UNIVERSITY OF ZULULAND

Structural, biochemical characterisation of *Saccharomyces cerevisiae* Really Interesting New Gene finger domain towards drug discovery

By

Makhosazana Siduduzile MATHENJWA

(2000951523)

Thesis submitted in fulfilment of the requirement for the degree **DOCTOR OF PHILOSOPHY (PhD) IN BIOCHEMISTRY**

Department of Biochemistry and Microbiology, Faculty of Science and Agriculture, University of Zululand, KwaDlangezwa, KwaZulu Natal

Supervisor: Prof Abidemi Paul Kappo

Co-Supervisor: Prof Albertus Kotze Basson

March 2021

Structural, biochemical characterisation of Saccharomyces cerevisiae RING finger domain towards drug discovery

Makhosazana Siduduzile Mathenjwa

(200951523)

A thesis submitted in fulfilment of the requirements for the degree of Doctor of Philosophy in the Department of Biochemistry and Microbiology, Faculty of Science and Agriculture, KwaDlangezwa, South Africa

Supervisor: **Prof Abidemi Paul Kappo**

November 2021

Abstract

World-wide cancer is rated as the deadliest disease, more than TB, HIV and malaria combined. Parveen *et al*., 2015 reports that unless drastic measures are taken, cancer will statistically increase by 78% by 2030. Although different cancer therapies have been employed to combat cancer development, immunotherapy remains the current hope as a manipulative therapy to provide immune boost, more specificity and effective drug design against cancer (Vanneman and Dranoff, 2012). Really Interesting New Gene (RING) finger protein is a truncated domain from RBBP6 known to function independently in protein-protein interaction and cellular processing signalling. It is a 10 kDa protein defined by its hallmark residues cysteine and histidine that binds two zinc ions for structural stability and integrity (Dominguez., 2004; Krishna *et al.,* 2003). When human RING finger domain was used as a searching bait to fish out other possible RING finger domains from other organisms, *Saccharomyces cerevisiae* RING (*Sc*RING) finger domain stood out amongst others as it lacked the first zinc binding site which is abnormal for RING finger domain (Kappo *et al*., 2012). *Sc*RING finger domain is cysteine rich domain from a Mpe-1 RBBP6 homologue that binds zinc ion for structural stability. Therefore, biochemical, structural and *In silico* characterisation was done for this study to provide basic insight into the hypothesis that *Sc*RING finger protein is a RING finger-like protein or an abnormal RING finger protein and to determine target inhibitors for RING finger-type E3 namely, Mmd2 and MdmX.

Expression, purification, Far-UV CD, Intrinsic tryptophan fluorescence, SE-HPL chromatography and ANS fluorescence were used to biochemically and structurally characterise *Sc*RING finger protein. *In silico* inhibitory studies, for Mdm2 and MdmX complexes with RO-2443 as a control, and ZINC16046951, as the experimental ligand, were conducted using molecular dynamics tools, namely RMSD, RMSF and Rog. MM/PBSA and MM/GBSA were utilized to calculate binding free energy contributions per complex amino acids.

Far-UV CD confirmed *Sc*RING finger domain is a protein dominated by β-sheets which are a major contributing factor in protein stability and integrity. ANS fluorescence indicated the presence of hydrophobic binding pockets required for protein-protein interaction. These results provided direction towards regarding *Sc*RING finger protein as a RING finger domain rather than a RING finger-like domain. *In silico* inhibitory studies showed ZINC16046951 as the highest energy binding and Table ligand to Mmd2 compared to other complexes. These results provided the baseline information needed for drug discovery and structural determination.

Keywords: *Sc*RING finger protein, Mdm2, MdmX, p53, cancer, ubiquitination, drug discovery, ZINC16046951

DECLARATION

I declare that ***"Structural, biochemical characterisation of Saccharomyces cerevisiae RING finger domain towards drug discovery"*** is my own research and has not been submitted for examination in any other university for degree. All the work written has been indicated or acknowledged by complete references.

Makhosazana Siduduzile Mathenjwa November 2021

Signed……  ……

ACKNOWLEDGEMENTS

It is with greatest joy to write this part of my thesis for without these people this work would not have been done to completion.

Firstly, I would like to thank the chief commander of this mission, the Lord Almighty Jesus Christ who had approved my plans and gave me provision for the completion of the work.

To the Mathenjwa empire, BT Mathenjwa (my late Daddy), GZ Mathenjwa (Mommy), Sihle (sister), Thando (brother) and Khetha (Nephew), thank you for your unfailing love and support in all the spheres of life.

