mather, 2005). Recently, Ying *et al.* (2011) have classified RING domains found in *D. melanogaster* proteins into eight sub-divisions, of which the U-box domain is a member. Specific mention was made of SNAMA in that its RING domain was so uncommon, that it was grouped along with the "RING-G"-type, with a sequence of C3HGC3. To further complicate matters, Kappo *et al.* (2012) demonstrate that the *H. sapiens* RBBP6 RING domain is of the C4C4-type and not C3NC4 as was always thought to be the case. They further demonstrate, through NMR analysis, that the RING domain is stabilized by hydrogen bonding between backbone residues – as occurs in U-box domains, as well as through zinc-ion chelation – which is the hallmark of RING domains. Thus they suggest that the *H. sapiens* RBBP6 RING domain is an intermediary between classical RING and U-box domains. This view is supported by the work of Antunes (2008), in which SNAMA auto-ubiquitination was shown to occur both in the presence and absence of zinc-ions, *in vitro*. Furthermore, ScanProsite analysis classifies the *C.*

elegans zinc finger-like domain as a U-box (Table 1.1). Thus, uncertainty exists with regards to the classification of the SNAMA zinc finger-like domain: whether it is a C3SC4 RING, a C3HGC3 RING, a C4C4 RING, or a hybrid between a RING and U-box.

Downstream of the DCM, there is a minor RS-domain common to both predicted SNAMA proteins, in which five RS-dipeptides are present. This 69 residue stretch, comprising residues 386-454, contains 24 serine and arginine residues. While SNAMA-PB terminates shortly after this region, this point represents only around one third of the SNAMA-PA sequence. The C-terminal two thirds of SNAMA-PA is highly enriched in SR/RS dipeptides. Of the 92 residues over the region 998-1089, 46 are arginine or serine, thus qualifying this region as a true RS-domain. SNAMA-PA also contains two lysine-rich tracts and five bipartite nuclear localisation signals. Combined, arginine and lysine residues account for over 20 % of the SNAMA-PA sequence, thus imparting SNAMA-PA with an extremely basic pI of 10.16 while SNAMA-PB, which lacks this C-terminal extension, has a pI of 6.74. Overall, the difference in the C-terminal domain, as well as the stark difference in pI, between SNAMA-PA and SNAMA-PB suggests that these two putative proteins localize to different cellular locations and may, possibly, also perform different biological functions.

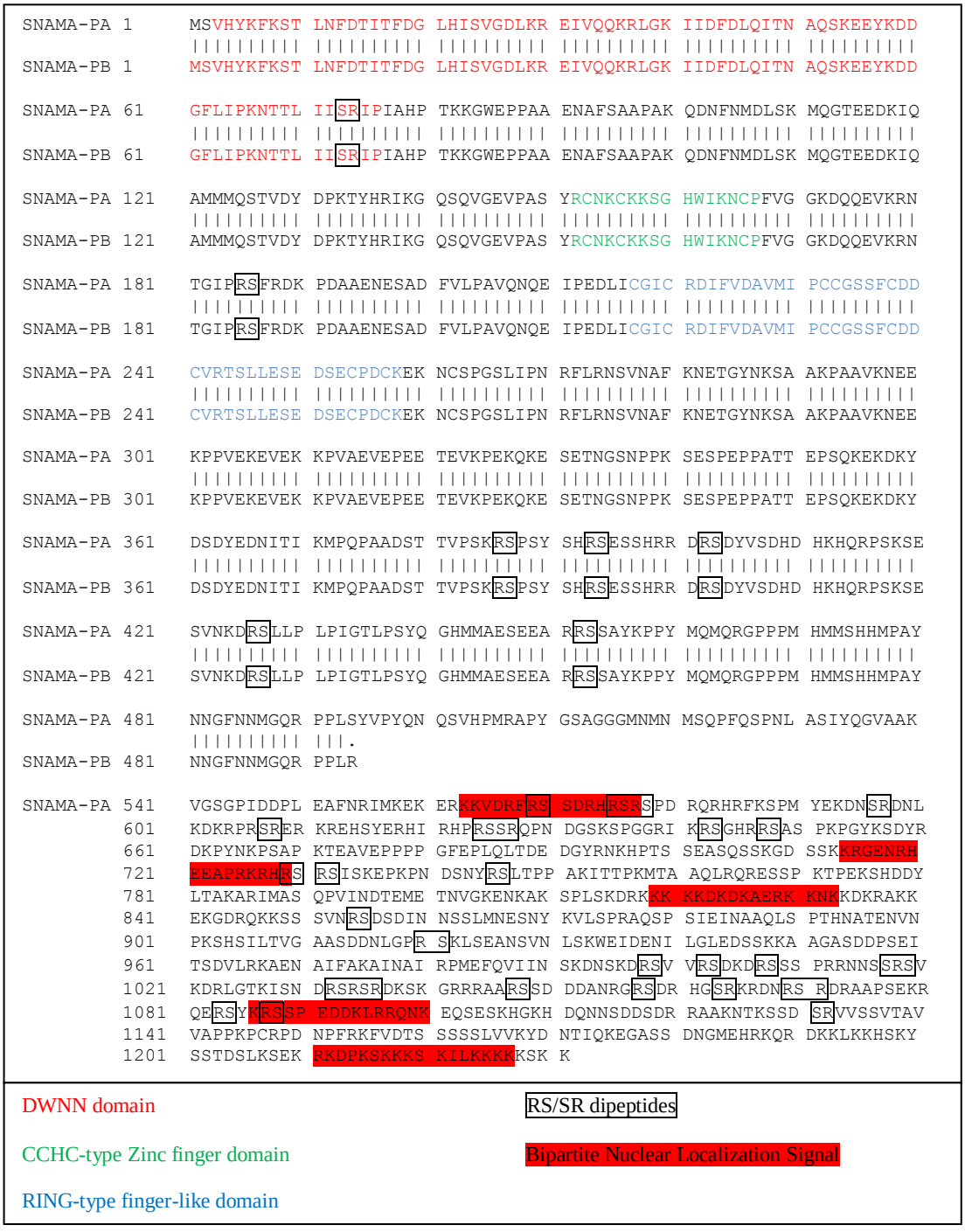

Figure 1.4: Domains and features of putative SNAMA proteins. Sequence alignment between the two predicted SNAMA polypeptides demonstrates that SNAMA-PB is a C-terminally truncated, alternatively-spliced variant of SNAMA-PA. SNAMA-PB therefore lacks the C-terminal extension containing the RS-domain and various other features described in-text.

1.4 RBBP6 FUNCTIONS AND INTERACTING PROTEINS

RBBP6 proteins are known to function in a wide range of processes. When the first RBBP6 protein, RBQ, was discovered, interest was mostly focussed on the role played by RBBP6 proteins in cell-cycle control. RBBP6 null mutations have been shown to be lethal (Vo *et al.*, 2001; Mather, 2005; Mather *et al.*, 2005; Li *et al.*, 2007; Huang *et al.*, 2013) with effects such as p53 accumulation and altered cell-cycle profiles, such as polyploidy, manifesting (Jones e*t al.*, 2006; Huang et al., 2013). These effects were due, in part, to the fact that both *H. sapiens* and *M. musculus* RBBP6 proteins were shown to bind Rb (Saijo *et al.*, 1995; Sakai *et al.*, 1995; Simons *et al.*, 1997; Witte & Scott, 1997) and p53 (Simons *et al.*, 1997; Gao *et al.*, 2002) tumour suppressor proteins. Surprisingly, only mammalian RBBP6 proteins have Rb and p53 binding domains (Pugh *et al.*, 2006).

RBBP6 proteins have been shown to be highly expressed in testis and / or oocytes (Simons *et al.*, 1997; Yoshitake *et al.*, 2004; Mather, 2005; Huang *et al.*, 2013) as well as intestine (Huang *et al.*, 2013). Expression of RBBP6 proteins is highest in proliferating tissues and lowest in differentiated cells (Witte & Scott, 1997; Jones *et al.*, 2006). RBBP6 has a role to play in transcription-related activities as RBBP6 forms part of the estrogen-transcription complex (Peidis *et al.*, 2010) and localizes to nuclear speckles (Gao & Scott 2002; Scott *et al.*, 2003; Huang *et al.*, 2013). Further RBBP6 activities occurring in nuclear speckles include mRNA processing (Vo *et al.*, 2001), specifically as a component of the Yeast Cleavage and Polyadenylation Factor (CPF). These nuclear-related activities likely involve the RS-domain of RBBP6 proteins.

More recently, attention has shifted towards a role for RBBP6 proteins in protein quality control (Li *et al.*, 2007; Antunes, 2008; Chibi *et al.*, 2008; Kappo *et al.*, 2012). RBBP6 has been shown to facilitate HDM2-mediated ubiquitination of p53 (Li *et al.*, 2007), thereby suggesting one mechanism in which mammalian RBBP6 proteins interact with tumour suppressors to regulate cell cycle control. HDM2-mediated ubiquitination occurred independently of a functional RING domain

thus suggesting that RBBP6 acts as an E4 enzyme (Li *et al.*, 2007). RBBP6 proteins have been shown to possess E3 ubiquitin ligase activity: Chibi *et al*. (2008) demonstrated that RBBP6 could bind Y-box binding protein 1 (YB-1) through its RING finger-like motif, *in vivo*. They went on further to demonstrate that YB-1 was subsequently ubiquitinated and degraded via the proteasome. Furthermore, Antunes (2008) demonstrated auto-ubiquitination of the SNAMA DCM, *in vitro*. As mentioned previously, although RBBP6 proteins are identified by an N-terminal DWNN domain which is thought to be a putative form of ubiquitin-like modification, DWNNylation remains unproven. Although the ubiquitin-related activities described in this paragraph are more reliant on the DCM than the unstructured domains of the C-terminus, RBBP6 proteins likely have multiple roles within a cell. Known binding partners of RBBP6 proteins are presented in Table 1.2, with numerous others possibly awaiting discovery.

Table 1.2: Experimentally verified protein interaction partners of RBBP6 proteins

BINDING PROTEIN	**REFERENCE**
RB (Retinoblastoma)	Saijo *et al.*, 1995; Sakai *et al.*, 1995; Simons *et al.*, 1997; Witte & Scott, 1997
P53	Li *et al.*, 2007; Simons *et al.*, 1997
HDM2 (Human-double-minute 2)	Li *et al.*, 2007
YB-1 (Y-box binding protein 1)	Chibi *et al.*, 2008
Cdc2	Scott *et al.*, 2003
SAF-B (scaffold attachment factor – B)	Scott *et al.*, 2003
Nucleolin	Scott *et al.*, 2003
SRPK1/ SRSF (SR-rich splicing factor protein kinase)	Scott *et al.*, 2003
ERα (Estrogen receptor alpha)	Peidis *et al.*, 2010
SRC-1 (steroid receptor co-activator 1)	Peidis *et al.*, 2010

The few proteins listed here represent only experimentally verified, direct interaction partners of RBBP6 proteins. These proteins are mostly involved in activities centred on cell-cycle control and gene transcription.

1.5 AIMS AND OBJECTIVES

The primary aim of this study was to characterize the RBBP6 family of proteins in *Drosophila melanogaster* and to test the hypothesis that SNAMA may be involved in apoptosis in a p53-dependant manner.

The specific objectives of this study were to:

- Heterologously express and purify a fusion protein comprised of the SNAMA DCM fused to the GST and HIS expression tags of pET41a(+) vector. This objective also requires the expression and purification of a negative control protein, which is comprised of only the expression tags of the pET41a(+) vector.
- Prepare and purify chicken antibodies raised against the SNAMA DCM protein.
- Determine, through Western blot analysis, the number of SNAMA isoforms as well as their spatio-temporal regulation.
- Identify, through co-immunoprecipitation assays, potential interaction partners of SNAMA proteins.

CHAPTER TWO

2. MATERIALS AND METHODS

All chemicals utilized in this project were, unless otherwise stated, purchased from either Sigma-Aldrich or Merck Biosciences.

2.1 KITS

Table 2.1: Commercial kits or kit components utilized throughout this study

DESCRIPTION	SUPPLIER	CATALOGUE NUMBER
Aldehyde-agarose	Sigma-Aldrich	A9951
Anti-Chicken IgY Peroxidase conjugate	Sigma-Aldrich	A9046
Bio-Rad protein assay	Bio-Rad	500-001
BioMax® light film	Sigma-Aldrich	Z373508
Chicken IgY Precipitating Resin	GenScript	L00405
Chromatography columns 30 ml (Econo-Pac)	Bio-Rad	732-1010
DNAse	Sigma-Aldrich	AMP-D1
Glutathione-Agarose	Sigma-Aldrich	G4510
HIS-Select® Nickel Affinity Gel	Sigma-Aldrich	P6611
PageRuler™Prestained Protein Ladder*	Fermentas Life Sciences	SM0671
Polyscreen PVDF transfer membrane	Perkin Elmer Life Sciences, Inc	NEF1002
Protease Inhibitor Cocktail Set II	Calbiochem (Merck Biosciences)	539132
SnakeSkin™ Pleated Dialysis Tubing	Thermo Scientific	68035
Stirred cell ultrafiltration unit, model 8003	Amicon Merck	5125

Stirred cell ultrafiltration filters, NMWL 10000	Amicon Merck	PLGC02510
Sodium cyanoborohydride, 5 M solution in aqueous 1 M sodium hydroxide	Sigma-Aldrich	296945
SuperBlock® Dry Blend Blocking Buffer in TBS	Pierce	37545
SuperSignal® West Pico Chemiluminescent Substrate Kit	Pierce	34080

* See Figure A1 for band sizes of the PageRuler™ molecular weight marker.

2.2 BACTERIAL STRAIN AND VECTOR

Table 2.2: Vector and bacterial strain utilized in this study

VECTOR / BACTERIA	STRAIN	ANTIBIOTIC RESISTANCE	SUPPLIER	CATA-LOGUE NUMBER
pET-41a(+)*	Not applicable	Kanamycin	Novagen (Merck Biosciences)	70556-3
Escherichia coli BL21 (DE3) pLysS	F- *ompT hsdSB* (rB - mB-) *gal dcm* (DE3) pLysS (Cam^R)	Chloramphenicol	Invitrogen	C6060-10

* See Figure B1 for an overview of the various features of the pET-41a(+) expression vector.

2.3 MEDIA

Table 2.3: Composition of media utilized in organism maintenance and growth

MEDIA	COMPOSITION
Apple juice agar plates	5 % sucrose; 8 % agar; 25 % apple juice; 0.6 % methyl paraben; 0.1 % Streptomycin
Cornmeal molasses	1 % sucrose; 1.05 % agar; 3 % cornmeal; 5 % molasses 2 % dry yeast; 0.15 % methyl paraben

Luria-Bertani (LB) agar plates	1 % tryptone; 0.5 % yeast extract; 1 % sodium chloride; 1.5 % agar
LB broth	1 % tryptone; 0.5 % yeast extract; 1 % sodium chloride

2.4 MAINTENANCE OF RESEARCH ORGANISMS

2.4.1 Maintenance of bacterial clones

Bacterial clones utilized in this project were maintained on LB agar plates containing 30 µg/ml Kanamycin and 100 µg/ml Chloramphenicol. Plates were restreaked every three weeks.

2.4.2 Maintenance of *Drosophila melanogaster*

Wild-type (Oregon-R) flies, reared at 25 ºC, were utilized in this study. The stock fly strain was maintained in cornmeal molasses vials, changed weekly. The vials were supplemented with fresh yeast prior to transfer of flies. For experiments, however, between 5000 and 10000 adult flies were housed in a cage and provided with apple juice agar plates containing fresh yeast spread on the surface of the plate. Plates were changed every 12 hours during maintenance, or alternatively, as stipulated when collecting embryos for experiments. In experiments requiring stages of development post-embryo stage, the embryos were transferred to cornmeal molasses media to allow for further development.

2.5 GENERAL MOLECULAR BIOLOGY METHODS

The methodologies described – and buffers listed – in this section, are common molecular biology techniques which were performed in multiple experiments of this study. Sections 2.6-2.9 detail more specific methodologies, and buffer components, pertaining to the specific investigation.

2.5.1 General buffers and solutions

Table 2.4: General buffers and solutions utilized in various experiments throughout this study

BUFFER / SOLUTION	COMPOSITION
Immobilized metal-ion affinity chromatography	
IMAC cleansing buffer	50 mM sodium phosphate buffer, pH 8.0; 6 M guanidine hydrochloride
IMAC elution buffer	50 mM sodium phosphate buffer, pH 6.5; 0.1 M NaCl; 8 M urea; 350 mM imidazole
IMAC equilibration buffer	50 mM sodium phosphate buffer, pH 8.0; 0.1 M NaCl; 8 M urea
IMAC wash buffer	50 mM sodium phosphate buffer, pH 6.5; 0.1 M NaCl; 8 M urea
Glutathione affinity chromatography	
GSH cleansing buffer 1	0.1 M borate buffer, pH 8.5; 0.5 M NaCl
GSH cleansing buffer 2	0.1 M acetate buffer, pH 4.5; 0.5 M NaCl
GSH elution buffer	0.1 M sodium phosphate buffer, pH 8.0; 0.1 M NaCl; 15 mM reduced-glutathione
GSH wash buffer	0.1 M sodium phosphate buffer, pH 8.0; 0.1 M NaCl; 0.1 % Triton™ X-100
SDS-PAGE	
SDS-PAGE electrode running buffer (5X)	125 mM Tris base; 1 M glycine; 17 mM SDS; pH 8.2
SDS-PAGE sample buffer (5X)	0.313 M Tris-HCl, pH 6.8; 10 % glycerol; 2 % SDS; 5 % β-mercaptoethanol; 0.05 % bromophenol blue
SDS-PAGE separating layer	0.375 M Tris-HCl, pH 8.8; 3.5 mM SDS; 12 % (37.5:1) acrylamide/bis-acrylamide; 2.2 mM ammonium persulphate; 0.1 % TEMED
SDS-PAGE stacking layer	62.5 mM Tris-HCl, pH 6.8; 3.5 mM SDS; 4 % (37.5:1) acrylamide/bis-acrylamide; 4.4 mM ammonium persulphate; 0.2 % TEMED
Gel staining and destaining	
Coomassie brilliant blue staining solution	0.1 % Coomassie blue R-250; 40 % ethanol; 10 % glacial acetic acid; 50 % ddH_20
Destain solution	40 % ethanol; 10 % glacial acetic acid; 50 % ddH_20

Silver stain developer	2 % Na_2CO_3; 0.02 % formalin
Silver stain fixative	30 % ethanol; 10 % glacial acetic acid
Silver stain sensitizer solution	0.02 % sodium thiosulphate
Silver stain staining solution	0.1 % silver nitrate
Silver stain wash solution	20 % ethanol
	Immunoblotting
PVDF membrane wash buffer	1X PBS buffer, pH 7.4; 0.1 % Tween®20
Towbin buffer	25 mM Tris base; 192 mM glycine; 0.1 % SDS; 20 % methanol; pH 8.2
	Generic
Chromatography column storage buffer	50 mM Tris-HCl, pH 7.5; 0.1 M NaCl; 1mM NaN_3
PBS buffer (10X)	1.4 M NaCl; 27 mM KCl; 0.1 M Na_2HPO_4; 18 mM KH_2PO_4; pH 7.4

2.5.2 Purification of tagged proteins

Column affinity chromatography was used to purify recombinantly-expressed proteins produced in this study. All parts of the procedure in which sample was present were performed under gravity, while a peristaltic pump was used at all other times. Prepared columns were preserved in chromatography column storage buffer between experiments.

2.5.2.1 Immobilized metal-ion affinity chromatography

A 30 ml chromatography column, containing 2 ml of HIS-Select® nickel affinity gel, was prepared according to the manufacturers' instructions. Purification of SNAMA DCM protein via its two HIS-tags was performed according to the resin manufacturer's instructions.

2.5.2.2 Glutathione affinity chromatography

A 30 ml chromatography column, containing 1 ml of Glutathione-Agarose was prepared according to the manufacturers' instructions. Purification of GST-tagged proteins was performed according to the resin manufacturer's instructions.