To my academic drivers, Prof. Kappo (Supervisor) thank you for trusting me with your project. Prof. Basson (Co-supervisor) for believing in me. Prof. Opoku (Andy) thank you for your encouraging words and believing in my dream. To the University of Witwatersrands (PSFRU) group, Dr. IA Achilonu for allowing me to work in his lab and polishing my skill in molecular biology. To Dr NN Mhlongo from University of KwaZulu Natal, for the skills in bioinformatic studies.

Thank you to University of Zululand for academic funds that helped with my travelling. To National Research Fund (NRF) for providing me with funds that assisted me towards finishing my degree with financial stress free.

I would also like to express my gratitude to my church (TWCI) for their prayers and my friends, Kho, Philly, Taiwo, Sibonelo and Nathi for keeping the spirit of education and socialising raised high. Lastly, I would like to thank my colleagues; it has been great experience being with you guys until we meet again.

"For I know the plans I have for you," says the Lord, plans to proper you and not to harm you, plans to give you hope and a future" *– Jeremiah 29 vs 11*

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS

^{113}Cd, Cd^{2+}	Cadmium ion
Amp	Ampicillin
ANS	8-Anilino-1-naphthalene sulfonate
APS	Ammonium persulphate
Cv	Column volume
Cys, C	Cysteine
Da	Dalton
dH_2O	Deionised water
DNA	Deoxyribonucleic acid
DTT	Dithiothreitol
DWNN	Domain With No Name
E1	Ubiquitin-activating enzyme
E2	Ubiquitin-conjugating enzyme
E3	Ubiquitin protein ligase
EDTA	Ethylene diamine tetra acetic acid
EEA1	Early endosome antigen1
Fab1p	Formation of aploid and binu
Far-UV CD	Far-Ultraviolet Circular dichroism
FTIR	Fourier Transformation Infrared spectroscopy
FYVE	Fab1p, YOTB, Vac1p, EEA1
GST	Glutathione S-transferase
Hdm2	Human double minute 2
HECT	Homologous to E_6-AP carboxyl terminus

His, H	Histidine
Hr	Hour
HSQC	Heteronuclear Single Quantum Coherence
Ile, I	Isoleucine
IPTG	Isopropyl β-D-thiogalactoside
kDa	Kilodalton
LB	Luria Bertani broth
Leu, L	Leucine
LIM	Lin11/ Isl-1/ Mec-3
Lys	Lysine
Mdm2	Mouse double minute 2
MdmX	Murine double minute X
Met, M	Methionine
Min	Minutes
MM/GBSA	Molecular Mechanics/ Generalised born surface area
MM/PBSA	Molecular Mechanics/ Poisson-Boltzmann surface area
mRNA	Messenger RNA
nm	Nanometer
NMR	Nuclear Magnetic Resonance
NOESY	Nuclear overhauser enhancement spectroscopy
N-terminus	Amino terminus
P2P-R	Proliferation potential protein-related
p44	Interferon-induced protein 44
Pab1p	Poly(A)-binding protein 1

PACT	p53-associated cellular protein-testes
PAGE	Polyacrylamide gel electrophoresis
PBS	Phosphate buffer saline
PDB	Protein database bank
PHD	Plant homeodomain
Phe, F	Phenylalanine
PI(3)P	Phosphatidylinositol-3 phosphate
PINCH	Particularly Interesting New Cysteine-Histidine protein
PMSF	Phenymethylsulphonyl fluoride
ppm	parts per million
pRb	Retinoblastoma gene product
Rad6	Radiation protein 6
RBBP6	Retinoblastoma binding protein 6
RBQ-1	Retinoblastoma-binding Q-protein 1
RING	Really Interesting New Gene
RMSD	Root mean square deviation
RMSF	Root mean square fluctuation
RoG	Radius of gyration
s	Seconds
SDS	Sodium dodecyl sulphate
SE-HPLC	Size Exclusion High Pressure Liquid Chromatography
SR	Serine/Arginine
TEMED	N’,N’,N’-tetramethylethylenediamine
Tris	2-amino-2-hydroxymethyprone-1,3-diol

UFD2	Ubiquitin fusion degradation protein 2
UV	Ultraviolet
YB-1	Y-box binding protein 1
Zn^{2+}	Zinc ion