2.5.3 Protein concentration

Samples were concentrated via stirred-cell ultrafiltration, using filters with a MWCO of 10 kDa, according to the manufacturer's instructions.

2.5.4 Protein quantification

Protein concentration was determined via the Bio-Rad protein assay according to the manufacturer's instructions.

2.5.5 Sodium dodecyl sulphate – polyacrylamide gel electrophoresis

The denaturing, discontinuous gel method of Laemmli (1970) was followed, consisting of 12 % separating and 4 % stacking layers. Samples were boiled in 1 X SDS-PAGE sample buffer for 5 minutes prior to loading. Unless otherwise stated, 10 µg of protein, as determined in Section 2.5.4 prior to SDS-PAGE analysis, was loaded per lane. PageRuler® Prestained protein ladder was added to all gels for size comparison. Electrophoresis was performed at a constant voltage of 100 V.

2.5.6 Coomassie brilliant blue staining and destaining

Polyacrylamide gels were stained overnight in Coomassie brilliant blue staining solution and destained, the following morning, in destain solution until desired.

2.5.7 Silver staining

Silver staining was performed according to the MS-compatible methodology of Shevchenko *et al.* (1996).

2.5.8 Immunoblotting

Following SDS-PAGE separation, samples were transferred to Polyscreen® PVDF membrane using Towbin buffer (Towbin *et al.*, 1979) at a constant current of 295 mA for 90 minutes. The membrane was placed in SuperBlock® blocking buffer for at least one hour at 4 ºC before being subjected to three successive, 40 ml changes of PVDF membrane wash buffer for 10 minutes at a time. The membrane was probed for one hour with a 1: 40000 primary antibody dilution; this equated

to a total of 0.45 µg of anti-SNAMA DCM antibody, or alternatively, 3.16 µg of pre-immune serum which was used as a negative control probe. The membrane was washed thrice each time after exposure to primary and secondary antibodies. Anti-Chicken IgY Peroxidase conjugate was used as a secondary antibody, at a dilution of 1: 60000. SuperSignal® Chemiluminescent substrate kit was utilized for signal generation, which was captured on BioMax® light film.

2.5.9 Analysis of gels and films

All gels and films were scanned on a Bio-Rad GS800 calibrated densitometer (cat. no. 170-7980). Analysis, such as detection of bands and determination of band sizes, was performed using the accompanying Quantity One® software.

2.6 EXPRESSION, EXTRACTION AND REFOLDING OF SNAMA DCM PROTEIN

SNAMA DCM protein was prepared in order to perform pulldown assays, with a view towards investigating the biological functions of RBBP6 proteins in *Drosophila*.

Table 2.5: Buffers and solutions used in refolding experiments

BUFFER/SOLUTION	COMPOSITION
Protein extraction and inclusion body treatment	
Bacteria lysis buffer 1	50 mM Tris-HCl, pH 7.5; 2 mM EDTA, pH 7.5; 200 mM NaCl; 0.1 mM PMSF; 200 µg/ml lysozyme
Inclusion body solubilisation buffer 1	50 mM sodium phosphate buffer, pH 8.0; 8 M urea; 0.1 M NaCl; 1 mM PMSF; 2 mM DTT
Inclusion body wash buffer 1	50 mM Tris-HCl, pH 7.5; 2 mM EDTA, pH 7.5; 200 mM NaCl; 0.1 mM PMSF; 1 % Triton™ X-100

Refolding of SNAMA DCM protein via dialysis	
Dialysis buffer 1 (No additives)	50 mM sodium phosphate buffer, pH 8.0; 0.1 M NaCl
Dialysis buffer 2 (With additives)	50 mM sodium phosphate buffer, pH 8.0; 0.1 M NaCl; 5 mM EDTA; 5 mM β-cyclodextrin; 50 mM arginine
Dialysis buffer 3 (Redox)	50 mM sodium phosphate buffer, pH 8.0; 0.1 M NaCl; 5 mM β-cyclodextrin; 50 mM arginine; 0.1 mM $ZnCl_2$; 1 mM reduced glutathione; 0.3 mM oxidized glutathione

2.6.1 Revival of bacterial clones from glycerol stocks and protein expression

The "control" and "SNAMA DCM" bacterial clones utilized throughout this study were prepared by a former colleague (Antunes, 2008). This project was begun by reviving glycerol stocks of the two aforementioned bacterial clones, which were thereafter maintained on LB agar plates. For protein expression, a 2 ml starter culture of LB broth was inoculated with the required bacterial clone, which was then incubated overnight at 37 ºC with orbital rotation of 250 revolutions per minute. The following day, the 2 ml starter culture was transferred to a two litre Erlenmeyer flask containing 500 ml LB broth supplemented with 10 µg/ml Kanamycin. Cells were grown to an $O.D_{600}$ of between 0.5 and 0.6 before protein expression was induced via the addition of IPTG to a concentration of 0.2 mM. Induction lasted for 5 hours, after which the culture was aliquoted into polypropylene tubes and centrifuged at 1500 x g for 10 minutes at 4 ºC. After removal of the LB broth supernatant, bacterial pellets were frozen at -20 ºC until further processed.

2.6.2 Protein extraction and inclusion body treatment

Thawed cells, prepared as described in Section 2.6.1, were resuspended in bacteria lysis buffer 1 in a ratio of 10 ml lysis buffer per 100 ml original culture volume. The extract was maintained on ice for the entire protein extraction procedure.

Three microlitres of DNAse was added to the extract which was then subjected to mild sonication until all cellular components were solubilized. After a one hour incubation period with brief, periodic agitation, the extract was centrifuged at 11000 x g for 10 minutes at 4 ºC after which the supernatant was discarded.

The resultant pellet fraction was resuspended in inclusion body wash buffer 1, in a volume equal to that used of bacterial cell lysis buffer 1, and subjected to end-over-end mixing for 15 minutes. The suspension was centrifuged at 8000 x g, for 10 minutes at 4 ºC, after which the supernatant was discarded. These steps were repeated once more, before resuspending the inclusion bodies in a 1-in-2 dilution of inclusion body wash buffer 1, with subsequent steps repeated. The treated inclusion body mass was then resuspended overnight, at 4 ºC via end-over-end mixing, in 5 ml of inclusion body solubilisation buffer 1, followed the next morning by clarification at 11000 x g for 15 minutes at 4 ºC. The inclusion bodies were then subjected to immobilized metal-ion affinity chromatography, to further select for SNAMA DCM protein.

2.6.3 Refolding of SNAMA DCM protein via dialysis

Refolding via dialysis was performed according to the two methodologies [(i) and (ii) below] reported in Antunes (2008) as yielding correctly refolded SNAMA DCM protein. Further attempts [(iii) below] were made at refolding the SNAMA DCM protein using experimental approaches based on literature available at REFOLD (Chow *et al.*, 2006). In all instances, approximately 5 ml of IMAC-purified SNAMA DCM protein (Section 2.6.2) was placed in SnakeSkin™ pleated dialysis tubing and subjected to one of the procedures detailed below. Dialysis was performed at 4 ºC overnight; a magnetic stirrer bar was utilized to facilitate diffusion.

(i) Dialysis without additives

The sample was subjected to a one-step, 100-fold volume of dialysis buffer 1.

(ii) Dialysis with additives

The sample was subjected to a one-step, 100-fold volume of dialysis buffer 2.

(iii) Dialysis with redox buffer

The sample was subjected, for 4 hours at a time, to 3 sequential changes of dialysis buffer 3, in which the first buffer was supplemented with 4 M urea, the second buffer with 2 M urea, while the final buffer lacked urea. The sample was then dialyzed overnight in a fourth change of dialysis buffer 3. In each buffer, the dialysate was 100 times the volume of sample.

2.6.4 Recovery of correctly-refolded SNAMA DCM protein

Dialyzed samples were clarified by centrifugation at 10000 x g for 10 minutes at 4 ^{o}C prior to being processed via glutathione affinity chromatography. The standardized manufacturer's instructions were mostly followed, with two exceptions: resins were equilibrated, and bound proteins washed, with the relevant dialysate used in the refolding attempt. The fractions generated were then analyzed via SDS-PAGE.

2.7 PREPARATION AND PURIFICATION OF ANTI-SNAMA DCM ANTIBODIES FROM CHICKEN SERUM

Anti-SNAMA DCM antibodies were sought in order to verify the existence of the two putative SNAMA proteins via immunoblotting and form the backbone of further experimental approaches in this study.

Table 2.6: Buffers and solutions used in the preparation and purification of anti-SNAMA DCM antibodies

BUFFER/SOLUTION	COMPOSITION
Protein extraction and inclusion body preparation	
Bacteria lysis buffer 2	0.1 M sodium phosphate buffer, pH 7.4; 0.15 M NaCl; 0.2 % Triton™ X-100; 100 µg/ml lysozyme

Inclusion body solubilisation buffer 2	50 mM sodium phosphate buffer, pH 7.4; 0.1 M NaCl
Inclusion body wash buffer 2	0.1 M sodium phosphate buffer, pH 7.4; 1.5 M NaCl; 0.5 M urea; 2 % Triton™ X-100
Cross-linking of control proteins to aldehyde-agarose	
Blocking buffer	0.1 M Tris-HCl, pH 8.8
Reducing agent	Sodium cyanoborohydride, 5 M solution in aqueous 1 M sodium hydroxide – see Table 2.1
Resin equilibration buffer	50 mM sodium phosphate buffer, pH 7.4; 0.1 M NaCl
Purification of anti-SNAMA DCM antibodies	
Antibody elution buffer	0.1 M glycine-HCl, pH 3.0
Antibody neutralization buffer	1 ml of 1X PBS buffer, pH 7.4; 100 µl of 1 M Na_2HPO_4, pH 8.8
Antibody purification buffer	1X PBS buffer, pH 7.4; 0.1 % Tween®20

2.7.1 Preparation of immunogen

SNAMA DCM protein was prepared as described in Sections 2.6.2 and 2.6.3. Prepared protein was concentrated to 200 µg/ml via stirred cell ultrafiltration and provided to Professor Theresa Coetzer, of the University of KwaZulu-Natal, for polyclonal antibody proliferation in adult chickens. As the author was not involved in any further procedures related to antibody production (e.g., injection of immunogen and adjuvant, bleeds, etc.), this falls outside the scope of this project and is not described.

2.7.2 Preparation of control proteins for use in antibody purification and as controls in immunoblotting.

2.7.2.1 Positive control

Pelleted SNAMA DCM cells were prepared as described in Section 2.6.1. Thawed cells were lysed in 5 ml bacteria lysis buffer 2 per 100 ml original culture volume, along with gentle sonication. Three microlitres of DNAse was added per 5ml extract, which was placed on ice for a further hour. After centrifugation at 10000

x g for 10 minutes, the supernatant was discarded and the inclusion body mass resuspended, via gentle pipetting, in 1 ml inclusion body wash buffer 2 per 5 ml original protein extract. The sample was centrifuged at 500 x g for 5 minutes before repeating the wash and low speed centrifugation steps twice more.

The previous resuspension and low speed centrifugation steps were repeated a further three times using 1ml inclusion body solubilisation buffer 2, although this was effectively a wash to remove residual compounds such as urea which would interfere with cross-linking. The inclusion body mass was re-suspended a final time, via gentle pipetting, in 1 ml inclusion body solubilisation buffer 2 and subjected to end-over-end mixing, overnight, at 4 ºC. Following clarification at 3000 x g for 10 minutes, the supernatant was transferred to a new tube and the protein concentration determined via the Bio-Rad protein assay. Sample to be cross-linked to aldehyde-agarose was utilized immediately, whereas sample destined for use as a positive control in immunoblots was made up to 30 % with glycerol and stored at -30 ºC.

2.7.2.2 Negative control

Pelleted control cells were prepared as described in Section 2.6.1. Thawed cells were lysed in 5 ml bacteria lysis buffer 2 per 100 ml original culture volume, along with gentle sonication. Three microlitres of DNAse was added per 5ml extract, which was placed on ice for a further hour. The extract was centrifuged at 3000 x g for 10 minutes and the supernatant transferred to a new tube. Protein concentration was determined via the Bio-Rad protein assay. Sample to be cross-linked to aldehyde-agarose was utilized immediately. However, for sample to be used as a negative control in immunoblots, glycerol was added to 30 % before storage at -30 ºC.

2.7.3 Cross-linking of control proteins to aldehyde-agarose

Two 30 ml chromatography columns, each containing 1 ml of aldehyde-agarose, were prepared according to the manufacturers' instructions. Cross-linking of control proteins to their respective resins was performed according to the resin

manufacturer's instructions. In total, 4.5 mg of positive control protein extract was cross-linked to the positive control resin, while 6.5 mg of negative control protein extract along with 2 mg dialyzed, glutathione affinity purified control protein was cross-linked to the negative control resin.

2.7.4 Selection of anti-SNAMA DCM antibodies from post-immune serum

The positive control column, containing a frit only below the resin, was equilibrated with antibody purification buffer and thereafter plugged at the drainage spout. The resin was resuspended, through gentle vortexing, in a 2 ml solution consisting of 1 ml post-immune serum and 1 ml antibody purification buffer. The resin was allowed to settle before gently resuspending it once more. Resuspension was repeated several times, at random, over a three hour period, after which the column was drained of the serum solution.

The column was again plugged and the resin resuspended in 2 ml antibody purification buffer, which was drained as soon as the resin had settled. This wash procedure was repeated a further four times. Thereafter, the resin was resuspended in 1 ml of antibody elution buffer which was drained, as soon as the resin had settled, into a 15 ml tube containing 1.1 ml of antibody neutralization buffer. This was immediately followed by resuspending the resin in 1 ml of antibody purification buffer, which was then also drained into the same 15 ml tube as soon as the resin had settled. The eluted antibody fraction was further processed via pre-adsorption (Section 2.7.5) whereas the positive control column was cleaned with GSH cleansing buffers 1 & 2 and stored in chromatography column storage buffer for future use.

2.7.5 Adsorption of contaminating antibodies from partially purified anti-SNAMA DCM antibody solution as well as the pre-immune serum

The negative control column was equilibrated in antibody purification buffer, after which the drainage spout was sealed. The resin was resuspended, via gentle vortexing, in either the partially purified anti-SNAMA DCM antibody solution (Section 2.7.4), or pre-immune serum diluted in an equal volume of antibody

purification buffer, and allowed to settle. Resuspension was performed at random and repeated several times over a two hour period, after which the pre-adsorbed antibody (anti-SNAMA DCM or pre-immune serum) was drained from the column. The negative control column was stripped of adsorbed antibodies through use of GSH cleansing buffers 1 & 2, after which the column was either equilibrated for a second round of pre-adsorption of the recently drained antibody solution, or else stored in chromatography column storage buffer between use. The protein concentration of the purified antibody (either SNAMA DCM or pre-immune serum) was determined via the Bio-Rad protein assay, after which glycerol was added to 30 %; a working aliquot was stored at 4 ºC with remaining aliquots stored at -30 ºC.

2.7.6 MALDI-TOF confirmation of anti-SNAMA DCM antibody specificity

Positive control protein was separated via SDS-PAGE and stained with Coomassie brilliant blue stain. The 63 kDa SNAMA DCM band, as determined via scanning densitometry, was excised from the gel and transferred to a tube containing 5 % acetic acid solution. Further preparation steps, as well as MALDI-TOF sequencing and analysis, were performed by the Plateforme Protéomique, Université de Strasbourg, Institut de Biologie Moléculaire et Cellulaire, Strasbourg, France. See Figure E1 for sequencing results.

2.8 INVESTIGATION OF THE SPATIO-TEMPORAL REGULATION OF SNAMA PROTEINS

Immunoblotting was performed in order to determine the stages in the *Drosophila* life-cycle, as well as the likely cellular compartments, in which SNAMA proteins are functional.

Table 2.7: Buffers and solutions used in investigating the spatio-temporal expression of SNAMA proteins

BUFFER/SOLUTION	**COMPOSITION**
Subcellular fractionation of 0-6 hour old embryos	
Embryo grinding buffer	15 mM HEPES-KOH, pH 7.4; 10 mM KCl; 0.2 mM EDTA; 1 mM DTT; 0.35 M sucrose; 1 µl/ml Protease inhibitor cocktail
Embryo wash solution	0.7 % NaCl; 0.05 % Triton™ X-100
Organelle solubilisation buffer	50 mM HEPES-KOH, pH 7.4; 100 mM KCl; 0.2 mM EDTA; 1 mM DTT; 0.05 % Triton™ X-100; 1 µl/ml Protease inhibitor cocktail; 1 µl/ml DNAse
***In vitro* modification of protein extracts**	
Acetylation	50 % sample to be acetylated; 25 % 25 mM $NaHCO_3$, pH 8.0; 25 % 10 µg/ml acetyl *N*-hydroxysuccinamide ester in *N*,*N*-dimethylformamide
Phosphorylation	1 ml sample to be phosphorylated; 16 µl 0.1 M ATP; 10 µl 0.5 M creatine phosphate
Mg^{2+}-precipitation	20 µl 1 M $MgCl_2$
Exposure of flies to camptothecin as a means of up-regulating *Snama*	
Camptothecin treatment	1 mM camptothecin dissolved in dimethyl sulfoxide

2.8.1 Classification and preparation of samples according to age

Protein extracts were prepared from commonly classified morphological stages of development covering the entire *Drosophila* life cycle. The number of hours of development for each morphological stage, post-deposition of an egg on an apple juice agar plate, was in accordance with Strickberger (1962). All samples were prepared by homogenizing 30 mg of organism, from a relevant stage of development, in 200 µl of 5 X SDS-PAGE sample buffer. The extract was

clarified by centrifugation at 10000 x g, after which the supernatant was transferred to a new tube and boiled at 95 ºC for 5 minutes. As SDS renders protein quantification impractical, ten microlitres of prepared sample was separated via SDS-PAGE and analyzed via immunoblotting.

2.8.2 Subcellular fractionation of 0-6 hour old embryos via differential velocity centrifugation

Zero to six hour old embryos were fractionated based on the methodology of Tie *et al.* (2001), with modifications. Briefly, embryos were dechorionated using household bleach and rinsed twice thereafter: first with embryo wash, followed by ddH_2O. The embryos were homogenized in an eppendorf tube with a plastic pestle in 0.5 ml embryo grinding buffer per 200 mg dechorionated embryos. The extract was centrifuged at 1000 x g, at 4 ºC, for 10 minutes, after which both the supernatant fraction and pelleted nuclei were extracted to separate tubes.

The nuclei were resuspended in ice-cold organelle solubilisation buffer in a volume equal to that used of embryo grinding buffer and further homogenized using a dounce homogenizer. The now solubilized nuclei represent the nuclear fraction.

The supernatant fraction was centrifuged at 1000 x g, for 5 minutes, at 4 ºC and transferred to a new tube. This procedure was repeated a further three times in order to discard floating lipid contaminants as well as to ensure complete removal of cellular debris and potentially remaining nuclei. The supernatant was centrifuged at 18000 x g for 30 minutes at 4 ºC and transferred to a new tube. This supernatant fraction represents the soluble cytosolic fraction. The pellet resulting from the 18000 x g centrifugation step was resuspended and homogenized exactly as was done with pelleted nuclei and represents the post-nuclear fraction. This fraction is expected to contain organelles such as ribosomes, endoplasmic reticuli and Golgi apparati. The protein concentration was determined for all three fractions which were then separated via SDS-PAGE and analyzed by immunoblotting.

2.8.3 *In vitro* modification of protein extracts

Whole cell protein extract was prepared from 0-6 hour embryos which were homogenized, using a dounce homogenizer, in 0.5 ml organelle solubilisation buffer per 200 mg dechorionated embryos. After clarification at 13000 x g for 5 minutes at 4 ºC, the extract was subjected to either, or both, of the treatments detailed below. Following *in vitro* modification, the extracts were separated via SDS-PAGE and analyzed via immunoblotting. Note: samples were intentionally overloaded in an attempt to discover SNAMA proteins.

2.8.3.1 Acetylation

Acetylation was performed as described by Simons *et al.* (1997), which was based on Bayer *et al.* (1996).

2.8.3.2 Phosphorylation and Mg^{2+}-precipitation

Phosphorylation and subsequent Mg^{2+}-precipitation were performed as described by Blencowe *et al.* (1995).

2.8.4 Exposure of flies to camptothecin as a means of up-regulating *Snama*

Up-regulation of *Snama* via camptothecin treatment was performed as described by Hull & Ntwasa (2010), with modifications. In order to avoid immuno-reactivity attributable to embryos, only five day old male flies were used in this experiment, with a negative control consisting of flies raised on yeast which was reconstituted in DMSO lacking camptothecin. One hundred flies from each treatment were separately homogenized, using a dounce homogenizer, in 0.5 ml of ice-cold, organelle solubilisation buffer. The extracts were clarified by centrifugation at 13000 x g for 10 minutes. The supernatant was collected and acetylated as described in Section 2.8.3.1. The samples were then subjected to SDS-PAGE and immunoblotting analysis.

2.9 INVESTIGATION OF POTENTIAL INTERACTION PARTNERS OF SNAMA PROTEINS

Anti-SNAMA DCM antibodies were employed in co-immunoprecipitation assays, using buffers based on Xu *et al.* (2009), in order to discover potential interaction partners of SNAMA proteins.

Table 2.8: Buffers and solutions used in investigating potential interaction partners of SNAMA proteins

BUFFER/SOLUTION	COMPOSITION
Preparation of protein extracts	
COIP lysis buffer	50 mM Tris-HCl, pH 8.0; 150 mM NaCl; 1 mM EDTA; 1 % NP-40; 1 µl/ml protease inhibitor cocktail; 1 µl/ml DNAse
COIP wash buffer	50 mM Tris-HCl, pH 8.0; 150 mM NaCl; 1 mM EDTA; 1 % NP-40
Extraction of potential interaction partners	
Spot destain solution	500 µl of 30 mM potassium ferricyanide; 500 µl of 0.1 M sodium thiosulphate
Spot storage solution	200 mM ammonium bicarbonate

2.9.1 Preparation of protein extracts and co-immunoprecipitation reactions

Protein samples were prepared from 0-6 hour old dechorionated embryos, mixed larval stages, and male-only flies. Samples were homogenized in 1 ml COIP lysis buffer, on ice, over a period of 2 hours. Homogenates were clarified via multiple rounds of 5 minute long centrifugation at 18000 x g and transferred to a new tube between each spin. The protein concentration was determined for all three samples after which eppendorf tubes, containing one milligram of embryo, larvae, or adult protein extract, were set up in duplicate. Fifteen micrograms of anti-SNAMA DCM antibody was added to one set of samples, while 15 µg of preadsorbed pre-immune serum was added to the other set, for use as a negative

control. A positive control reaction was also set up, consisting of 100 µg Positive control protein and 15 µg anti-SNAMA DCM antibody. Each of the seven reactions' volumes were made up to 1 ml with COIP lysis buffer and subjected to end-over-end mixing overnight at 4 ºC.

2.9.2 Preparation of columns for binding immune-reactive complexes

Due to the minute volumes used, seven home-made chromatography columns were prepared by placing a custom-fitted frit at the bottom of a 1 ml syringe. Each column was packed with 20 µl of Chicken IgY Precipitating Resin, equilibrated in COIP wash buffer, after which the spouts were sealed and the seven individual assays each added to a separate column. The columns were sealed at the top with a plunger and subjected to end-over-end mixing at 4 ºC for 4 hours.

2.9.3 Purification and analysis of immune-reactive complexes

The columns were drained of extract and the resins washed with four column volumes of COIP wash buffer, followed by a further two washes lacking detergent. Residual liquid within the resins was removed with blotting paper and the spouts were again sealed. The resins were resuspended in 30 µl 1X SDS-PAGE sample buffer, via vortexing, over a period of 30 minutes during which time the columns were also placed in a heating block set at 60 ºC. The SDS-PAGE sample buffer elution was extracted from the columns and analyzed via SDS-PAGE.

2.9.4 Extraction of potential interaction partners from gels and identification via LC-MS/MS

After analysis via scanning densitometry, bands of interest were punched out of gels using individual, cut pipette tips. Coomassie-stained samples were stored in 5 % acetic acid whereas silver-stained samples were first destained according to Gharahdaghi *et al.* (1999), before storing in spot destain solution. Samples were further processed and identified by LC-MS/MS by the Molecular and Biomedical Technology Platform, at the Council for Scientific and Industrial Research, Pretoria, South Africa.

2.10 BIOINFORMATICS

All nucleotide and peptide sequences, related to *Drosophila melanogaster*, were obtained from FlyBase (St. Pierre *et al.*, 2014). Non Fly-related sequences were obtained from Ensembl (Flicek *et al.*, 2014). Sequence alignments were performed using DNAMAN (Lynnon BioSoft, version 4.03). Mass spectrometry data analysis was performed using ProteinPilot™ software (AB SCIEX, version 4.0.8085). Protein accession numbers for all samples identified via mass spectrometry were acquired from UniProt (The UniProt Consortium, 2014).

CHAPTER THREE

3. RESULTS

The single-copy *RBBP6* gene is thought to have evolved shortly after the emergence of the eukaryotic cell (Pugh *et al.*, 2006). Generally, more evolutionarily ancient species produce only one RBBP6 protein from this gene, whereas later evolved species produce multiple proteins through alternative splicing. Two RBBP6 proteins are predicted to exist in *Drosophila melanogaster*: the 139 kDa SNAMA-PA and 55.6 kDa SNAMA-PB, although neither protein has thus far been proven experimentally. This study aims to investigate and characterize RBBP6 proteins in *Drosophila melanogaster*. The SNAMA DCM sequence, which is common to both putative SNAMA proteins, was previously cloned by Antunes (2008). In the present study, the SNAMA DCM protein was heterologously expressed in *E. coli*, with the intention of performing pulldown assays as a means of investigating potential interaction partners of SNAMA proteins. Polyclonal antibodies were also raised against the SNAMA DCM protein and purified. The antibodies were employed in immunoblots; firstly, to discover RBBP6 proteins and secondly, to investigate their spatio-temporal regulation. Lastly, the antibodies were used in co-immunoprecipitation assays as an alternate approach in investigating potential interaction partners of SNAMA proteins. Taken together, this study aims to further our understanding of SNAMA, which may in turn provide new insight into mammalian RBBP6 functionality.

3.1 PURIFICATION OF HETEROLOGOUSLY EXPRESSED PROTEINS

Heterologous protein expression was performed using *E. coli* BL21 (DE3) pLysS cells as the host and pET-41a(+) (Figure B1) as the expression vector. Upon IPTG-mediated induction of protein expression, the vector is transcribed from a start site, located at position 1166 b.p., until the terminator site located at positions 72-26 b.p. The resultant transcript is translated from an AUG corresponding to the ATG located at positions 1095-1093 b.p. on the vector. Translation is terminated

at a stop codon corresponding to positions 141-139 of the vector sequence. The resultant 35.675 kDa protein contains an N-terminal 25.65 kDa *Schistosoma japonicum* GST solubility and purification tag, followed by a hexa-histidine purification tag and both thrombin and enterokinase protein cleavage sites (Figure B2). At the nucleotide level, the sequence encoding the 35.675 kDa protein also contains an S-tag for sequencing analysis and a C-terminal multiple cloning site (MCS) for insertion of desired sequences.

The two bacterial clones utilized in this project were created by Antunes (2008). The first clone contains an unmodified pET-41a(+) vector and is referred to in this study as the "control" clone. Subsequently, induced protein expression of the control clone yields the 35.675 kDa protein described above, and is referred to in this study as the 36 kDa control protein (Figure C1). As the control protein contains a GST tag, glutathione affinity chromatography was used to purify the protein from the bacterial protein extract. Processing of the supernatant fraction led to the recovery of a single band, approximately 36 kDa in size, in the elution fraction (Figure 3.1) and is consistent with the calculated size of 35.675 kDa reported in Figure C1.

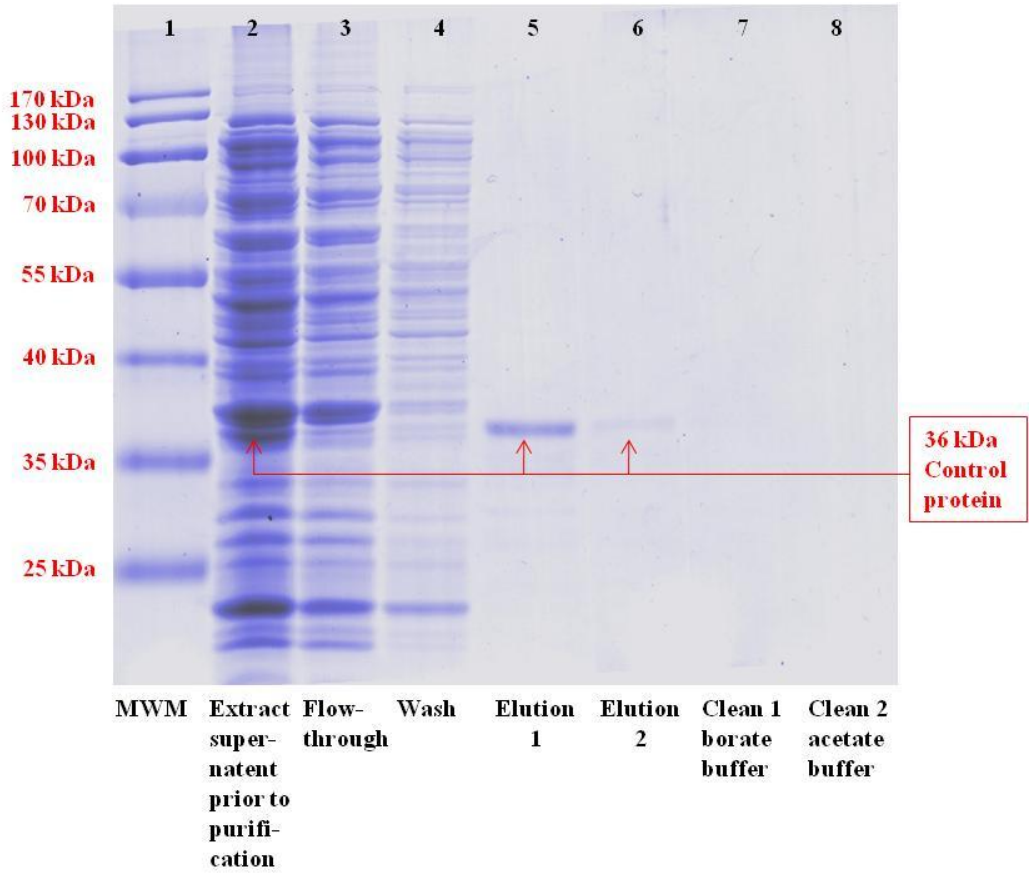

Figure 3.1: Glutathione affinity chromatography purification of the control protein. The 36 kDa control protein, as indicated by red arrows, is present in the supernatant fraction of induced control bacterial extract (lane 2) and recovered as a pure protein in the elution fraction (lane 5).

The second bacterial clone features a recombinant pET-41a(+) vector into which the SNAMA DCM sequence has been inserted between the *Bam*HI and *Hind*III restriction sites. This second clone, producing a 63 kDa protein (Figure D1), is referred to in this study as the SNAMA DCM clone / protein, respectively. The control protein and the SNAMA DCM protein are therefore identical in both sequence and features, from the amino-terminus up to the *Bam*HI restriction site; the sequences differ beyond this point, as the approximately 1.7 kDa sequence located between the *Bam*HI and *Hind*III sites has been replaced by the 29.2 kDa SNAMA DCM sequence. Furthermore, insertion of the SNAMA DCM sequence causes a frameshift, resulting in the region downstream of the *Hind*III site to change. As such, the SNAMA DCM protein codes for the sequence of pET-41b(+), as shown in Figure B1, which contains a C-terminus octa-HIS tag instead of the octa-proline tract found in the control protein.

In contrast to the control protein, and consistent with the results of Antunes (2008), the SNAMA DCM protein was recovered as an inclusion body following protein expression and extraction. Glutathione affinity chromatography could not be utilized in order to purify the SNAMA DCM protein, as the protein was not correctly folded and could therefore not interact with the GST binding site of the glutathione resin. As such, the SNAMA DCM protein had to be purified via its two HIS tags. The inclusion body pellet was first solubilized in a urea-based buffer and thereafter subjected to immobilized metal-ion affinity chromatography (IMAC) (Figure 3.2). The elution fraction contained predominantly SNAMA DCM protein but also contaminating bacterial proteins. Once correctly refolded, the SNAMA DCM protein can be purified from the contaminating proteins via glutathione affinity chromatography.

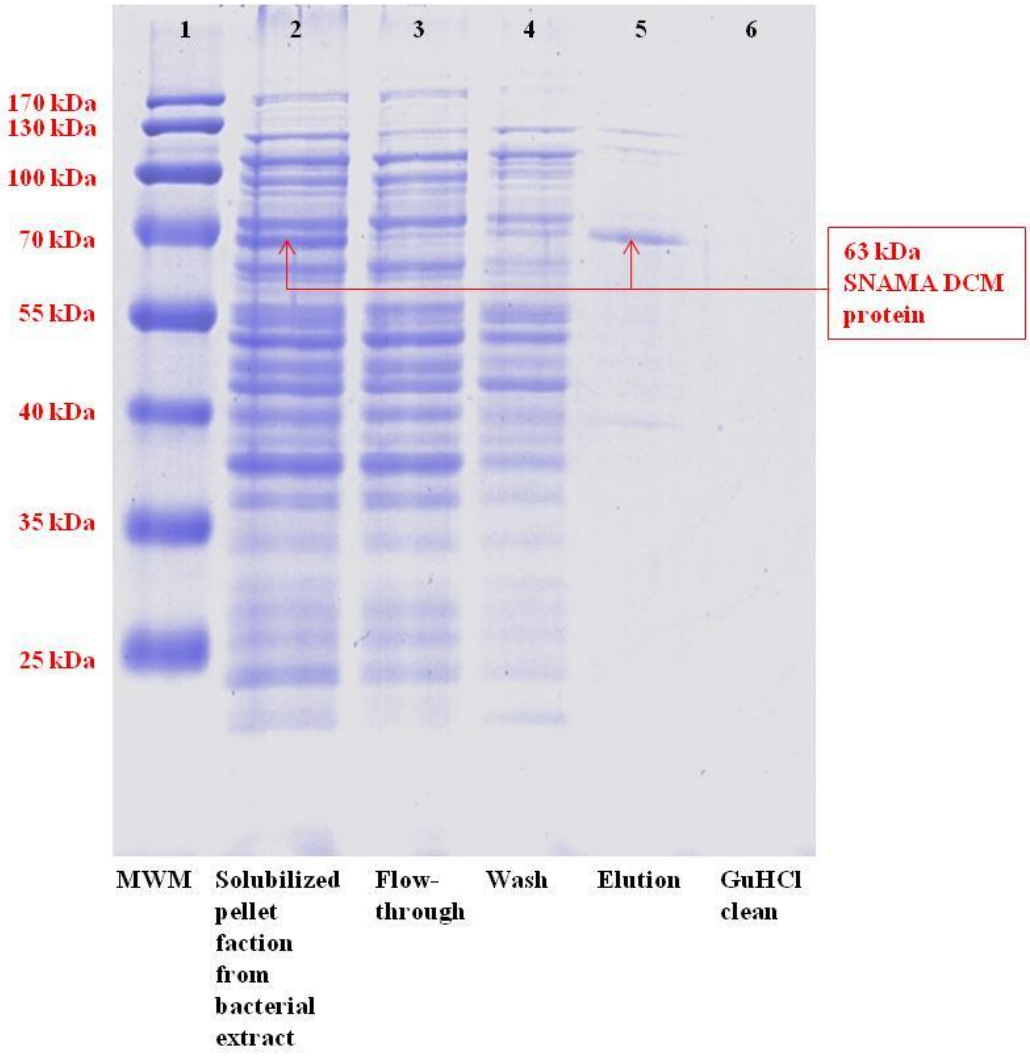

Figure 3.2: IMAC purification of SNAMA DCM protein. Co-purifying, positively-charged bacterial proteins contaminate the elution fraction, which predominantly contains SNAMA DCM protein.

3.2 DENATURED SNAMA DCM PROTEIN WAS UNABLE TO REFOLD VIA DIALYSIS

In order to perform pulldown assays, the SNAMA DCM protein must exist in its correct tertiary structure. The assay is thus dependant on the correct morphology of two components of the SNAMA DCM protein. Firstly, the N-terminal GST tag of the SNAMA DCM protein must be correctly folded, in order to anchor the protein to glutathione-agarose. Secondly, the SNAMA DCM component must

also be correctly refolded in order for it to be representative of native sites of wild-type SNAMA proteins, so as to allow credible interaction partners to bind.

IMAC-purified SNAMA DCM protein was refolded according to the two methodologies of Antunes (2008) which were shown to yield enzymatically-active SNAMA DCM protein. The two approaches differ only in the ex/inclusion of additives in the dialysis buffer, although the absence of additives severely reduces refolding efficiency (Antunes, 2008). In accordance with Antunes (2008), glutathione affinity chromatography (Figure 3.3) was used to evaluate whether the SNAMA DCM protein was correctly refolded following dialysis. The results of this study are not consistent with those of Antunes (2008). In both refolding strategies, SNAMA DCM protein was not recovered in the elution fractions; instead, it was present in the unbound and wash fractions. This indicates that the GST tag of the SNAMA DCM protein was unable to interact with the glutathione resin, as the tag was not refolded into its correct tertiary structure. This is supported by the fact that the same columns, used for recovering refolded SNAMA DCM protein, were able to purify control protein. Low efficiency of refolding was ruled out after failing to recover any refolded SNAMA DCM protein, after processing a twenty-fold excess compared to the amount processed by Antunes (2008).

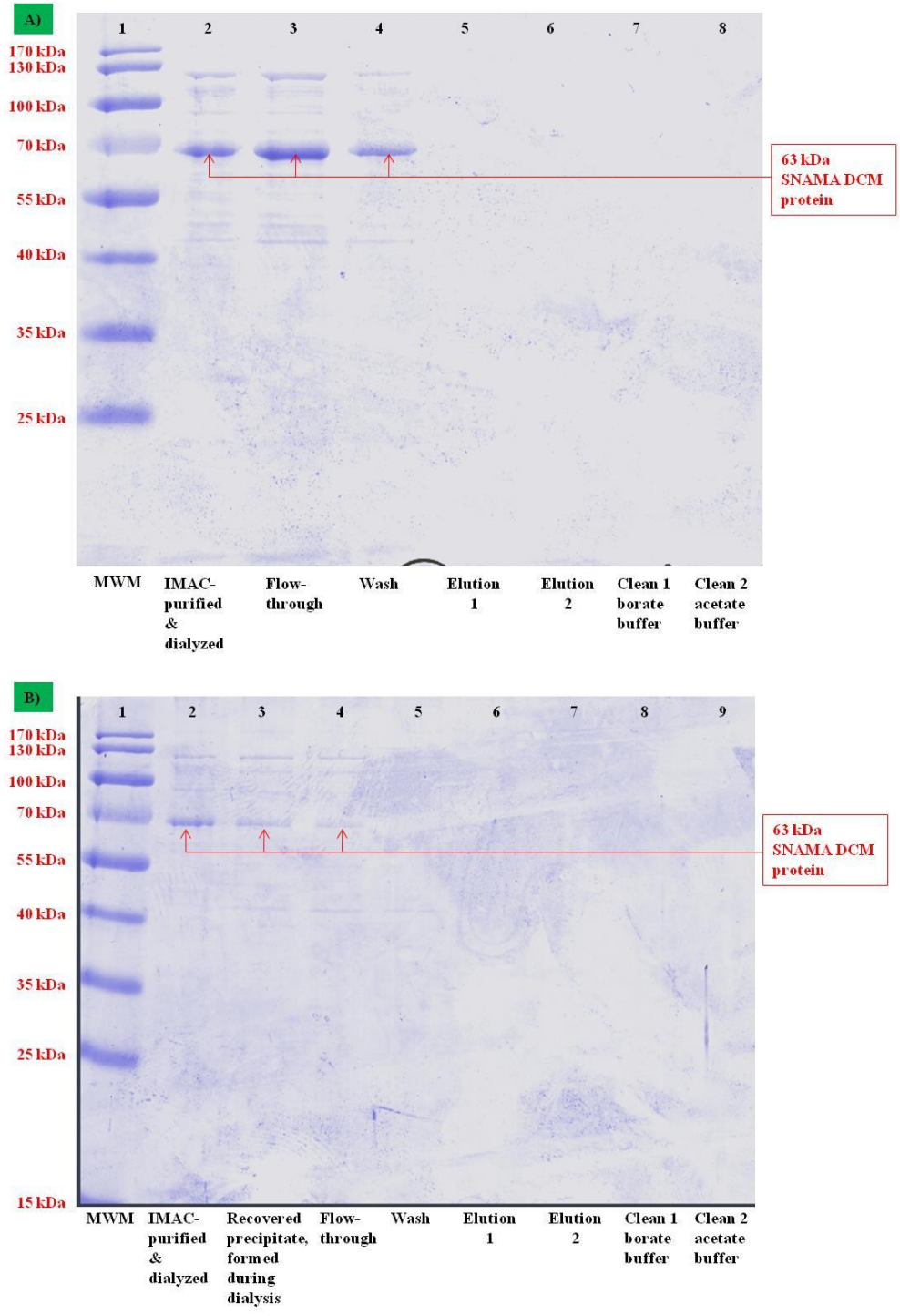

Figure 3.3: Recovery of refolded SNAMA DCM protein via glutathione affinity chromatography. Denatured SNAMA DCM protein was dialyzed according to the two methodologies of Antunes (2008) reported to yield enzymatically active SNAMA DCM protein: A) with additives, B) without additives. In both strategies, SNAMA DCM protein is absent from the elution fraction and instead recovered in the flow-through and wash fractions, thereby indicating that the SNAMA DCM protein had not been correctly refolded.