CHAPTER 1

INTRODUCTION

INTRODUCTION

Cancer is one of the deadly diseases and it kills more people than AIDS, TB and malaria combined. Globally, it has been estimated to increase to 78% by 2030 which puts us at very high risk (Parveen *et al*., 2015). There are many factors that propel cancer which may either be poor life-style choices, environmental factors, infections or/and sometimes inherited within family (Parveen *et al*., 2015). However, different techniques have been applied to different patients to treat against various types of cancer. These cancer treatments include radiation therapy, chemotherapy, immunotherapy, target therapy and, in some cases, hormone therapy (Vanneman and Dranoff, 2012; Green *et al*., 2001). Radiation therapy is concomitant to chemotherapy because after radiation doses have been used to destroy cancer cells, chemotherapy, in drug form, is used to destroy cancer cells that might remain, to prevent cancer from returning (Green *et al*., 2001). Immunotherapy is the use of drugs designed from living organisms to boost the immune system to fight against cancer on its own, while target therapy studies the mechanism of action for cancer and designer drugs that will stop or block certain processes leading to the modification of DNA forming cancer cells (Vanneman and Dranoff, 2012). All these cancer treatments aim to destroy the DNA in cancer cells preventing growth and replication (Baskar *et al*., 2012). However, they are not specific to the type of cells they destroy which may lead to the knock-down of normal body cells resulting in severe side effects. Therefore, in this study, we are interested in the immunotherapy form of cancer treatment.

According to Yoshitake and co-workers (2004), it was postulated that P2P-R can be a promising immunotherapy antigen against oesophageal cancer. This was because of the observations in the proliferation of cancer cells caused by P2P-R. P2P-R is a Retinoblastoma binding protein 6 (RBBP6) homologue highly expressed in testes, while moderate to low levels are also found in other organs of the body (Hull *et al*., 2015). Apart from P2P-R, there are other homologues of RBBP6; PACT and RBQ-1 found in human and mouse cells, Mpe-1 found in yeast, RBPL-1 found in a *Caenorhaditis elegans* worm and lastly SNAMA found in the female mosquito, *Drosophila melanogaster* (Hull *et al*., 2015).

RBBP6 is a multi-domain protein consisting of six domains, namely DWNN, zinc finger, serine arginine, Rb-binding, p53 binding and RING finger (Moela *et al*., 2014). All these domains contribute to the 250kDa size of this protein which makes it a multi-functional domain, involved in programmed cell death, formation of tumours, mRNA metabolism and cell cycle progression (Moela *et al*., 2014). One of the interesting characteristics of RBBP6 protein domains is that they can replicate and express independently, making it easy to study this protein in sections. Therefore, in this study, we chose RING finger domain because, in a previous genome-wide bioinformatics study done by Kappo and co-workers (2012), they used RING finger domain derived from human RBBP6 as bait to fish for all the RING finger domains present in other organisms. Surprisingly, it was found that in three organisms, the RING finger domain was mutated and one of these was *Saccharomyces cerevisiae*. Hence this study will focus on *S. cerevisiae* RING finger domain (Kappo *et al*., 2012).

RING finger domain is a 9.8kDa protein initially discovered by Freemont and co-workers (1991) in a human genome known as RING-1, hence the origin of its name. As

aforementioned, it's a self-replicating and self-expressing protein that is composed of 7 cysteines and 1 histidine that bind to two zinc ions in a cross-brace topology for the integrity and rigidity of its three-dimensional structure (Dominguez., 2004; Krishna *et al.,* 2002). Furthermore, RING finger domain has been implicated in different functions which may either be interaction between protein and protein or with nucleic acids, cellular processes etc. The major process is ubiquitination where it acts as an E_3 ligase enzyme recruiting ubiquitin-E_2 transfer enzyme complex and mutated substrate for degradation (Chasapis *et al.,* 2010; Dominguez, 2004).

In contrast, there are other RING finger domain variants that also have the cross-brace topology, some binding to zinc ions, but they are not RING finger domains. These protein domains are LIM, FYVE, PHD and U-box, collectively known as RING finger-like domains (Eitzen *et al*., 2019; Capili *et al., 2001*). However, even though these proteins may have similar structural configuration, they have different functions (Dudev and Lim, 2003; Capili *et al*., 2001). In addition, it was also noted that the RING finger domain coordination pattern was not the same in all organisms, hence the different types of RING finger domains, namely C_3HC_4, $C_3H_2C_3$, $C_2H_2C_4$ and C_4C_4 (Lian *et al.*, 2009).

Finally, the significance of this study explores different methods used to fully characterize the *S. cerevisiae* RING finger and acquired biological data, which will be useful prior to structural determination of the protein, and to use these data as baseline parameters for the rational discovery, design and development of new drugs against cancer. The study reports the secondary structure elements of *S. cerevisiae* RING finger to allow comparison with the other reported RING finger-like domains. This is very important to confirm the biological significance of the loss of the first zinc coordinating residues within the protein sequence of the *S. cerevisae* RING finger domain.

Determination of a quaternary structure and hydrophobic pockets using SE-HPLC and ANS florescence, respectively, provided a clear understating on the composition of *S. cerevisiae* RING finger. Obtaining these results will provide clarity on whether *S. cerevisiae* RING finger domain is a true RING finger protein or a RING-finger like protein.