In order to correct for misfolded SNAMA DCM protein, the refolding strategies of Antunes (2008) were modified. Refolding in the absence of additives was abolished due to the formation of precipitated SNAMA DCM protein (Figure 3.3B, lane 3), which did not occur when employing additives. Further modifications were made based on features of the SNAMA DCM protein. As the active site of the SNAMA DCM protein was expected to co-ordinate two zinc ions (Mather, 2005), EDTA was omitted from the refolding buffer which instead was supplemented with zinc chloride. Furthermore, seven cysteine residues are involved in disulfide bond formation in the active site (Mather, 2005), hence a redox system was introduced into the refolding buffer. Instead of dialyzing against a single buffer change, multiple buffer changes were utilized, including, in some attempts, a crude, continuous flow system. However, these modifications were ineffective in correcting for misfolded SNAMA DCM protein (Figure 3.4) and the procedure was abandoned. Instead, potential interaction partners of SNAMA proteins were investigated through co-immunoprecipitation assays (Section 3.5).

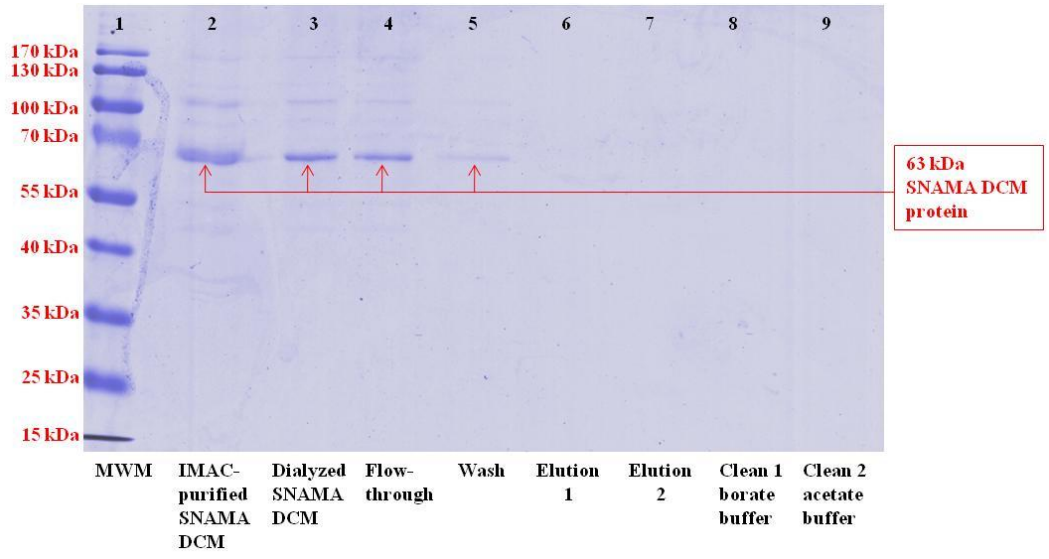

Figure 3.4: Glutathione affinity chromatography of SNAMA DCM protein following dialysis with redox-buffer. The refolding strategy of Antunes (2008) was modified in an attempt to correct for misfolded SNAMA DCM protein. Likewise, the modifications did not yield correctly refolded SNAMA DCM protein.

3.3 EVALUATION OF THE SPECIFICITY OF THE ANTI-SNAMA DCM ANTIBODIES

The author was provided with unpurified pre-and-post immune sera. In order to assess whether antibody proliferation was successful, two identical immunoblots, containing IMAC-purified SNAMA DCM protein, were probed with either unpurified serum. SNAMA DCM protein was detected exclusively by the post-immune serum (results not shown). This indicated that the pre-immune serum did not contain any antibodies which inherently cross-react with SNAMA DCM protein and further demonstrated that antibody production against SNAMA DCM protein had occurred post-administration of the immunogen. Antibodies had also been elicited against the co-purified, positively-charged bacterial proteins which contaminated the immunogen (Figure 3.5, lane 1). Furthermore, antibodies would also have been produced against antigenic sites arising from the tag components, i.e. amino acids 1-285 fused N-terminally and amino acids 546-559 fused C-terminally to the SNAMA DCM component protein.

The procedure for separating anti-SNAMA DCM antibodies from the post-immune serum was initiated by preparing protein extracts from both induced bacterial clones (Figure 3.6). As the samples were cross-linked via reductive amidation to aldehyde-agarose, the samples were prepared using buffers which lacked amine-containing compounds - thus accounting for the change in methodology compared to Section 3.1, for protein extraction and purification. The post-immune serum was applied to the column containing cross-linked positive control protein extract (Figure 3.6, lane 6), with the effect that only antibodies specific for the SNAMA DCM protein and bacterial proteins present in the immunogen, were pulled out. These selected antibodies were thereafter pre-adsorbed on the second column containing cross-linked negative control protein extract (Figure 3.6, lane 3), in which antibodies specific for the bacterial proteins as well as the N-terminal tag of the SNAMA DCM component protein were sequestered. This purification procedure yielded antibodies specific for the SNAMA DCM component, with the only contaminating antibodies being those

specific for the C-terminal tag of the SNAMA DCM protein. Given that the C-terminal tag is 5 % of the length of the SNAMA DCM, the contaminating antibodies are considered negligible. Thus, this resultant antibody sample represents the end-point of the purification procedure and is referred to as "anti-SNAMA DCM antibodies". Furthermore, the pre-immune serum was also pre-adsorbed on the negative control extract column to allow for its use as a negative control in various applications.

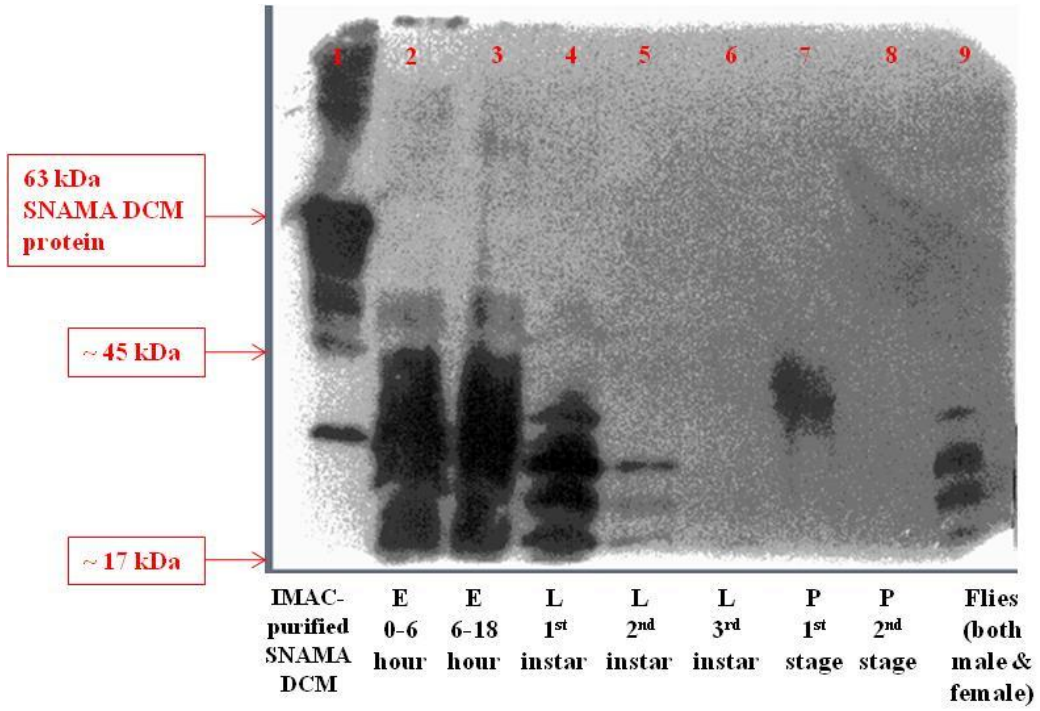

Figure 3.5: Immunoblot of developmental stages of *D. melanogaster* probed with unpurified post-immune serum. The SNAMA DCM protein is detected in lane 1 along with co-purified, positively-charged bacterial proteins. As a result, the immune-reactive bands in *D. melanogaster* samples may be attributable to cross-reactivity. Hence, the post-immune serum required purification in order to select for only anti-SNAMA DCM antibodies. E = embryo, L = larvae, P = pupae.

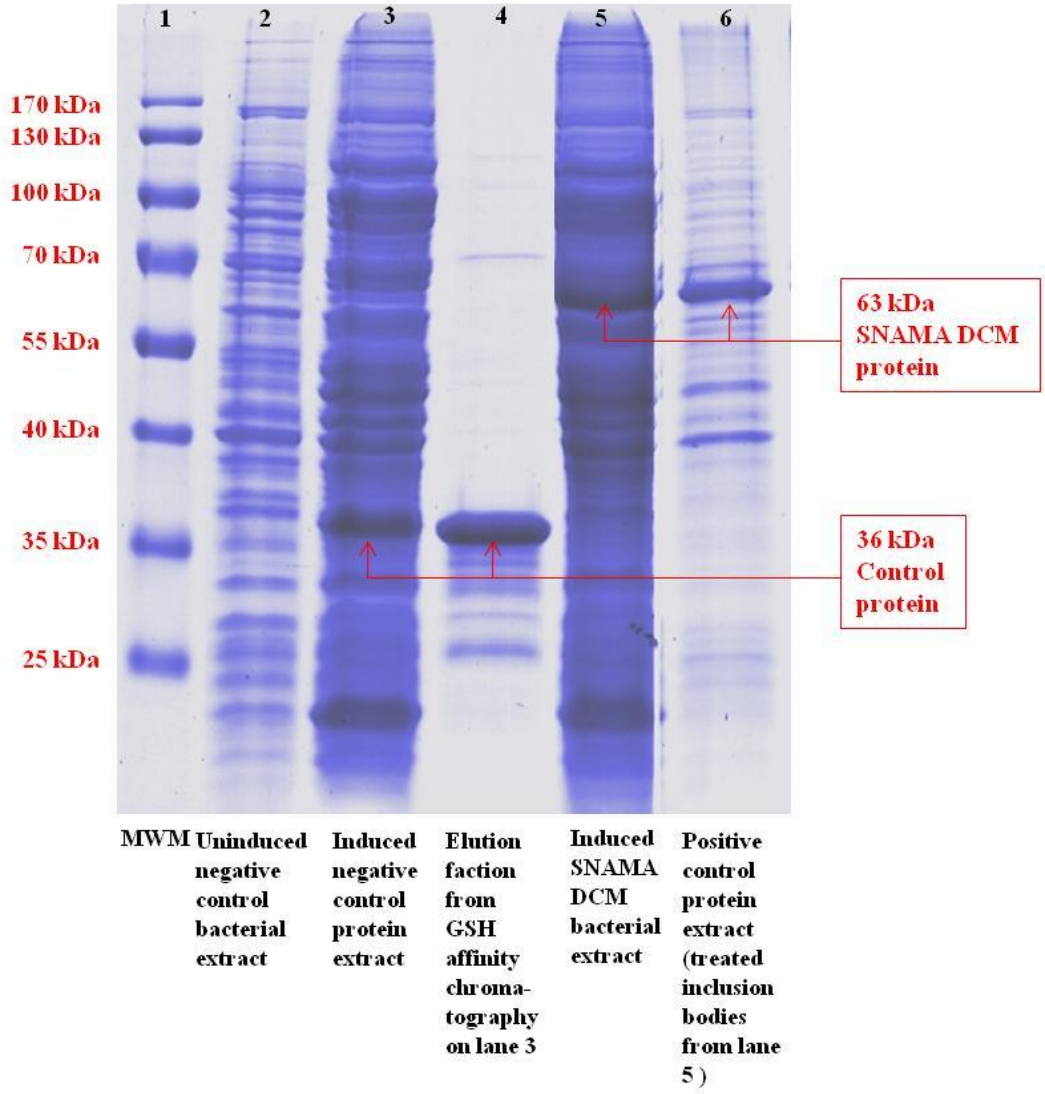

Figure 3.6: Protein extracts utilized in purifying anti-SNAMA DCM antibodies from post-immune serum. A protein extract from the uninduced control bacterial cells (lane 2) will yield an identical protein complement to uninduced SNAMA DCM protein extract. Upon IPTG induction of protein expression, the induced control (lane 3) and SNAMA DCM (lane 5) bacterial protein extracts differ only in the production of one protein: the 36 kDa control protein (lane 4) and 63 kDa SNAMA DCM protein, respectively. The first 285 amino acids of the control protein are identical to the N-terminal tag of the SNAMA DCM protein (refer to Figures C1 and D1). Thus, the control bacterial protein extract is representative of every possible antigenic site that may arise from the SNAMA DCM bacterial protein extract, with the exception of the SNAMA DCM component (amino acids 286-545) and C-terminal tag (amino acids 546-559) of the SNAMA DCM protein. This represents a strategy for purifying for only anti-SNAMA DCM antibodies.

3.4 SPATIO-TEMPORAL EXPRESSION OF SNAMA PROTEINS

3.4.1 Temporal expression of SNAMA proteins

In order to verify the existence of the two putative SNAMA proteins, as well as determine the temporal expression of any other unknown SNAMA proteins, protein extracts were prepared from the major morphological stages of *D. melanogaster* development, namely: embryo, larvae, pupae and adult. These extracts were analyzed via immunoblotting, utilizing anti-SNAMA DCM antibodies which had been effectively purified of all contaminating antibodies (Figure 3.7). The results of this experiment were mostly identical to those presented in Figure 3.5, which was probed with unpurified post-immune serum. In both Figures 3.5 and 3.7, SNAMA-PA was not detected, while the existence of SNAMA-PB – if present – is indeterminable from the immune-reactive bands which span from ~ 17 kDa to ~ 55 kDa. These immune-reactive proteins are most abundant during early stages of development and diminish with age, being completely absent in adult male flies (Figure 3.7) and virgin females (results not shown). The presence of immune-reactive proteins in the combined male and female fly extract (Figure 3.5, lane 9), is therefore attributable to the embryos of fertilised female flies. These immune-reactive proteins are therefore maternally deposited, which is in agreement with the observation that these immune-reactive proteins are most abundant during early stages of development.

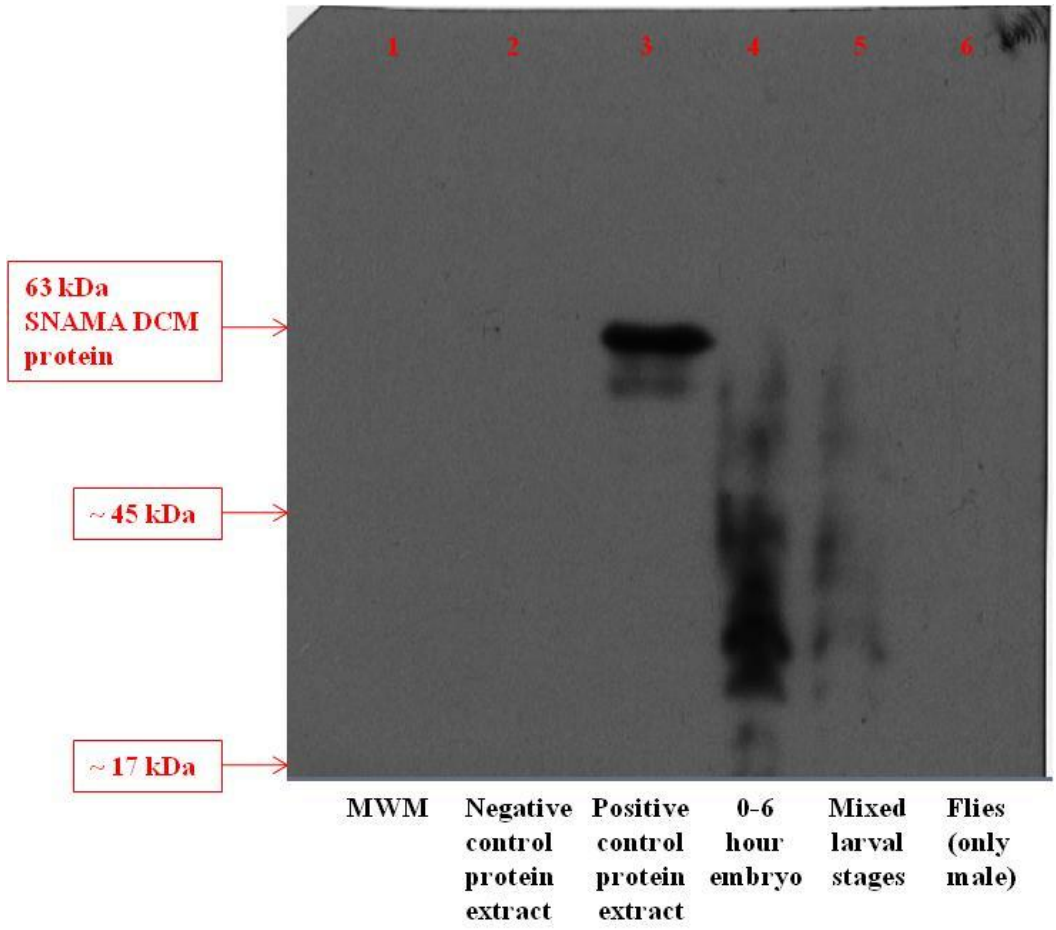

Figure 3.7: Immunoblot of developmental stages of *D. melanogaster* probed with anti-SNAMA DCM antibodies. Bacterial proteins were not detected thereby demonstrating the effectiveness of the antibody purification strategy as well as the specificity of the anti-SNAMA DCM antibodies. The results of this experiment are consistent with those of Figure 3.5: SNAMA-PA is not detected whilst the presence of SNAMA-PB is indeterminable. Furthermore, the immune-reactive proteins which span from ~ 17 kDa to ~ 55 kDa are most abundant during early stages of development and diminish with age.