CHAPTER 2

LITERATURE REVIEW

2.0. LITERATURE REVIEW

2.1. RBBP6

RBBP6 is a 250kDa multi-domain protein that is found in all eukaryotes but not in prokaryotes. It is a protein that is found in the nucleoli of the interphase cells and, during mitosis, it can be found at the periphery of chromosomes in mitotic cells that lack nucleoli (Scott *et al*., 2003). Hull and co-workers (2015) documented that RBBP6 can be upregulated in the lungs, colon, oesophagus and hepatic cells of the liver. This upregulation suggests that RBBP6 may promote tumour cell proliferation (Motadi *et al.,* 2010). In normal conditions, RBBP6 facilitates DNA replication in such a way that, in the absence of RBBP6 protein, the replication fork progresses slowly. Hence, this may be an indication that RBBP6 plays a significant role in animal physiology (Hull *et al*., 2015).

RBBP6 was first discovered in humans, however several truncated versions have been found in different organisms, such as RBQ-1, PACT, P2P-R, Mpe-1 and SNAMA; these are called RBBP6 homologues.

2.2 Different Homologues of RBBP6

2.2.1. RBQ-1

RBQ-1 is a cDNA clone of RBBP6, which was initially probed using pRb protein isolated from human small lung carcinoma cells (H69c), that correspond to 150-1146 residues. It is a 140kDa protein that is composed of 948 amino acids located at chromosome 16p12.2 (Sakai *et al*., 1995). Recently, RBQ-1 has not attracted much interest, however a previous study has shown that it binds to underphosphorylated pRb and p53 in both human and mouse cells (Saijo *et al.,* 1995). Other variants such as RBQ-2 and RBQ-3 have also been reported to bind underphosphorylated pRb but their biological

significance is still unclear. The study also reported that RBQ (1, 2 and 3) competes with the adenovirus early region 1A (E1A) that binds to the pocket domain of pRb (Saijo *et al*., 1995; Simons *et al*., 1997). Lastly, sequence analysis has revealed that RBQ-1 sequence is 94% identical to PACT which may imply that RBQ-1 has a role in cell cycle regulation (Simons *et al*., 1997).

2.2.2. PACT (p53 Associated Cell Testis)

A p53 Associated Cell Testis (PACT) is a RING finger-containing protein characterized by a rich serine/arginine region and a basic lysine rich C-terminal (Simons *et al*., 1997). Through using a p53 probe, it was found that PACT expresses in the mice testis, hence the origin of its name (Simons *et al*., 1997). PACT is a 250kDa protein that is composed of 1583 amino acids which accounts for the approximate 1577bp length, which is also attached to a 437bp non-coding region polyadenylation tail (Simons *et al*., 1997). Physiologically, PACT is highly expressed in oesophageal cancer, and according to the study conducted by Li and co-workers (2007), it was observed that PACT has a significant effect on p53 function, silencing PACT in mice, resulting in early embryonic lethality and uncontrollable apoptosis. Additionally, the interaction of PACT and p53 mediates the formation of tumour and embryonic development (Li *et al*., 2007).

2.2.3. P2P-R (Proliferation Potential Protein-Related)

Proliferation Potential Protein-Related (P2P-R) is a highly basic nuclear protein (pI 9.6) cloned decades ago from a mouse cDNA library. Study has shown that it is in the nucleoli of the interphase cells (Witte and Scott, 1997). It is a 156.9kDa protein, 5173bp long, with a reading frame of 4142bp (Witte and Scott, 1997). Structurally, it is composed of four domains, namely N-terminal RING finger domain, proline rich domain, RS region and a C-terminal lysine-rich domain (Gao *et al.,* 2002; Scott *et al*., 2003). The

alternative spliced product of PACT-P2P-R gene, lacking an exon of 34 amino acids, has been said to increase immunoreactivity during mitosis (Scott *et al*., 2003). The overexpression of P2P-R blocks mitosis in prometaphase. This process promotes apoptosis, and a study has revealed that P2P-R binds to tumour suppressing proteins, known as p53 and Rb1 (Hu *et al*., 2014), which may suggest that it is also involved in cellular processes. Different functions have been attributed to P2P-R, including the cellular processes in mitosis, mRNA processing, translation, growth control and differentiation (Van Hooser *et al*., 2005). According to Yoshitake and co-workers (2004), P2P-R is highly expressed in oesophageal cancer making it a druggable protein candidate in treating oesophageal cancer. There are also some vital interactions that were observed between P2P-R and Pum2 (Pumilio homolog 2), which regulate the P2P-R expression, and the regulation gene-specific translation, in association with translation factors, located at the 3' untranslated region of their mRNA (Scott *et al*., 2005).