3.4.2 Spatial localization of SNAMA proteins

The temporal expression studies described in Section 3.4.1 did not reveal either putative SNAMA protein, although the anti-SNAMA DCM antibodies did detect multiple immune-reactive proteins in early stages of *D. melanogaster* development. In order to further analyze these proteins, 0-6 hour old embryos were fractionated into nuclear, post-nuclear and remaining cytosolic fractions, via differential centrifugation. Two identical immunoblots, containing these protein

fractions, were probed with either pre-adsorbed pre-immune serum or anti-SNAMA DCM antibodies (Figure 3.8). This experiment verifies an earlier statement that the pre-immune serum does not detect the SNAMA DCM protein. However, the results demonstrate inherent cross-reactivity of the chicken antibodies for various *D. melanogaster* proteins, due to the detection of common bands by both pre-adsorbed pre-immune serum and anti-SNAMA DCM antibodies. Immune-reactive proteins range from ~ 17 kDa to ~ 55 kDa – consistent with previous results (Figures 3.5 and 3.7). Six unique bands (yellow arrows) are detected exclusively by the anti-SNAMA DCM antibodies. These bands all localize to the post-nuclear fraction and are of the following approximate sizes: 20, 28, 33, 35, 37 and 55 kDa. It is tempting to speculate that the ~55 kDa band may be putative 55.6 kDa SNAMA-PB.

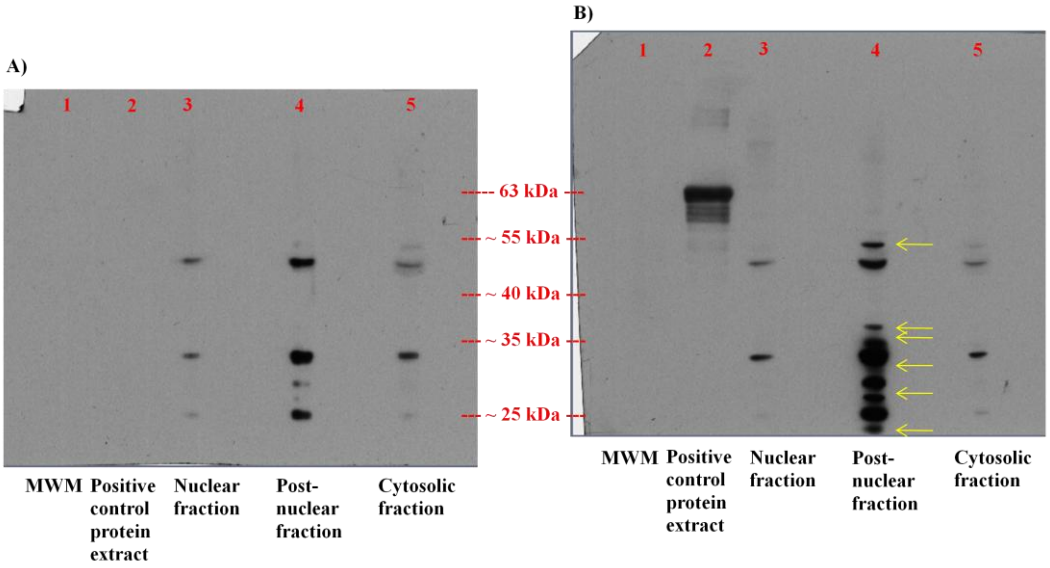

Figure 3.8: Immunoblot of subcellular fractionated, 0 – 6 hour old *D. melanogaster* embryos. Two identical blots were probed with either (A) pre-adsorbed pre-immune serum, or (B) anti-SNAMA DCM antibodies. Detection of identical bands in both blots indicates antibody cross-reactivity. Consequently, any band detected in blot A would negate its corresponding band in blot B. Six unique bands are detected in blot B - as indicated by the yellow arrows - and are all present in the post-nuclear fraction.

3.4.3 *In vitro* acetylation of protein extracts did not result in the detection of SNAMA-PA

The *M. musculus* RBBP6 protein, PACT, could only be detected via immunoblotting after it was first acetylated *in vitro* (Simons *et al.*, 1997). Acetylation neutralizes the positive charge contributed by the lysine-rich tracts, which are thought to interfere with the movement of RBBP6 proteins into the separating layer of an SDS-PAGE gel. Furthermore, PACT was also shown to co-precipitate along with SR proteins, when using a selective *in vitro* precipitation method specific for RS-domain-containing proteins (Simons *et al.*, 1997). The basis for this selective precipitation technique is phosphorylation of serine-rich tracts constituting the RS-domain, with Mg^{2+} then forming intra-and-inter ionic cross-links between phosphoserines, with resultant precipitation. Both these procedures were performed on protein extracts from 0-6 hour old embryos, which were thereafter analyzed via immunoblotting (Figure 3.9) as a means of detecting, specifically, SNAMA-PA. Cross-reactivity of the anti-SNAMA DCM antibodies for SR proteins is confirmed in both lanes 4 and 6 of Figure 3.9, owing to the detection of SR proteins in the pellet fraction of sample treated with this selective precipitation technique. A similar range of immune-reactive bands is observed in all lanes containing embryo protein extract, regardless of whether the samples were subjected to acetylation, selective precipitation, or both procedures. Immune-reactive proteins are most discernible in lane 6 of Figure 3.9, which clearly shows seven distinct bands ranging from approximately 25 kDa to 45 kDa. These bands mostly appear consistent with the bands detected in previous results (see, for example, Figure 3.5, lane 2). Despite severely overloading protein extracts, SNAMA-PA was not detected. SNAMA-PB would not be expected as it lacks the majority of the C-terminal RS-domain present in SNAMA-PA.

TREATMENT	ORDER IN WHICH PROCEDURES WERE PERFORMED							
Extraction in organelle solubilisation buffer	**N/A**	**N/A**	**N/A**	**1**	**1**	**1, 3***	**1**	**1**
Phosphorylation & Mg^{2+}-precipitation	**N/A**	**N/A**	**N/A**	**2**	**2**	**2**	**-**	**-**
Acetylation	**N/A**	**N/A**	**N/A**	**-**	**-**	**4**	**-**	**2**
LANE	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**

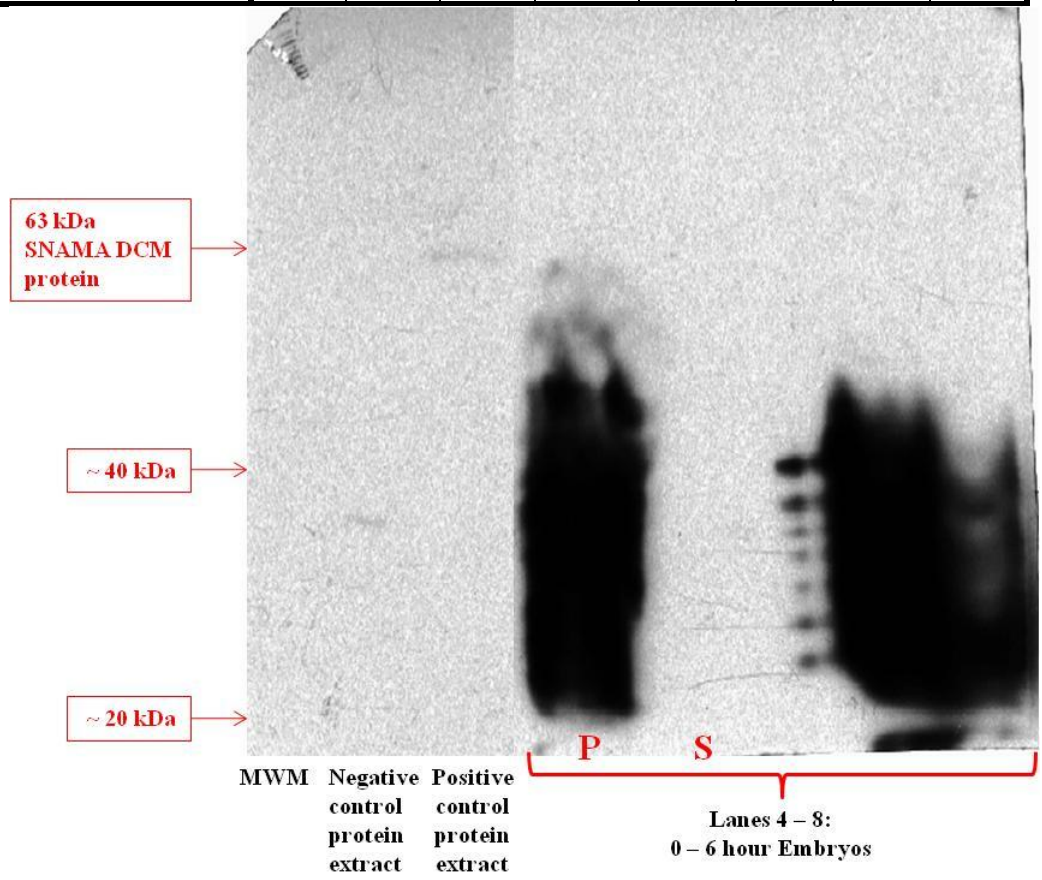

Figure 3.9: Immunoblot of acetylated protein extracts from 0-6 hour old *D. melanogaster* embryos. Samples were subjected to acetylation and/or a selective precipitation procedure, in order to detect SNAMA-PA - specifically. These procedures, however, did not result in the detection of SNAMA-PA.* The SR-precipitated pellet fraction was resusupended in HEPES buffer, prior to acetylation. P = pellet fraction; S = supernatant fraction.

3.4.4 Up-regulation of *Snama-RA* in male flies did not result in the detection of SNAMA proteins

Hull and Ntwasa (2010) demonstrated an inverse relationship between *Dmp53* and *Snama* mRNA expression levels. After subjecting flies to camptothecin-laced food for 48 hours, their results showed that *Dmp53* mRNA is up-regulated in response to DNA damage. However, following a subsequent recovery period of 24 hours, *Snama* mRNA is up-regulated while *Dmp53* mRNA is down-regulated (Hull and Ntwasa, 2010). Thus, in this experiment, protein extracts were prepared from flies 24 hours after camptothecin exposure, as yet another approach towards discovering SNAMA proteins. In order to exclude bands attributable to antibody cross-reactivity, as seen in embryo extracts, male-only protein extracts were prepared and thereafter acetylated, prior to analysis via SDS-PAGE and immunoblotting (Figure 3.10). Acetylated SNAMA DCM protein (Figure 3.10, lane 4) is slightly larger than non-acetylated SNAMA DCM protein (Figure 3.10, lane 3) due to the covalent attachment of acetyl groups, thus demonstrating the viability of the acetylation methodology. While false-positives have been effectively avoided, no immune-reactive proteins were detected. The absence of either putative SNAMA protein, as well as any other immune-reactive protein, in protein extracts prepared from male flies, is therefore consistent with previous results (Figure 3.7, lane 6). Failure to detect either SNAMA protein in this experiment could possibly suggest that camptothecin treatment was ineffective in up-regulating SNAMA proteins.

MANIPULATION	DURATION (HOURS)								
Camptothecin in DMSO	N/A	N/A	N/A	N/A	-	72	48	-	-
DMSO only	N/A	N/A	N/A	N/A	-	-	-	72	48
Recovery	N/A	N/A	N/A	N/A	-	-	24	-	24
Acetylated	-	-	-	+	+	+	+	+	+
LANE	1	2	3	4	5	6	7	8	9

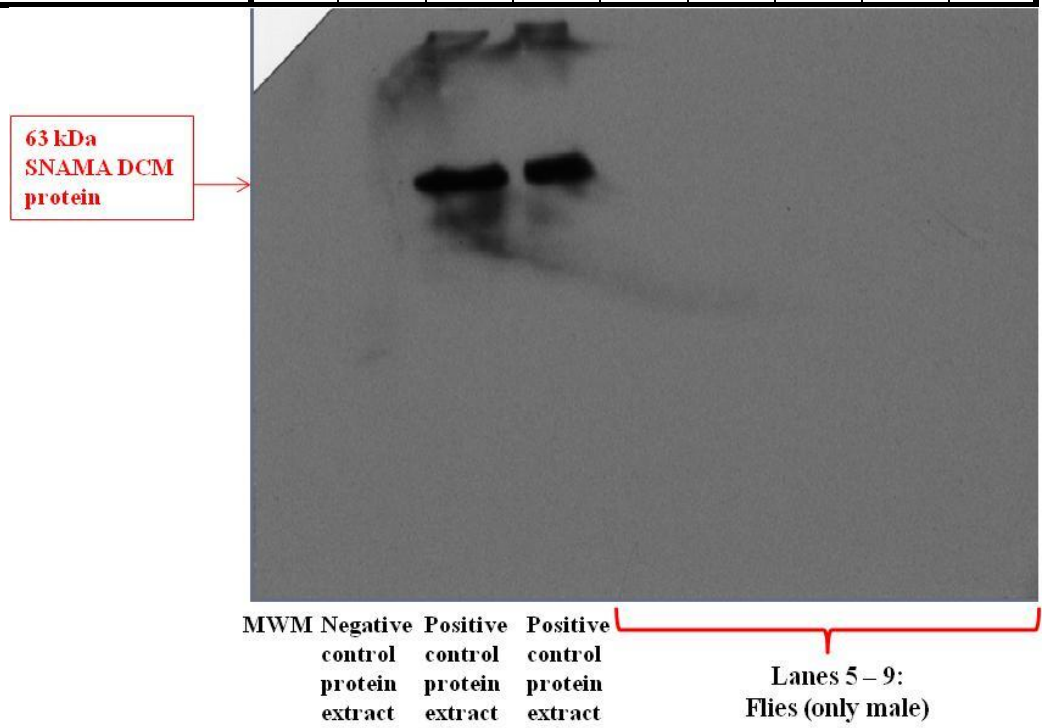

Figure 3.10: Immunoblot of protein extracts prepared from male flies exposed to camptothecin. Flies were fed camptothecin in order to indirectly up-regulate *Snama* mRNA, and potentially, subsequent SNAMA proteins. This measure also did not result in the detection of either putative SNAMA protein.

3.5 CO-IMMUNOPRECIPITATION ASSAYS AND LC-MS/MS IDENTIFICATION OF IMMUNE-COMPLEX PROTEINS

The inability to refold SNAMA DCM protein as described in Section 3.2 ruled out the option of performing pull-down assays. Instead, co-immunoprecipitation

assays were performed as a means of investigating potential interaction partners of SNAMA proteins. Protein extracts were prepared from various developmental stages and incubated in separate reactions with either anti-SNAMA DCM antibodies or, for the purpose of identifying false-positives, pre-adsorbed pre-immune serum. Captured immune-complexes were analyzed via SDS-PAGE (Figure 3.11). The 68 kDa band common to lanes 2-9 in Figure 3.11 represents the heavy chain of the chicken antibodies utilized in each assay. Recovery of SNAMA DCM protein in a positive control reaction (Figure 3.11, lane 9) demonstrates the viability of the methodology. The embryo sample (Figure 3.11, lane 3) was the only stage of development in which unique bands were observed. This result was consistent among biological replicates (Figure 3.11c). In agreement with the findings of Figure 3.8, a unique band was detected in Figure 3.11c at 55 kDa while no band was detected at 139 kDa – the anticipated sizes of the two putative SNAMA proteins.

Unique bands were manually excised from gels and sent for identification via mass spectrometry (Section 2.9.4). Negative control bands were also excised and identified in cases where uncertainty existed regarding corresponding control and test bands – e.g. two bands in lane 3 and a single band in lane 2 of Figure 3.11a are all approximately 40 kDa in size. This approach allowed for the exclusion of several false-positives (data not shown); the identities of the remaining unique proteins, across replicate experiments, are reported in Table 3.1.

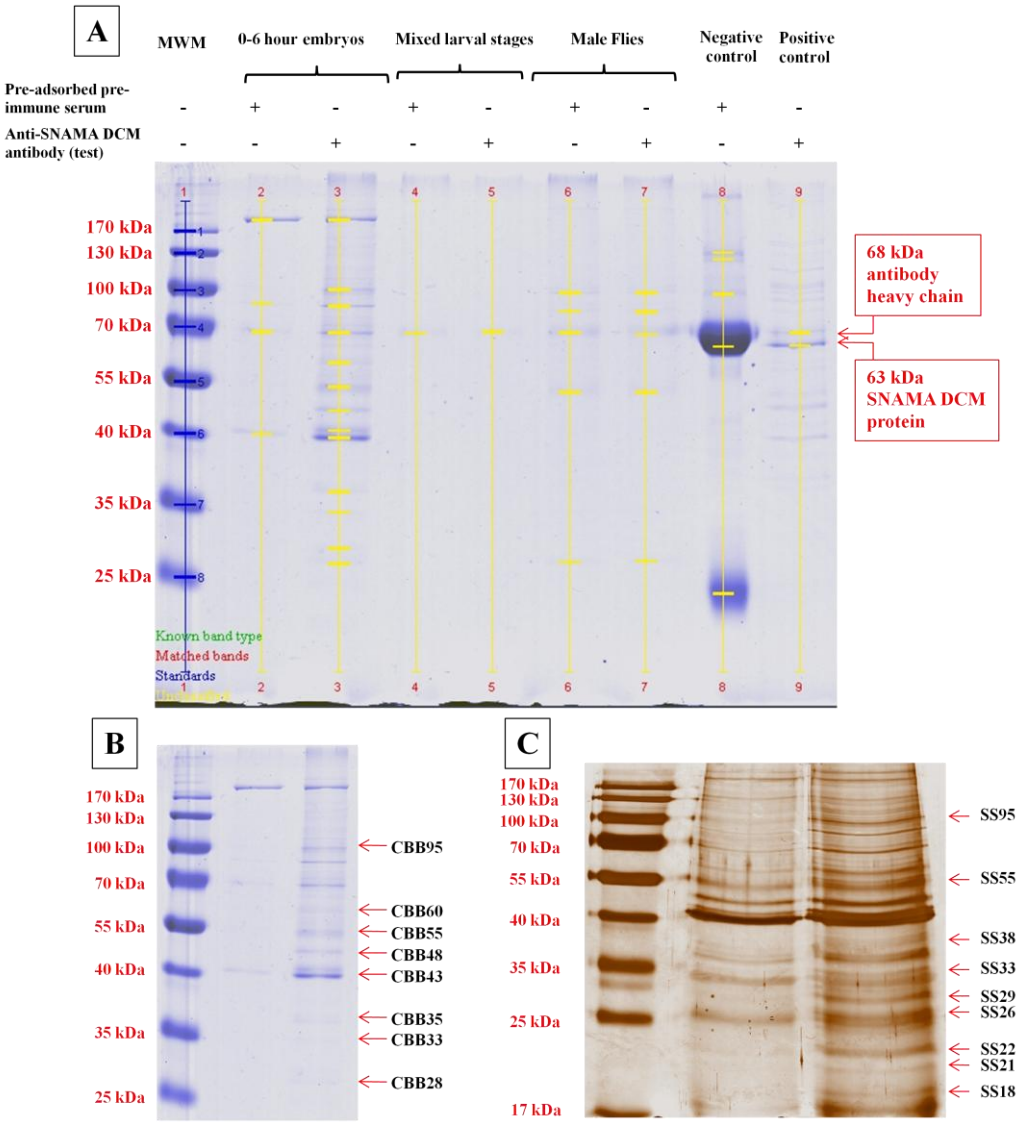

Figure 3.11: SDS-PAGE analysis of immune-complex proteins isolated from co-immunoprecipitation assays. (A) Samples were prepared from 0-6 hour embryos (lanes 2 & 3), combined larval stages (lanes 4 & 5), and male-only flies (lanes 6 & 7). For each stage of development, a negative control assay was performed (lanes 2, 4, 6). Equivalent bands between both test and control lanes, for a given stage of development, are negated. Unique bands were detected exclusively in the embryo sample (lane 3). (B & C) Bands identified by LC-MS/MS. B) Lanes 2 and 3 from the gel in (A), whereas C) is a replicate gel stained with silver-stain. The proteins identified in these two gels are presented in Table 3.1.

Table 3.1: Identities of unique proteins isolated from co-immunoprecipitation assays

SAMPLE[1]	UNIPROT ACCESSION CODE	PROTEIN NAME[2]	Mr (kDa)	PROTSCORE[3]
		TRANSCRIPTION AND TRANSCRIPTIONAL REPRESSION		
SS22a	Q94883_DROME	DNA REPLICATION-RELATED ELEMENT FACTOR (DREF)	80.727	1.24
SS95a	HP1_DROME	HETEROCHROMATIN PROTEIN 1	23.185	1.31
SS18a & SS21 & SS22b	Q24156_DROME	STONEWALL	112.913	0.87 0.94 1.08
		mRNA PROCESSING		
SS18b & CBB33	Q9W2K4_DROME	CG4266-PA	130.427	1.74 0.89
SS18c & SS26b	Q9W1K4_DROME	EGALITARIAN	112.129	1.43 1.15
SS95c	Q8SYG4_DROME	PHOSPHORYLATED ADAPTER FOR RNA EXPORT (PHAX)	53.965	1.17
		PROTEIN SYNTHESIS AND FOLDING		
CBB55a	Q9TWZ1_DROME	D-ERp60 / PROTEIN DISULFIDE ISOMERASE (PDI)	55.373	2.26
SS29	EF1A2_DROME	ELONGATION FACTOR 1-ALPHA 2	50.663	2.5
CBB43a	EF1G_DROME	ELONGATION FACTOR 1-GAMMA	48.968	6
CBB90a	EF2_DROME	ELONGATION FACTOR 2	94.459	27.74
SS26a	Q4V5H1_DROME	PEPTIDYL-PROLYL CIS-TRANS ISOMERASE	20.182	1.62
CBB43b	RL4_DROME	60S RIBOSOMAL PROTEIN L4	45.026	2.37
CBB35	Q9UAN1_DROME	60S RIBOSOMAL PROTEIN L22	30.611	2
		METABOLISM		
CBB90b	PYG_DROME	GLYCOGEN PHOSPHORYLASE	96.997	2
CBB55b	KPYK_DROME	PYRUVATE KINASE	57.44	2.75
		OTHER		
SS38	Q9VSY1_DROME	CG4022-PA	57.294	2
SS22c	Q9VQ83_DROME	CG10874-PA	35.041	1.57
SS95b	Q6IGH5_DROME	HDC06258	17.442	0.78

CBB28a	ADT_DROME	ADP, ATP CARRIER PROTEIN	34.215	5.64
CBB28b	HSP27_DROME	HEAT SHOCK PROTEIN 27	23.617	4
SS33	Q9VXL1_DROME	MIND-MELD, isoform C	150.547	0.87
SS22d	Q4ABH1_DROME	MUSCLE-SPECIFIC PROTEIN 300, isoform D	1405.148	0.97
SS95d	Q9VKQ3_DROME	RIBOSOME BIOGENESIS PROTEIN WDR12 HOMOLOG	47.222	1.47
SS95e	Q9W003_DROME	SPINOPHILIN	233.209	1.86
CBB90c	TERA_DROME	TRANSITIONAL ENDOPLASMIC RETICULUM ATPASE TER94	88.859	9.72
CBB55c	TBB1_DROME	TUBULIN BETA-1 CHAIN	50.147	16.55
CBB43c	VIT2_DROME	VITELLOGENIN-2	49.66	14.35
CBB28c	VDAC_DROME	VOLTAGE-DEPENDENT ANION-SELECTIVE CHANNEL	30.55	8.13

[1] Samples beginning with "CBB" represent bands extracted from the gel depicted in Figure 3.11b, whereas samples beginning with "SS" were isolated from the gel shown in Figure 3.11c. The number associated with each sample, represents the approximate mass (kDa) of the protein when extracted from a gel. The lower-case alphabetical letter differentiates between multiple proteins identified from a single band. [2] Proteins are listed alphabetically within broadly-defined groups, based on their biological activity reported at UniProt. [3] ProtScore is a relative score for determining and comparing the overall coverage of a protein. The score is based on the cumulative confidence scores of all peptide fragments detected for an individual protein. A score of ≥ 2 represents ≥ 99 % confidence, while a score of ≥ 1 is representative of ≥ 90 % confidence. Further details about ProtScore, including formulae and further ProtScore confidence approximations, can be found in the ProteinPilot™ software Help file.

CHAPTER FOUR

4. DISCUSSION

In this study, immunological and biochemical approaches were undertaken in order to characterize SNAMA proteins. The spatio-temporal regulation of SNAMA proteins was investigated through immunoblotting, while potential interaction partners were investigated through co-immunoprecipitation assays. These studies were performed with a view towards elucidating the roles of SNAMA and thereby also possibly those of RBBP6 proteins in general.

4.1 THE PURIFIED ANTI-SNAMA DCM ANTIBODIES DID NOT DETECT SNAMA-PA BUT LIKELY DETECTED SNAMA-PB

This study sought to identify and characterize RBBP6 proteins in *D. melanogaster*. As the two putative SNAMA proteins share the same N-terminal 493 residues, polyclonal chicken antibodies were produced against a recombinant protein comprising the first 260 residues - which includes the three structured domains referred to as the DCM (Figure 1.4). Anti-SNAMA DCM antibodies were successfully purified from the post-immune serum and were shown to be specific for the positive control protein (Figure 3.7). Further validation of antibody specificity is provided by the mass spectrometry result in Figure E1, which verifies the identity of the 63 kDa positive control protein as recombinant SNAMA DCM protein.

Utilization of the anti-SNAMA DCM antibodies in immunoblots did not result in the detection of a 139 kDa band approximating SNAMA-PA. This is not surprising, given the problems encountered by other researchers working on RBBP6 proteins. Jones *et al.* (2006) reported raising antibodies against SNAMA in rabbit, although they also failed to detect SNAMA-PA. At the time, the existence of SNAMA-PB was not known, thus no mention was made regarding detection of SNAMA-PB with these rabbit antibodies.

In addition, detection of *M. musculus* RBBP6 protein has been demonstrated (Simons *et al.*, 1997), although this was only possible after the sample was first chemically modified through acetylation, in order for it to be visualized on a SDS-PAGE gel. Likewise, in the present study, acetylation was therefore also performed. This procedure would effectively neutralize the positively-charged lysine residues found primarily at the C-terminus of SNAMA-PA in the form of two lysine-rich tracts (Table 1.1), which was suspected by Simons *et al.*, (1997) to be responsible for the difficulties in visualizing *M. musculus* RBBP6. The acetylation procedure also modifies all primary amine-containing residues and hydroxyl-containing residues such as serine, threonine and tyrosine (Abello *et al.*, 2007), which would drastically decrease the pI of SNAMA-PA. Although this procedure was demonstrated to have worked correctly due to the acetylated positive control protein having a slightly larger size than the non-acetylated positive control (Figure 3.9, lanes 3 & 4), SNAMA-PA was still undetectable. Furthermore, the results of this study appear to echo the results of a study on *H. sapiens* RBBP6, in which anti-DWNN antibodies could not detect native RBBP6 proteins in extracts from HEK293 cells (Chibi *et al.*, 2008).

As the difficulties experienced in detecting RBBP6 proteins are not confined to one organism or research group, it is likely that the domains and subsequent properties of RBBP6 proteins may interfere with the ability to detect these proteins. For instance, the RING domain present in RBBP6 proteins is known to be highly susceptible to aggregation (Borden, 2000), as is the RS-domain. Thus, considering the positive control protein lacks the RS-domain and the associated physicochemical implications such as increased susceptibility to aggregation and an extremely high isoelectric point, this could account for why it could be detected but not native SNAMA proteins.

In the event that SNAMA-PA could not be detected due to low expression levels, two techniques were performed as a means of facilitating SNAMA-PA detection. Indirect up-regulation of *Snama-RA* through camptothecin treatment has been

demonstrated by Hull & Ntwasa (2010), although it does not necessarily mean that SNAMA-PA would, in turn, be up-regulated. Regardless, this approach did not result in the detection of SNAMA-PA (Figure 3.10). As the second biochemical approach employed in this study, protein extracts were subjected to phosphorylation followed by Mg^{2+} precipitation. This technique has been shown to select for RS-domain-containing proteins (Blencowe, 1995), which therefore includes SNAMA-PA. However, this approach did not result in the detection of SNAMA-PA either.

SNAMA-PB lacks the majority of the RS-domain, thus its isoelectric point is far lower than that of SNAMA-PA (Table 1.1). A band of approximately 55 kDa is detected in Figures 3.8 and 3.11, which may be SNAMA-PB. Thus, this 55 kDa band is detected via both immunoblotting and co-immunoprecipitation assays, which is suggestive of a direct interaction with the antibody. In support of this, the 55 kDa band was not detected in the pellet fraction of Figure 3.9, which selected for RS-domain-containing proteins. Unfortunately, however, SNAMA-PB was not one of the proteins identified from the 55 kDa band sequenced via mass spectrometry (Table 3.1), although this does not discount the possibility of SNAMA-PB being identified in future, possibly even through a replicate experiment. This protein localizes to the post-nuclear fraction (Figure 3.8, lane 4), as do all unique bands identified in Figure 3.8. While this would be unusual for SNAMA-PA, due to the multiple bipartite NLS as well as nucleus-targeting sequences within the RS-domain, as SNAMA-PB lacks these regions, it seems reasonable that it could localise to other organelles, typically found in the post-nuclear fraction. Examples include mitochondria, ribosomes and Golgi apparati – the localization of SNAMA-PB to the ribosomes or Golgi apparati is consistent with a role in protein folding and quality control, as suggested by Kappo *et al.* (2012) for RBBP6 proteins. This banding pattern could be reminiscent of DWNNylation, although it would have shown in the mass spectrometry results (Table 3.1) as a DWNN domain present on all of the proteins. Thus, the MS data has demonstrated that these bands are not attributable to DWNNylation. This potential SNAMA-PB is expressed at high levels in the early stages of

development and is absent in adults (Figures 3.5, 3.8 & 3.11). Furthermore, the detection of SNAMA-PB in adult female samples is most likely due to the presence of embryos in fertilised female flies. This agrees with Northern blot analysis conducted by Mather (2005), which also showed higher levels of SNAMA transcripts in the early stages of development. Moreover, this is also in agreement with the high-throughput data of Daines *et al.* (2011) (Table 4.1), which shows that *Snama* is most actively transcribed during the early stages of development, and decreases thereafter. *Snama* is thereafter only up-regulated in female flies.

Table 4.1: Relative transcription levels for *Snama* throughout the *Drosophila melanogaster* life-cycle

STAGE	**EMBRYO 2-4 hours**	**EMBRYO 14-16 hours**	**LARVAE (average of 3 instar stages)**	**PUPAE (stage 1)**	**PUPAE (stage 2)**	**FLY (male) 72 hours**	**FLY (female) 72 hours**
RPKM	**31**	**9**	**4**	**6**	**5**	**8**	**25**

The data in this table is derived from the results of Daines *et al.* (2011). RPKM: reads per kilobase per million reads – refer to Daines *et al.* (2011) for a detailed explanation of RPKM values. RPKM values have been rounded off to the nearest whole number and indicate that *Snama* is most abundantly transcribed during early stages of embryonic development and in female flies.

4.2 ANTI-SNAMA DCM ANTIBODIES ALSO CROSS-REACT WITH RS-DOMAIN-CONTAINING PROTEINS

The original technique of Mg^{2+} precipitation (Zahler *et al.*, 1992) resulted in the detection of a conserved set of proteins, with sizes of 20, 30, 40, 55 and 75 kDa, collectively referred to as SR proteins. This Mg^{2+}-dependent precipitation procedure is thought to occur through Mg^{2+} mediating ionic cross-linking between phosphoserines on adjacent RS-domain containing proteins, with resultant protein precipitation (Zahler *et al.*, 1992). While modifications to the original procedure by others resulted in additional RS-domain-containing proteins being discovered, Simons *et al.* (1997) have demonstrated that the *M. musculus* RBBP6 protein,

PACT, can be selectively purified via this approach. Unfortunately, this technique did not result in the detection of SNAMA-PA.

This procedure was demonstrated to have worked correctly, as well as demonstrate cross-reactivity of the anti-SNAMA DCM antibodies for RS-domain-containing proteins, through the immunological detection of Mg^{2+}-precipitated proteins in the pellet fraction of Figure 3.9. In addition, some of these proteins had sizes equivalent to those shown by Blencowe *et al.* (1995), thus further supporting unintended cross-reactivity of anti-SNAMA DCM antibodies for RS-domain-containing proteins. Furthermore, the identification of one of these immune-reactive proteins as an SR protein was confirmed by the identification of a *Drosophila* SR protein, CG4266 (Shepard & Hertel, 2009) in Table 3.1. This cross-reactivity could be due to an epitope having formed by either the SR or RS dipeptide in the SNAMA DCM sequence (Figure 1.4) located at residues 74-75 and 185-186, respectively, which could resemble a phospho-epitope present on CG4266. This would not be unusual as anti-PACT antibodies raised against an epitope generated on the RS-domain cross-reacted with other RS-domain-containing proteins (Simons *et al.*, 1997).

The presence of CG4266 could also be as a result of a genuine interaction between either of the two SNAMA proteins and CG4266 - either directly, or indirectly through complex association. Like all SR proteins, CG4266 is predicted to function in mRNA processing. Despite this being a function common to all SR proteins, individual family members do appear to have distinct functions (Zhong, 2009). Some of these functions include the facilitation of RNA transport (Sapra *et al.*, 2009). Indeed the anti-SNAMA DCM antibodies also interacted with PHOSPHORYLATED ADAPTER FOR RNA EXPORT (PHAX). Furthermore, as is shown in Table 1.2, Scott *et al.* (2003) demonstrate that the SR-rich splicing factor kinase (SRPK1) phosphorylates *M. musculus* RBBP6. Thus SNAMA proteins can be expected to interact with SR proteins, either directly, or as part of a protein complex, as both RBBP6 proteins and SR proteins are known to localize to nuclear speckles.

4.3 SNAMA MAY INTERACT WITH A WIDE RANGE OF PROTEINS – MANY OF WHICH FUNCTION IN TRANSCRIPTION AND mRNA PROCESSING

Our group has recently attempted to identify potential protein transcription factors and regulatory proteins that bind to the promoter of *Snama* (Mokgohloa, L., unpublished). One of the proteins identified was the the transcription regulatory factor DREF (DNA replication-related element-binding factor). This transcription factor was also identified in this study (Table 3.1). DREF is known to control the expression of proteins playing a role in cell proliferation through their function as DNA replication, transcription and cell cycle regulators, and has appropriately been labelled the "master key for cell proliferation" (Matsukage *et al.*, 2008). Interaction between either of the two SNAMA proteins with DREF fits in well with the results of Jones *et al.* (2006), who demonstrate the importance of SNAMA in cell-cycle control and cell proliferation. The fact that DREF was captured in this study, as well as another study from our group involving a completely different approach, strongly suggests a direct interaction between DREF and SNAMA.

The proteins listed in Table 3.1 have diverse roles within a cell. A fair number of proteins could easily be surmised to accommodate roles played by RBBP6 proteins. For instance, numerous proteins involved in transcription and mRNA processing are present. These are known functions of RBBP6 proteins. Furthermore, it could be speculated that the presence of Vitellogenin-2 in Table 3.1 may also imply that SNAMA proteins are involved in nutrient synthesis, as RBPL1 was shown to directly affect Vitellogenin-2 protein levels (Huang *et al.*, 2013). As the two SNAMA proteins have vastly different isoelectric points as well as C-terminal domains, the two SNAMA proteins are likely to localize to different cellular locations and perform different cellular roles. For instance, the RS-domain of SNAMA-PA may result in it localizing to the nucleus where it could perform transcriptional activities or mRNA processing activities, whereas SNAMA-PB, which lacks the RS-domain and nuclear localization signals, may

remain within the cytosol or on organelle membranes. Table 3.1 contains numerous proteins involved in protein folding, such as HSP27 and PDI. Considering that proteasomes localize to the outer surface of endoplasmic reticuli (Wojcik & DeMartino, 2003) and ER are expected in the post-nuclear fraction of Figure 3.8, the 55 kDa band could be SNAMA-PB involved in a protein quality control capacity, as suggested by Kappo *et al.* (2012). No evidence was found for DWNNylation, as the mass spectrometry analysis would have identified a DWNN domain on at least some of the proteins isolated, which was not the case. Thus further investigation is needed to evaluate the credibility of each potential interaction partner.

4.4 CONCLUSIONS

- The polyclonal anti-SNAMA DCM antibodies produced in this study are specific for the SNAMA DCM protein but also cross-react with other RS-domain-containing proteins. This cross-reactivity is most likely due to a common epitope formed by an RS or SR dipeptide located in the DCM.

- The putative SNAMA-PA protein was not detected in this study, despite attempts at up-regulating SNAMA-PA levels and biochemical approaches aimed at enriching extracts for SNAMA-PA through selective precipitation. It is thought that the protein's C-terminal structure may interfere with the ability to visualize it via SDS-PAGE, which could not be overcome through acetylation as is the case for *M. musculus* RBBP6.

- It is likely that the putative 55.6 kDa SNAMA-PB protein has been detected via the immunological approach undertaken in this study. The spatio-temporal regulation of an immune-reactive 55 kDa protein is consistent with what is expected for SNAMA-PB as are the physico-chemical properties of this protein as determined by biochemical approaches.

- Potential interaction partners of SNAMA proteins have been isolated, although these await further confirmation through experimental validation. A strong candidate is DREF, which was isolated in this study as well as by our group in an independent study. DREF has a role consistent with known functions of *Snama*.

4.5 FUTURE WORK

In order to verify the existence of both SNAMA proteins, the co-immunoprecipitation assays of this study should be repeated. Instead of electrophoresing the eluted immune complexes via SDS-PAGE, the elutions should be sequenced via LC-MS/MS, directly. Given the capabilities of software used for identifying protein fragments, a search can be optimised for a protein of interest (Stoychev, S., personal communication) – in this case, the search should be optimised for SNAMA proteins. This approach would confirm the existence of any SNAMA proteins and would also clarify whether the absence of SNAMA-PA is attributable to its surmised inability to migrate on standard Laemmli gels.

Finally, given that DREF was isolated in both this study and in an independent study by our group, as well as the credible reasons given in Section 4.3, for it being a potential interaction partner for SNAMA proteins, the possibility of these two proteins being interaction partners should be investigated.

CHAPTER FIVE

5. APPENDICES

APPENDIX A

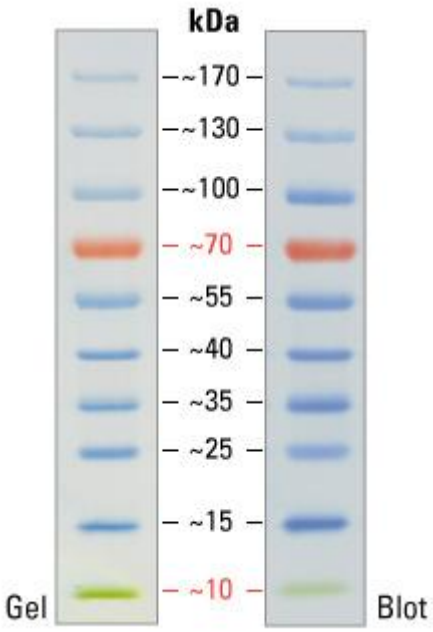

Figure A1: Fermentas PageRuler^{TM} Prestained Protein Ladder.