2.2.4. SNAMA (Something That Stick Like Glue)

SNAMA is a protein which was initially characterised by Ntwasa *et al*. (2005) and is a homologue to mammalian RBBP6, composed of 1231 amino acids residue. It has been proposed to be involved in cellular processes, such as apoptosis, that regulates the degradation of overly matured cells so that new ones can be synthesised again (Ntwasa *et al*, 2005). This was confirmed in the study carried out by Hull *et al*. (2015), where it was discovered that the deletion of SNAMA results in an abnormal apoptosis. Structurally, SNAMA is formed by the incorporation of DWNN, zinc knuckle, p53 (known as Dp53), Rb and RING finger domains, though Dp53 lacks Mdm2 (Ntwasa *et al*., 2005). There are various functions in which the SNAMA protein have been implicated,

some of which include up-regulation at the formation of embryos and rapid increase in cellular formation during the development of eyes. In addition, SNAMA is also involved in the survival of cells at the anterior during shaping of a furrow, which is facilitated by the Hedgehog signalling protein (Ntwasa *et al*., 2005). A Drosophila homologue of Rb, known as RBF, plays a role in the downstream process of Hedgehog in the fruit fly retina during cell differentiation, where it mediates cell cycle arrest in a Notch-dependent pathway. In a Notch-dependent pathway, RBF antagonization results in the release of Drosophila E2F into the nucleus, and the activity of E2F is inhibited during it interaction with Rb. However, SNAMA in Hedgehog pathways controls the inhibition by binding to RBF (Ntwasa *et al*., 2005).

SNAMA is involved in RNA metabolism and regulation of cellular homeostasis (Ntwasa *et al*., 2005). There are four gene products involved in Drosophila apoptosis, namely reaper (rpr), head involution defective (hid), grim and sickle (skl) (Lee *et al*., 2013; Goyal *et al*., 2000). Collectively rpr, hid and grim are known as RHG genes which are located at chromosome H99 and are defined as Df (3L)H99 (Lee *et al*., 2013). Skl was discovered after RHG genes and was found to be located close to rpr, just outside H99 chromosome. RHG genes work collectively to prevent accidental apoptosis; however, their expression patterns are not the same (Lee *et al*., 2013). Hid is a 43kDa molecular machinery that induce apoptosis in mammals and Drosophila by the activation of caspase pathway, arresting the activity of the inhibitor of apoptosis protein (IAP) (Lee *et al*., 2013). Grim is a major cell death gene for apoptosis of the ventral nerve cord corazonin (vCrz) and EW3-sibling cells (EW3-sib), however it needs skl as a supporting proapoptotic factor. According to Lee and co-workers (2013), skl acts in a superfluous manner with other death genes (RHG) and plays a minor role in vCrz neuronal death

(Lee *et al*., 2013). In addition, the reaper gene also helps in recruiting the tumour suppressing proteins, either by initiating apoptosis or through cell development arrest (Lee *et al*., 2013; Goyal *et al*., 2000).

2.2.5. RBPL-1(Retinoblastoma Binding Protein Like-1)

The sixth RBBP6 homologue discovered later is RBPL-1, found in the round worm known as *Caenorhaditis elegans* (Hull *et al*., 2015). It is a 154kDa nucleolus protein made up of DWNN, zinc knuckle and RING finger domains (Huang *et al*., 2013). Unlike RBBP6, RBPL-1 lacks the tumour suppressor proteins (p53 and Rb) and SR domain. It is highly expressed in the intestine, neurons, spermatheca and vulva. In the study conducted by Huang and colleagues (2013), they inhibit the expression of RBPL-1 through RNA interference. This study enabled them to discover various functions possessed by this protein, some of which include its developmental role in embryos and the production of gametes (Huang *et al*., 2013). It was further averred that worm, with the silenced RBPL-1 eggs, die before hatching and have defects in oocyte functioning when compared to the control. Also, the production of oocytes and the development of the germline is negatively affected (Huang *et al*., 2013). Moreover, the progeny brood size and fertility were observed to be reduced, confirmed by physical observation as the oocytes appeared weak, sickly and smaller, and the intestine, which is the major organ where RBPL-1 is regulated, almost disappeared compared to the control worms (Huang *et al*., 2013). This overall observation indicates that RBPL-1 is implicated in the up-regulation of nutrient up-take. Lastly, like Mpe-1 that regulates the cleavage and polyadenylation processes, bioinformatical interaction studies predicted proteins (involving the cleavage and polyadenylation processes) interact with RBPL-1; indicating

that RBPL-1 may also functions as an mRNA cleavage and polyadenylation factor to control the expression of a wide range of genes *C. elegans* (Huang *et al*., 2013).