APPENDIX B

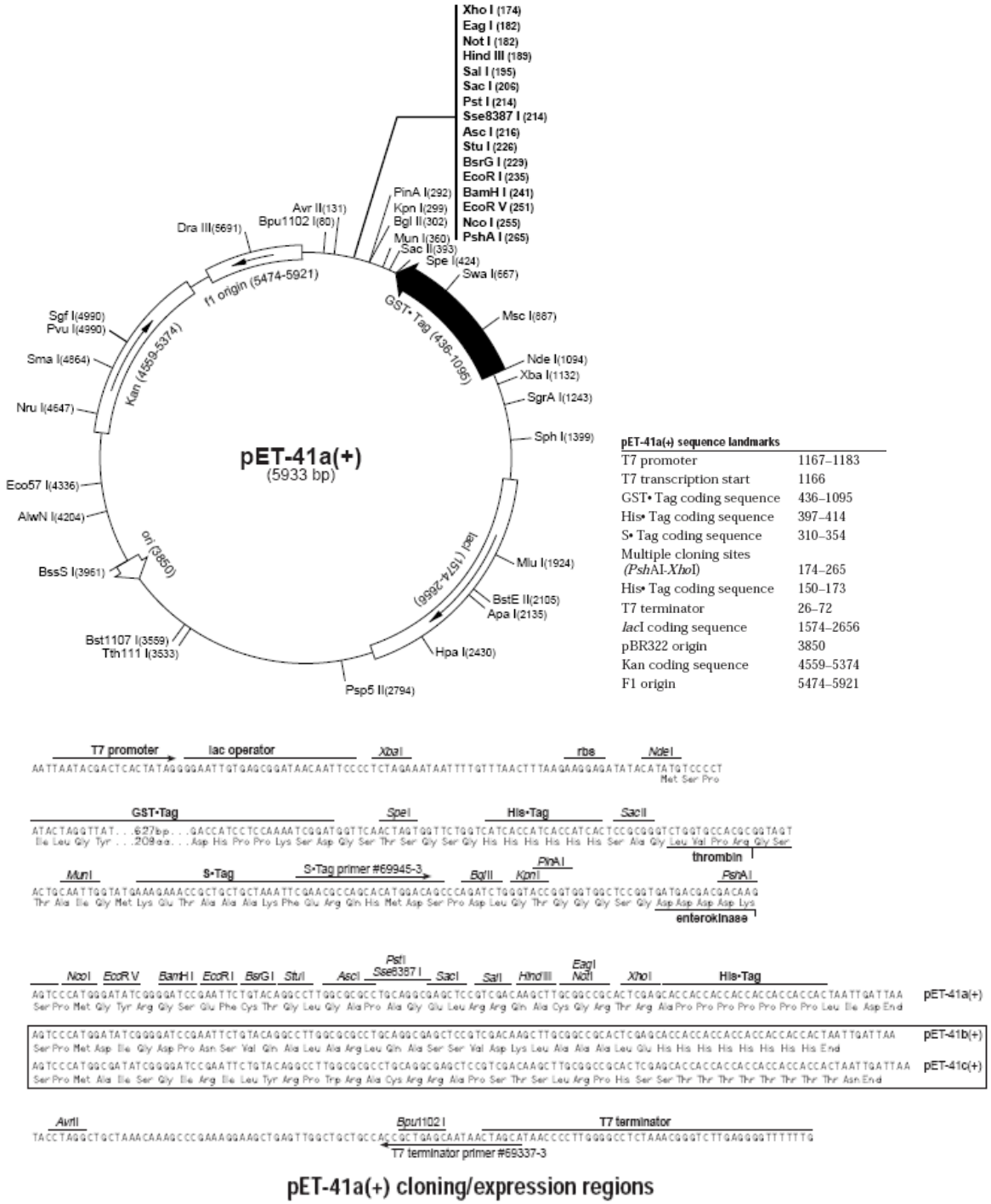

Figure B1: Novagen pET-41a(+) expression vector. The vector is transcribed from the transcriptional start site located at position 1166 b.p. until the terminator region located at positions 72-26 b.p.

ATGTCCCCTATACTAGGTTATTGGAAAATTAAGGGCCTTGTGCAACCCACTCGACT
TCTTTTGGAATATCTTGAAGAAAAATATGAAGAGCATTTGTATGAGCGCGATGAAG
GTGATAAATGGCGAAACAAAAAGTTTGAATTGGGTTTGGAGTTTCCCAATCTTCCT
TATTATATTGATGGTGATGTTAAATTAACACAGTCTATGGCCATCATACGTTATAT
AGCTGACAAGCACAACATGTTGGGTGGTTGTCCAAAAGAGCGTGCAGAGATTTCAA
TGCTTGAAGGAGCGGTTTTGGATATTAGATACGGTGTTTCGAGAATTGCATATAGT
AAAGACTTTGAAACTCTCAAAGTTGATTTTCTTAGCAAGCTACCTGAAATGCTGAA
AATGTTCGAAGATCGTTTATGTCATAAAACATATTTAAATGGTGATCATGTAACCC
ATCCTGACTTCATGTTGTATGACGCTCTTGATGTTGTTTTATACATGGACCCAATG
TGCCTGGATGCGTTCCCAAAATTAGTTTGTTTTAAAAAACGTATTGAAGCTATCCC
ACAAATTGATAAGTACTTGAAATCCAGCAAGTATATAGCATGGCCTTTGCAGGGCT
GGCAAGCCACGTTTGGTGGTGGCGACCATCCTCCAAAATCGGATGGTTCAACTAGT
GGTTCTGGTCATCACCATCACCATCACTCCGCGGGTCTGGTGCCACGCGGTAGTAC
TGCAATTGGTATGAAAGAAACCGCTGCTGCTAAATTCGAACGCCAGCACATGGACA
GCCCAGATCTGGGTACCGGTGGTGGCTCCGGTGATGACGACGACAAGAGTCCCATG
GGATATCGGGGATCCGAATTCTGTACAGGCCTTGGCGCGCCTGCAGGCGAGCTCCG
TCGACAAGCTTGCGGCCGCACTCGAGCACCACCACCACCACCACCACCACTAATTG
AT

KEY:

ATG...GAT (position 1095 – 436 of vector)	= GST tag
CATCACCATCACCATCAC (414 - 397)	= Hexa-HIS tag
CTGGTGCCACGCGGTAGT (387 - 370)	= Thrombin cleavage site
AAA...AGC (354 - 310)	= S-tag
GATGACGACGACAAG (274 - 260)	= Enterokinase cleavage site
Underlined sequences (GAC...GAG) (265 - 174)	= Multiple cloning site
G'GATCC	= *Bam*HI
A'AGCTT	= *Hind*III

Figure B2: Tags and cloning features of the pET-41a(+) vector over the region 1095-142 basepairs. While this sequence forms only part of the open reading frame of the pET-41a(+) vector (refer to Figure B1), following induction of protein expression, only the resultant transcript of the above sequence is translated. The ATG at the beginning of the GST-tag corresponds to the AUG from which translation is initiated, which terminates immediately after the GAT at the end of this sequence.

APPENDIX C

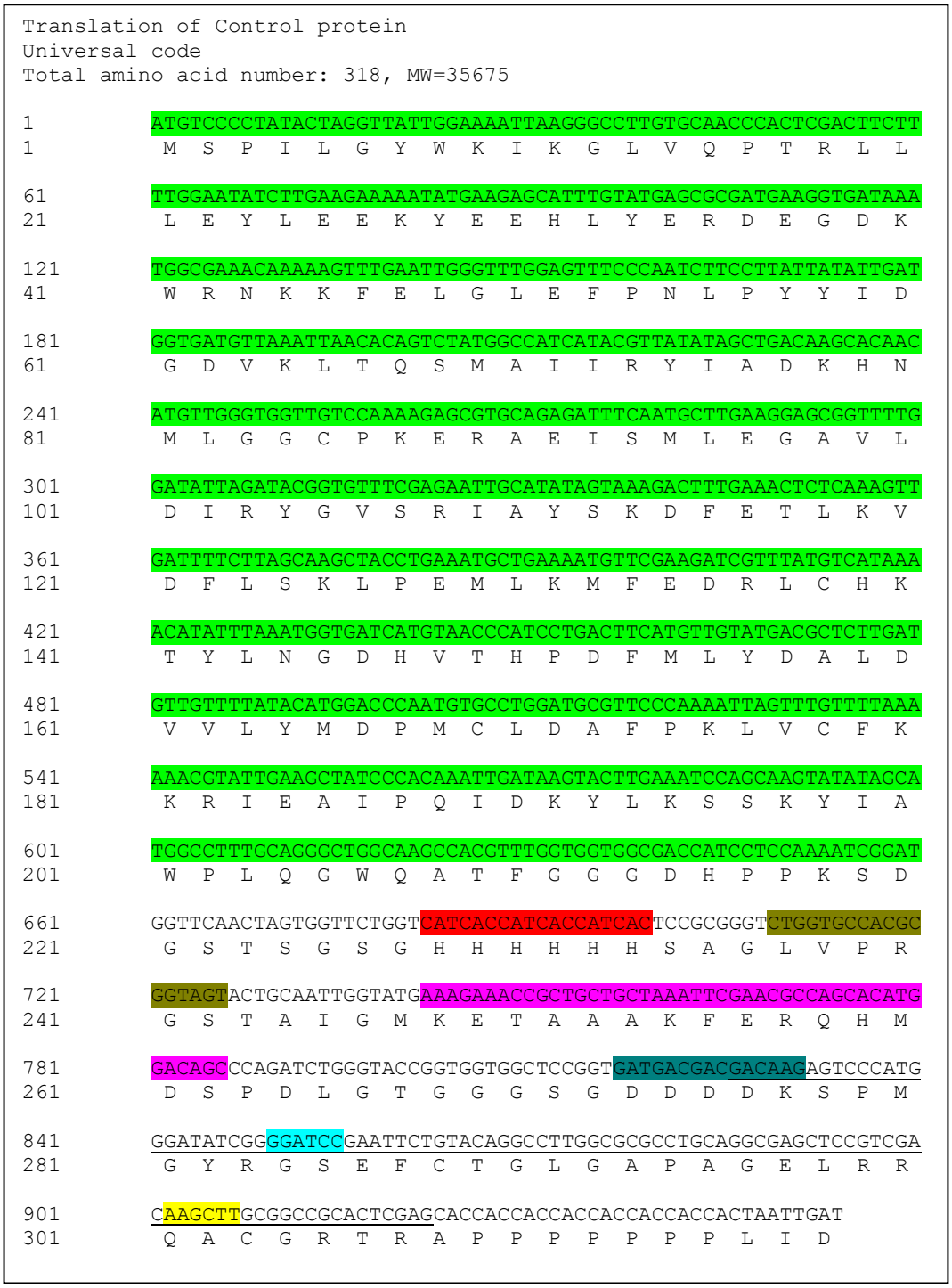

Figure C1: Amino acid sequence of the 36 kDa control protein. The colour-coded nucleotide sequence (refer to the key in Figure B2) is shown in the line above the amino acid sequence.

APPENDIX D

```
Translation of SNAMA DCM protein
Universal code
Total amino acid number: 559 AA, MW=63120

1    ATGTCCCCTATACTAGGTTATTGGAAAATTAAGGGCCTTGTGCAACCCACTCGACTTCTT
1     M  S  P  I  L  G  Y  W  K  I  K  G  L  V  Q  P  T  R  L  L

61   TTGGAATATCTTGAAGAAAAATATGAAGAGCATTTGTATGAGCGCGATGAAGGTGATAAA
21    L  E  Y  L  E  E  K  Y  E  E  H  L  Y  E  R  D  E  G  D  K

121  TGGCGAAACAAAAAGTTTGAATTGGGTTTGGAGTTTCCCAATCTTCCTTATTATATTGAT
41    W  R  N  K  K  F  E  L  G  L  E  F  P  N  L  P  Y  Y  I  D

181  GGTGATGTTAAATTAACACAGTCTATGGCCATCATACGTTATATAGCTGACAAGCACAAC
61    G  D  V  K  L  T  Q  S  M  A  I  I  R  Y  I  A  D  K  H  N

241  ATGTTGGGTGGTTGTCCAAAAGAGCGTGCAGAGATTTCAATGCTTGAAGGAGCGGTTTTG
81    M  L  G  G  C  P  K  E  R  A  E  I  S  M  L  E  G  A  V  L

301  GATATTAGATACGGTGTTTCGAGAATTGCATATAGTAAAGACTTTGAAACTCTCAAAGTT
101   D  I  R  Y  G  V  S  R  I  A  Y  S  K  D  F  E  T  L  K  V

361  GATTTTCTTAGCAAGCTACCTGAAATGCTGAAAATGTTCGAAGATCGTTTATGTCATAAA
121   D  F  L  S  K  L  P  E  M  L  K  M  F  E  D  R  L  C  H  K

421  ACATATTTAAATGGTGATCATGTAACCCATCCTGACTTCATGTTGTATGACGCTCTTGAT
141   T  Y  L  N  G  D  H  V  T  H  P  D  F  M  L  Y  D  A  L  D

481  GTTGTTTTATACATGGACCCAATGTGCCTGGATGCGTTCCCAAAATTAGTTTGTTTTAAA
161   V  V  L  Y  M  D  P  M  C  L  D  A  F  P  K  L  V  C  F  K

541  AAACGTATTGAAGCTATCCCACAAATTGATAAGTACTTGAAATCCAGCAAGTATATAGCA
181   K  R  I  E  A  I  P  Q  I  D  K  Y  L  K  S  S  K  Y  I  A

601  TGGCCTTTGCAGGGCTGGCAAGCCACGTTTGGTGGTGGCGACCATCCTCCAAAATCGGAT
201   W  P  L  Q  G  W  Q  A  T  F  G  G  G  D  H  P  P  K  S  D

661  GGTTCAACTAGTGGTTCTGGTCATCACCATCACCATCACTCCGCGGGTCTGGTGCCACGC
221   G  S  T  S  G  S  G  H  H  H  H  H  H  S  A  G  L  V  P  R

721  GGTAGTACTGCAATTGGTATGAAAGAAACCGCTGCTGCTAAATTCGAACGCCAGCACATG
241   G  S  T  A  I  G  M  K  E  T  A  A  A  K  F  E  R  Q  H  M

781  GACAGCCCAGATCTGGGTACCGGTGGTGGCTCCGGTGATGACGACGACAAGAGTCCCATG
261   D  S  P  D  L  G  T  G  G  G  S  G  D  D  D  D  K  S  P  M

841  GGATATCGGGGATCCATGTCGGTACACTATAAATTTAAGAGTACACTCAACTTTGATACA
281   G  Y  R  G  S  M  S  V  H  Y  K  F  K  S  T  L  N  F  D  T

901  ATTACTTTTGATGGACTTCACATTTCTGTCGGGGACTTAAAAAGGGAGATTGTGCAGCAG
301   I  T  F  D  G  L  H  I  S  V  G  D  L  K  R  E  I  V  Q  Q

961  AAGCGACTGGGCAAAATCATCGACTTTGATCTCCAAATAACAAATGCGCAGAGTAAAGAA
321   K  R  L  G  K  I  I  D  F  D  L  Q  I  T  N  A  Q  S  K  E
```

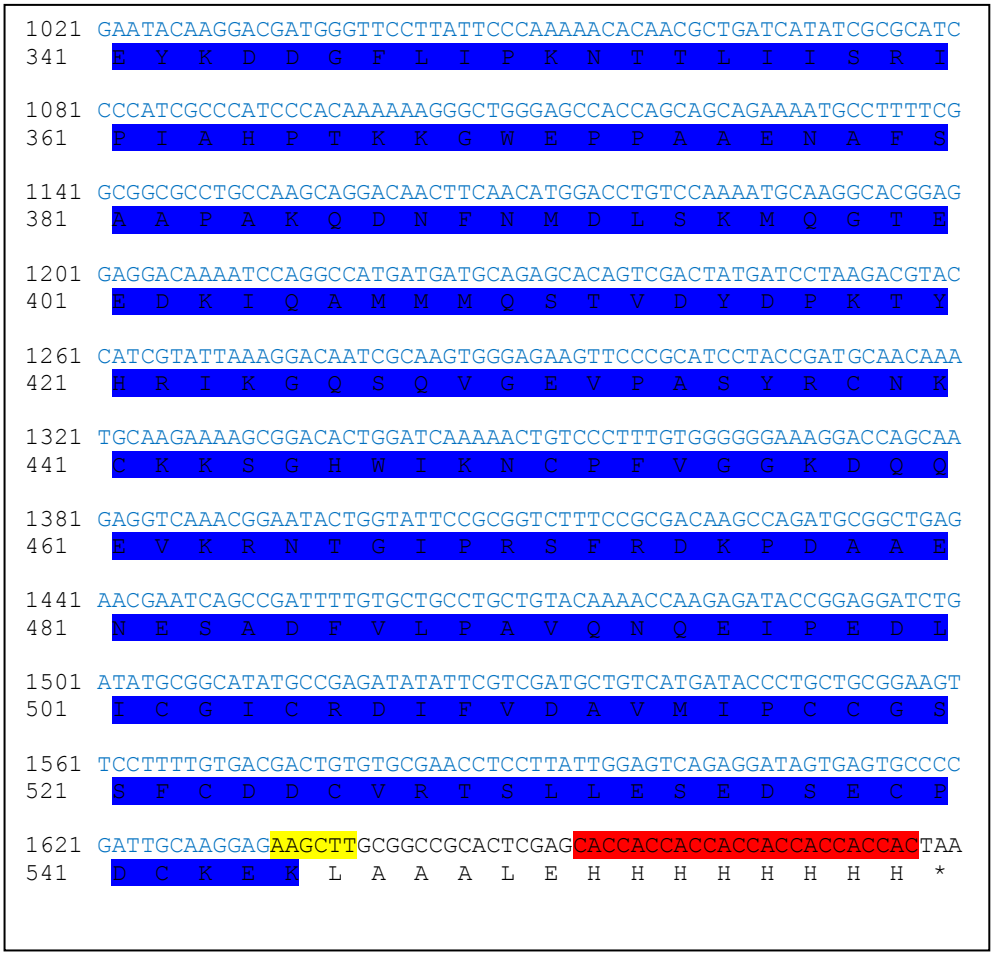

Figure D1: Amino acid sequence of the 63 kDa SNAMA DCM protein. The first 285 amino acids at the amino-terminus are identical to that of the control protein (refer to Figure C1 as well as the key in Figure B2 for colour-coded regions). Amino acids 286 to 545, high-lighted in dark blue with corresponding nucleotide sequence in blue lettering, code for the SNAMA DCM. A frameshift is introduced at the *Hind*III restriction site (high-lighted in yellow) where the SNAMA DCM carboxy-terminus is fused to the vector. As a result, the sequence downstream of *Hind*III differs between the control and SNAMA DCM proteins; the frameshift results in an additional octa-HIS-tag at the carboxy-terminus of the SNAMA DCM protein, followed by a stop-codon.