2.3. Mpe-1 (Mutant PCF11 Extragenic suppressor) homology

Mpe-1 is a cloned protein which was initially characterised by Vo and co-workers (2001) using a DNA isolated from *pcf11-2 mpe1-s* strain. It is a human RBBP6 homology from *Saccharomyces cerevisiae* in which the *Sc*RING finger domain originates. Studies have shown that the Mpe-1 protein contains 441 amino acid residues with a molecular weight of 49.5kDa (Vo *et al*., 2001; Ndabambi, 2004).

2.3.1. Mpe-1 Domains

Unlike human RBBP6 that is composed of six domains, Mpe-1 is a protein composed of three conserved domains. Figure 2.1 illustrates the positioning of the three domains, their comparison with other three organisms including the polypeptides encoded by three cDNAs from *homo sapiens*. The first domain is a DWNN which is located between 5 and 78 amino acid regions in Mpe-1 (Vo *et al.,* 2001; Di Giammartino *et al*., 2014). The second domain is the zinc knuckle which is found between 182 and 195 of the amino acids sequence (Vo *et al.,* 2001) and is said to be responsible for protein interaction and *si*DNA/RNA interaction (Vo *et al.,* 2001). Lastly, it's the cysteine-rich domain, which is located between 284 and 343, this is a RING finger domain better known an *Sc*RING finger domain (Vo *et al*., 2001). This domain is involved in protein-protein interactions and the ubiquitination pathway (Lee and Moore, 2014). The recorded identical score between *Sc*RING finger domain and *H. sapiens* RING finger domain was 29%, which was accepTable considering that Mpe-1 is a truncated homology of a Human RBBP6 (Vo *et al*., 2001). The latter may be postulated as one of the factors a *Sc*RING finger domain lacks the first zinc binding residue.

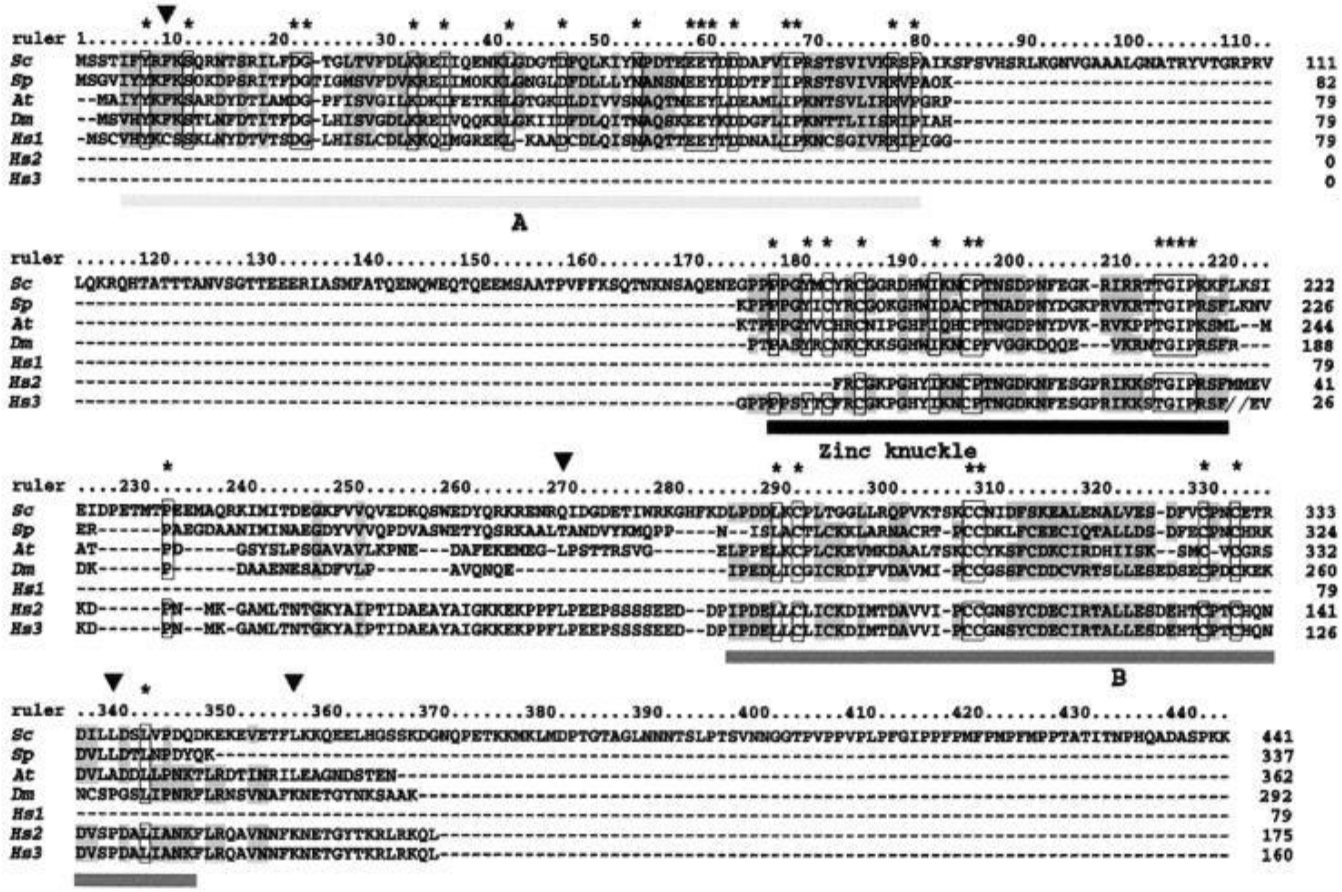


Figure 2.1: A human genome RBBP6 homologue Mpe-1. The three conserved domains sequences were compared to *Saccharomyces cerevisiae, Schizosaccharomyces pombe, Arabidopsis thaliana, Drosophila melanogaster* and three polypeptides encoded by cDNAs from *Homo sapiens* (denoted by *Hs*1-3). Residues highlighted in light grey represents similar residues, identical residues are framed by the asterisks and arrows points to the positions of the mutations in the mpe1-1 mutant (where, amino acid 9, phenylalanine changed to a serine; amino acid 268, glutamine changed to a lysine; amino acid 337, lysine changed to a phenylalanine; and amino acid 354, lysine changed to a stop codon). (Vo *et al., 2001*)

2.3.2. Mpe-1 encodes a subunit protein responsible for 3’ end formation

Mpe-1 is a conserved protein in all eukaryotes that encodes a subunit protein of cleavage factor 1(CF1) required for 3’ end processing (Lee and Moore, 2014). This is a shared characteristic with the N-terminal of RBBP6 as its also contained the three conserved domains observed in Mpe-1 (Di Giammartino *et al*., 2014). The 3’ end processing is a process required for the maturation of mRNA; the precursor of polypeptide chains, which transports from the nucleus to the cytoplasm, where they are translated into proteins (Darmon and Lutz, 2012). However, prior to the translational process, the mRNA needs to undergo an extensive co-transcriptional process. The efficiency of the 3’ end processing is completed by 14 protein factors in mammals and about 20 protein factors in yeast (Mandel *et al*., 2008). 3’ end processing is a very important process that maintains homeostasis in cellular growth and viability. Specifically, the 3’ end processing promotes the transfer of mRNAs from the nucleus to the cytoplasm and stabilizes mRNAs by adding the polyadenylation tail, since mRNAs are degraded initially from the 3’ end (Mandel *et al*., 2008). Furthermore, the interaction of the polyadenylation tail and polyadenylation binding protein (PABP) to methyl cap 5’ end enhances mRNA translation, while transcription initiation is regulated by an interaction complex of polymerase II C-terminal domain and transcription factors (Mandel *et al*., 2008). In addition, a proper polyadenylation signal is responsible for transcription termination. More so, the cleavage and the polyadenylation mechanisms are directed by sequence elements known as *cis*-acting elements and *trans*-acting protein factors (Mandel *et al*., 2008). Lastly, no much difference is observed between RBBP6 and Mpe-1 except for the structural observation in the lack of a C-terminal in Mpe-1 protein.