APPENDIX E

MATRIX SCIENCE Mascot Search Results

Protein View

Match to: **gi|19922884** Score: **247** Expect: **2e-018**
something that sticks like glue [Drosophila melanogaster]
Found in search of DATA.TXT

Nominal mass (M_r): **139695**; Calculated pI value: **9.74**
NCBI BLAST search of gi|19922884 against nr
Unformatted sequence string for pasting into other applications

Taxonomy: Drosophila melanogaster
Links to retrieve other entries containing this sequence from NCBI Entrez:
gi|4972740 from Drosophila melanogaster
gi|7291741 from Drosophila melanogaster

Fixed modifications: Carbamidomethyl (C)
Variable modifications: Oxidation (M)
Cleavage by Trypsin: cuts C-term side of KR unless next residue is P
Sequence Coverage: **11%**

Matched peptides shown in **Bold Red**

```
    1 MSVHYKFKST LNFDTITFDG LHISVGDLKR EIVQQKRLGK IIDFDLQITN
   51 AQSKEEYKDD GFLIPKNTTL IISRIPIAHP TKKGWEPPAA ENAFSAAPAK
  101 QDNFNMDLSK MQGTEEDKIQ AMMMQSTVDY DPKTYHRIKG QSQVGEVPAS
  151 YRCNKCKKSG HWIKNCPFVG GKDQQEVKRN TGIPRSFRDK PDAAENESAD
  201 FVLPAVQNQE IPEDLICGIC RDIFVDAVMI PCCGSSFCDD CVRTSLLESE
  251 DSECPDCKEK NCSPGSLIPN RFLRNSVNAF KNETGYNKSA AKPAAVKNEE
  301 KPPVEKEVEK KPVAEVEPEE TEVKPEKQKE SETNGSNFPK SESPEPPATT
  351 EPSQKEKDKY DSDYEDNITI KMPQPAADST TVPSKRSPSY SHRSESSHRR
  401 DRSDYVSDHD HKHQRPSKSE SVNKDRSLLP LPIGTLPSYQ GHMMAESEEA
  451 RRSSAYKPPY MQMQRGPPPM HMMSHHMPAY NNGFNNMGQR PPLSYVPYQN
  501 QSVHPMRAPY GSAGGGMNMN MSQPFQSPNL ASIYQGVAAK VGSGPIDDPL
  551 EAFNRIMKEK ERKKVDRFRS SDRHRSRSPD RQRHRFKSPM YEKDNSRDNL
  601 KDKRFRSRER KREHSYERHI RHPRSSRQPN DGSKSPGGRI KRSGHRRSAS
  651 PKPGYKSDYR DKFYNKPSAP KTEAVEPPPP GFEPLQLTDE DGYRNKHPTS
  701 SEASQSSKGD SSKKRGENRH EEAPRKRHRS RSISKEPKPN DSNYRSLTPP
  751 AKITTFKMTA AQLRQRESSP KTPEKSHDDY LTAKARIMAS QPVINDTEME
  801 TNVGKENKAK SPLSKDRKKK KKDKDKAERK KNKKDKRAKK EKGDRQKKSS
  851 SVNRSDSDIN NSSLMNESNY KVLSPRAQSP SIEINAAQLS PTHNATENVN
  901 PKSHSILTVG AASDDNLGPR SKLSEANSVN LSKWEIDENI LGLEDSSKKA
  951 AGASDDPSEI TSDVLRKAEN AIFAKAINAI RPMEPQVIIN SKDNSKDRSV
 1001 VRSDKDRSSS PRRNNSSRSV KDRLGTKISN DRSRSRDKSK GRRRAARSSD
 1051 DDANRGRSDR HGSRKRDNRS RDRAAPSEKR QERSYKRSSP EDDKLRRQNK
 1101 EQSESKHGKH DQNNSDDSDR RAAKNTKSSD SRVVSSVTAV VAPPKPCRPD
 1151 NPFRKFVDTS SSSSLVVKYD NTIQKEGASS DNGMEHRKQR DKKLKKHSKY
 1201 SSTDSLKSEK RKDPKSKKKS KILKKKKKSK K
```

Show predicted peptides also

Sort Peptides By (•) Residue Number ( ) Increasing Mass ( ) Decreasing Mass

Start - End	Observed	Mr(expt)	Mr(calc)	ppm	Miss	Sequence
9 - 30	2449.3181	2448.3108	2448.2649	19	1	K.STLNFDTITFDGLHISVGDLKR.E (No match)
41 - 54	1605.8578	1604.8505	1604.8410	6	0	K.IIDFDLQITNAQSK.E (No match)
55 - 66	1453.7236	1452.7163	1452.7136	2	1	K.EEYKDDGFLIPK.N (No match)
67 - 74	917.5549	916.5476	916.5342	15	0	K.NTTLIISR.I (No match)
138 - 152	1618.8838	1617.8765	1617.8475	18	1	R.IKGQSQVGEVPASYR.C (No match)
140 - 152	1377.6761	1376.6688	1376.6684	0	0	K.GQSQVGEVPASYR.C (No match)
180 - 185	657.3592	656.3519	656.3606	-13	0	R.NTGIPR.S (No match)
189 - 221	3713.8616	3712.8543	3712.7349	32	0	R.DKPDAAENESADFVLPAVQNQEIPEDLICGICR.D (Ions s
222 - 243	2623.1344	2622.1271	2622.0682	22	0	R.DIFVDAVMIPCCGSSFCDDCVR.T (Ions score 69)
588 - 593	754.3485	753.3413	753.3367	6	0	K.SPMYEK.D (No match)

Figure E1: MALDI-TOF identification of positive control protein. SNAMA is a Xhosa word meaning "something that sticks like glue" (Mather, 2005; Mather *et al.*, 2005). The 1231 a.a. sequence shown here is thus SNAMA-PA – refer to Figure 1.4. Residues in red lettering represent fragments of the positive control protein identified by MALDI-TOF. With the exception of the six residue long fragment spanning residues 588-593, all fragments correspond to the DCM region of SNAMA.

CHAPTER SIX

6. REFERENCES

- **Abello, N., Kerstjens, H.A.M., Postma, D.S. and Bischoff, R**. (2007). Selective acylation of primary amines in peptides and proteins. *Journal of Proteome Research* **6**, 4770-4776.

- **Antunes, R.J**. (2008). Characterization of the DWNN domain and RING finger-like motif within the DCM of the *Drosophila melanogaster* SNAMA protein. Master of Science Dissertation. University of the Witwatersrand, Johannesburg, South Africa.

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

- **Bayer, E.A., Ehrlich-Rogozinski, S. and Wilchek, M.** (1996). Sodium dodecyl sulfate-polyacrylamide gel electrophoretic method for assessing the quaternary state and comparative thermostability of avidin and streptavidin. *Electrophoresis* **17**, 1319-1324.

- **Blencowe, B.J., Issner, R., Kim, J., McCaw, P. and Sharp, P.A.** (1995). New proteins related to the Ser-Arg family of splicing factors. *RNA* **1**, 852-865.

- **Borden, K.L.B.** (2000). RING domains: master builders of molecular scaffolds? *Journal of Biological Chemistry* **295**, 1103-1112.

- **Bossis, G. and Melchior, F**. (2006). SUMO: regulating the regulator. *Cell Division* **1(13)**, 1-8.

- **Boucher, L., Ouzounis, C.A., Enright, A.J. and Blencowe, B.J.** (2001). A genome-wide survey of RS domain proteins. *RNA* **7**, 1693-1701.

- **Broemer, M. and Meier, P**. (2009). Ubiquitin-mediated regulation of apoptosis. *Trends in Cell Biology* **19**, 130-140.

- **Cajee, U.-F., Hull, R. and Ntwasa, M**. (2012). Modification by ubiquitin-like proteins: significance in apoptosis and autophagy pathways. *International Journal of Molecular Sciences* **13**, 11804-11831.

- **Chibi, M., Meyer, M., Skepu, A., Rees, D.J.G., Moolman-Smook, J.C. and Pugh, D.J.R.** (2008). RBBP6 interacts with multifunctional protein YB-1 through its RING finger domain, leading to ubiquitination and proteosomal degradation of YB-1. *Journal of Molecular Biology* **384**, 908-916.

- **Chow, M.K.M., Amin, A.A., Fulton, K.F., Fernando, T., Kamau, L., Batty, C., Louca, M., Ho, S., Whisstock, J.C., Bottomley, S.P. and Buckle, A.M.** (2006). The REFOLD database: a tool for the optimization of protein expression and refolding. *Nucleic Acids Research* **34**, 207-212.

- **Daines, B., Wang, H., Wang, L., Li, Y., Han, Y., Emmert, D., Gelbart, W., Wang, X., Li, W., Gibbs, R. and Chen, R.** (2011). The *Drosophila melanogaster* transcriptome by paired-end RNA sequencing. *Genome Research* **21(3)**, 15-24.

- **Desterro, J.M.P., Rodriguez, M.S., Kemp, G.D. and Hay, R.T**. (1999). Identification of the enzyme required for activation of the small ubiquitin-like protein SUMO-1. *The Journal of Biological Chemistry* **274**, 10618-10624.

- **Flicek, P., Amode, M.R., Barrell, D., Beal, K., Billis, K., Brent, S., Carvalho-Silva, C., Clapham, P., Coates, G., Fitzgerald, S., Gil, L.,**

Giron, C.G., Gordon, L., Hourlier, T., Hunt, S., Johnson, N., Juettemann, T., Kahari, A.K., Keenan, S., Kulesha, E., Martin, F.J., Maurel, T., McLaren, W.M., Murphy, D.N., Rag, R., Overduin, B., Pignatelli, M., Pritchard, B., Pritchard, E., Riat, H.S., Ruffier, M., Sheppard, D., Taylor, K., Thormann, A., Trevanion, S.J., Vullo, A., Wilder, S.P., Wilson, M., Zadissa, A., Aken, B.L., Birney, E., Cunningham, F., Harrow, J., Herrero, J., Hubbard, T.J.P., Kinsella, R., Muffato, M., Parker, A., Spudich, G., Yates, A., Zerbino, D.R. and Searle, S.M.J. (2014). Ensembl 2014. *Nucleic Acids Research* **42**, D749-D755.

- **Freemont, P.S.** (2000). Ubiquitination: RING for destruction? *Current Biology* **10**, R84-R87.

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

- **Gao, S. and Scott, R.E**. (2002). P2P-R protein overexpression restricts mitotic progression at prometaphase and promotes mitotic apoptosis. *Journal of Cellular Physiology* **193**, 199-207.

- **Gharahdaghi, F., Weinberg, C.R., Meagher, D.S., Imai, B.S. and Mische, S.M.** (1999). Mass spectrometric identification of proteins from silver-stained polyacrylamide gel: A method for the removal of silver ions to enhance sensitivity. *Electrophoresis* **20**, 601-605.

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

- **Hedley, M.L., Amrein, H. and Miniatis, T.** (1995). An amino acid sequence sufficient for subnuclear localization of an arginine/serine-rich

splicing factor. *Proceedings of the National Academy of Sciences* **92**, 11524-11528.

- **Hershko, A. and Ciechanover, A.** (1998). The ubiquitin system. *Annual Review of Biochemistry* **67**, 425-479.

- **Hochstrasser, M**. (2009). Origin and function of ubiquitin-like proteins. *Nature* **458**, 422-429.

- **Huang, P., Ma, X., Zhao, Y. and Miao, L.** (2013). The *C. elegans* homolog of RBBP6 (RBPL-1) regulates fertility through controlling cell proliferation in the germline and nutrient synthesis in the intestine. *Public Library of Science ONE* **8(3)**, e58736.

- **Hull, R. and Ntwasa, M.M**. (2010). Glycolytic flux occurs *in Drosophila melanogaster* recovering from camptothecin treatment. *Anti-Cancer Drugs* **21**, 945-957.

- **Hull, R.** (2012). Stress response to genotoxic agents and to infection. Doctor of Philosophy Thesis. University of the Witwatersrand, Johannesburg, South Africa.

- **Jones, C., Reifegerste, R. and Moses, K.** (2006). Characterization of *Drosophila mini-me*, a gene required for cell proliferation and survival. *Genetics* **173**, 793-808.

- **Jung, T., Catalgol, B. and Grune, T**. (2009). The proteasomal system. *Molecular Aspects of Medicine* **30**, 191-296.

- **Kappo, M.A., Eiso, A.B., Hassem, F., Atkinsin, R.A., Faro, A., Muleya, V., Mulaudzi, T., Poole, J.O., McKenzie, J.M., Chibi, M., Moolman-Smook, J.C., Rees, D.J.G. and Pugh, D.J.R.** (2012). Solution structure of RING finger-like domain of Retinoblastoma-binding protein-6

(RBBP6) suggests it functions as a U-box. *Journal of Biological Chemistry* **287**, 7146-7158.

- **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.

- **Kuhlbrodt, K., Mouysset, J. and Hoppe, T.** (2005). Orchestra for assembly and fate of polyubiquitin chains. *Essays in Biochemistry* **41**, 1-14.

- **Laemmli, U.K.** (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680-685.

- **Li, H. and Bingham, P.M.** (1991). Arginine/serine-rich domains of the *su(w^a)* and *tra* RNA processing regulators target proteins to a subnuclear compartment implicated in splicing. *Cell* **67**, 335-342.

- **Li, L., Deng, B., Xing, G., Teng, Y., Tian, C., Cheng, X., Yin, X., Yang, J., Gao, X., Zhu, Y., Sunj, Q., Zhang, L., Yang, X. and He, F**. (2007). PACT is a negative regulator of p53 and essential for cell growth and embryonic development. *Proceedings of the National Academy of Sciences* **104**, 7951-7956.

- **Lovering, R., Hanson, I.M., Borden, K.L., Martin, S., O'Reilly, N.J., Evan, G.I., Rahman, D., Pappin, D.J.C., Trowsdale, J. and Freemont, P.S**. (1993). Identification and preliminary characterization of a protein motif related to the zinc finger. *Proceedings of the National Academy of Sciences* **90**, 2112-2116.

- **Manley, J.L. and Krainer, A.R.** (2010). A rational nomenclature for serine / arginine-rich protein splicing factors (SR proteins). *Genes & Development* **24**, 1073-1074.

- **Mather, A.S**. (2005). The identification and characterisation of a novel apoptotic gene, *SNAMA*, in *Drosophila melanogaster*. Doctor of Philosophy Thesis. University of the Witwatersrand, Johannesburg, South Africa.

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

- **Matsukage, A., Hirose, F., Yoo, M.-A. and Yamaguchi, M**. (2008). The DRE/DREF transcriptional regulatory system: a master key for cell proliferation. *Biochemica et Biophysica Acta* **1779**, 81-89.

- **Ohi, M.D., Vander Kooi, C.W., Rosenberg, J.A., Chazin, W.J. and Gould, K.L.** (2003). Structural insights into the U-box, a domain associated with multi-ubiquitin. *Nature Structural & Molecular Biology* **10**, 250-255.

- **Okuma, T., Honda, R., Ichikawa, G., Tsumagari, N. and Yasuda, H**. (1999). *In vitro* SUMO-1 modification requires two enzymatic steps, E1 and E2. *Biochemical and Biophysical Research Communications* **254**, 693-698.

- **Passmore, L.A. and Barford, D**. (2004). Getting into position: the catalytic mechanisms of protein ubiquitylation. *Biochemistry Journal* **379**, 513-525.

- **Peidis, P., Giannakouros, T., Burow, M.E., Williams, R.W. and Scott, R.E**. (2010). Systems genetics analysis predicts a transcription role for P2P-R: Molecular confirmation that P2P-R is a transcriptional co-repressor. *BMC Systems Biology* **4(14)**, 1-11.

- **Priutt, K.D., Tatusova, T and Maglott, D.R.** (2005). NCBI Reference Sequence (RefSeq): a curated non-redundant sequence database of genomes, transcripts and proteins. *Nucleic Acids Research* **33**, D501-D504.

- **Pugh, D.J.R., Eiso A.B., Faro. A., Lutya, P.T., Hoffmann, E. and Rees, J.G**. (2006). DWNN, a novel ubiquitin-like domain, implicates RBBP6 in mRNA processing and ubiquitin-like pathways. *BMC Structural Biology* **6**, 1-32.

- **Rakgotho, P.M**. (2007). Identification of proteins that interact with the DWNN domain of SNAMA, a member of a novel superfamily. Master of Science Dissertation. University of the Witwatersrand, Johannesburg, South Africa.

- **Roehm, P.C. and Berg, J.M**. (1997). Sequential metal binding by the RING finger domain of BRCA1. *Biochemistry* **36**, 10240-10245.

- **Roth, M.B., Murphy, C. and Gall, J.G.** (1990). A monoclonal antibody that recognizes a phosphorylated epitope stains lampbrush chromosome loops and small granules in the amphibian germinal vesicle. *The Journal of Cell Biology* **111**, 2217-2223.

- **Roth, M.B., Zahler, A.M. and Stolk, J.A.** (1991). A conserved family of nuclear phosphoproteins localized to sites of polymerase II transcription. *The Journal of Cell Biology* **115**, 587-596.

- **Saijo, M., Sakai, Y., Kishino, T., Niikawa, N., Matsuura, Y., Morino, K., Tamai, K. and Taya, Y.** (1995). Molecular cloning of a human protein which binds to the retinoblastoma protein and chromosomal mapping. *Genomics* **27**, 511-519.

- **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.

- **Sapra, A.K., Anko, M.-L., Grishina, I., Lorenz, M., Pabis, M., Poser, I., Rollins, J., Weiland, E.-M. and Neugebauer, K.M**. (2009). SR protein family members display diverse activities in the formation of nacent and mature mRNPs *in vivo*. *Molecular Cell* **34**, 179-190.

- **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? *Journal of Cellular Biochemistry* **60**, 6-12.

- **Scott, R.E., White-Grindley, E., Ruley, H.E., Chesler, E.J. and Williams, R.W.** (2005). P2P-R expression is genetically coregulated with components of the translation machinery and with PUM2, a translational repressor that associates with the P2P-R mRNA. *Journal of Cellular Physiology* **204**, 99-105.

- **Seal, R.L., Gordon, S.M., Lush, M.J., Wright, M.W. and Bruford, E.A**. (2011). Genenames.org: the HGNC resources in 2011. *Nucleic Acids Research* **39**, D514-D519.

- **Shepard, P.J. and Hertel, K.J**. (2009). The SR protein family. *Genome Biology* **10 (242)**, 1-9.

- **Shevchecko, A., Wilm, M., Vorm, O. and Mann, M.** (1996). Mass spectrometric sequencing of proteins from silver-stained polyacrylamide gels. *Analytical Chemistry* **68**, 850-858.

- **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.

- **St. Pierre, S.E., Ponting, L., Stefancsik, R., McQuilton, P. and the FlyBase Consortium.** (2014). FlyBase 102 – advanced approaches to interrogating FlyBase. *Nucleic Acids Research* **42**, D780-D788.

- **Strickberger, M.W**. (1962). *Experiments in Genetics with Drosophila*. John Wiley & Sons, New York.

- **The UniProt Consortium**. (2014). Activities at the Universal Protein Resource (UniProt). *Nucleic Acids Research* **42**, D191-D198.

- **Tie, F., Furuyama, T., Prasad-Sinha, J., Jane, E. and Harte, P.J**. (2001). The Drosophila polycomb group proteins ESC and E(Z) are present in a complex containing the histone-binding protein p55 and the histone deacetylase RPD3. *Development* **128**, 275-286.

- **Towbin, H., Staehelin, T. and Gordon, J.** (1979). Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. *Proceedings of the National Academy of Sciences* **76**, 4350-4354.

- **Twyffels, L., Gueydan, C. and Kruys, V.** (2011). Shuttling SR proteins: more than splicing factors. *FEBS Journal* **278**, 3246-3255.

- **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. *Molecular and Cellular Biology* **21**, 8346-8356.

- **Welchman, R.L., Gordon, C. and Mayer, R.J**. (2005). Ubiquitin and ubiquitin-like proteins as multifunctional signals. *Nature Reviews Molecular Cell Biology* **6**, 599-609.

- **Witte, M.M. and Scott, R.E**. (1997). The proliferation potential protein-related (P2P-R) gene with domains encoding heterogenous nuclear ribonucleoprotein association and Rb1 binding shows repressed expression during terminal differentiation. *Proceedings of the National Academy of Sciences USA* **94**, 1212-1217.

- **Wojcik, C. and DeMartino, G.N**. (2003). Intracellular localization of proteasomes. *The International Journal of Biochemistry and Cell Biology* **35(5)**, 579-589.

- **Xu, Z., Gong, Q., Xia, B., Groves, B., Zimmermann, M., Mugler, C., Mu, D., Matsumoto, B., Seaman, M. and Ma, D**. (2009). A role of histone H3 lysine 4 methyltransferase components in endosomal trafficking. *Journal of Cell Biology* **186**, 343-353.

- **Ying, M., Huang, X., Zhoa, H., Wu, Y., Wan, F., Huang, C. and Jie, K.** (2011). Comprehensively surveying structure and function of RING domains from *Drosophila melanogaster*. *Public Library of Science One* **6(9)**, e23863.

- **Yoshitake, Y., Nakatsura, T.,Monji, M., Senju, S., Matsuyoshi, H., Tsukamoto, H., Hosaka, S., Komori, H., Fukuma, D., Ikuta, Y., Katagiri, T., Furukawa, Y., Ito, H., Shinohara, M., Nakamura, Y. and Nishimura, Y**. (2004). Proliferation potential-related protein, an ideal esophageal cancer antigen for immunotherapy, identified using

complementary DNA microarray analysis. *Clinical Cancer Research* **10**, 6437-6448.

- **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 & Development* **6**, 837-847.

- **Zhong, X.-Y., Wang, P., Han, J., Rosenfeld, M.G. and Fu, X.-D**. (2009). SR proteins in vertical integration of gene expression from transcription to RNA processing to translation. *Molecular Cell* **35(1)**, 1-10.