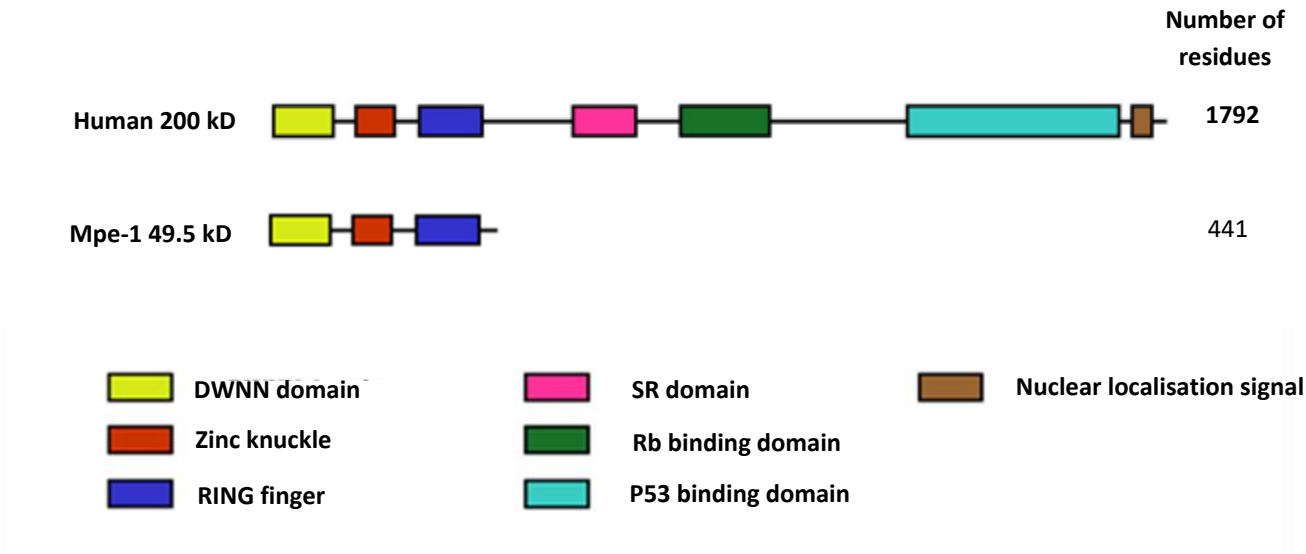


Figure 2.2: RBBP6 structural domains found in different organisms. DWNN (Domain With No Name), Zinc knuckle, RING (Really Interesting New Gene), SR (domain serine/arginine-rich region), Retinoblastoma (Rb)-binding domain, and the p53-binding domain, respectively. However, as shown in the diagram Mpe-1 lacks some of the domains in the C-terminal which, therefore, create a truncated version of the RBBP6. Updated from Pugh *et al.*, 2006.

2.4. RBBP6

RBBP6 is known as a multi-functionality protein because of the different domains it is composed of that have unique functions. In total, RBPP6 is made up of six domains, namely DWNN, zinc knuckle, serine/arginine, Rb-binding, p53 binding and RING finger domain (Pugh *et al*., 2006). These entire domains have interesting characteristics to express on their own and are very sTable as single entities, which makes it possible to study them individually (Pugh *et al*., 2006). This section gives a brief overview of the characteristics and biological roles of each domain.

2.4.1. DWNN (Domain With No Name)

Domain With No Name (DWNN) is a 20 kDa protein; it is the first domain in RBBP6 located at the N-terminal (Skepu *et al*., 2000). DWNN is found in all eukaryotes species as a single copy except in *Arabidopsis thaliana.* Looking at the makeup of this domain, it is seen to be attached to the tumour suppressor protein (p53 and Rb); this indicates its involvement in apoptosis (Skepu *et al*., 2000). The structure of DWNN is composed of 7 β-sheets and 1 α-helix and its topology resembles that of a ubiquitin-fold, making it an ubiquitin-like modifier (Pugh *et al*., 2006). This is confirmed bioinformatically by their z-scores, where DWNN and ubiquitin are sitting at 7.5 and 7.6 respectively and have 18% similarities in their amino acids (Pugh *et al*., 2006). Resemblance of DWNN to ubiquitin is divided into two; one being the ubiquitin domain protein (UDP) and the other, the ubiquitin-like domain (UBL) (Cajee *et al*., 2012). In the ubiquitination pathway, the cleavage protease recognises the GG motif in ubiquitin which forms the *di*glycine moiety during conjugating processes (Chatr-Aryamontri *et al*., 2018). The presence of this motif in humans and mouse DWNN insinuates their existence in the conjugation process informally known as “DWNNylation” (Pugh *et al*., 2006). However, the same motif is not

expressed in lower organisms, implying that DWNN is not a covalent modifier in lower organisms (Pugh *et al*., 2006).

2.4.2. ZINC KNUCKLE (Zinc finger domain)

Zinc is a metal ion that acts as an essential cofactor for both enzyme metabolism and transcription factors. Its binding site is divided into two, either for catalysation or structural stability (Dudev and Lim, 2003). For coordination purposes, zinc prefers soft ligands, such as cysteine and histidine, but is also found to coordinate other ligands like aspartic acid and glutamine. Normally, the zinc structure is coordinated in a tetrahedral manner; however, cases where it coordinates 5-6 ligands have also been reported in catalytic sites (Dudev and Lim, 2003). Most proteins are coordinated in a tetrahedral manner, this manner and this configuration has been proved as the most sTable compared to other coordination patterns (Krishna *et al*., 2003). This is confirmed by the inaccessibility of solvents and protein structural ability to transfer charge from the ligands to metal cations (Houben *et al*., 2005). Studies have shown that the substitution of zinc metal binding for other heavy metal ions can defect the stability of a protein tetrahedral structure, resulting in loss of protein and nucleic acid binding capability (Dudev and Lim, 2003). Furthermore, studies have demonstrated that the substitution of zinc by other heavy metals can be rectified by machineries known as metallothionein, removing heavy metals and replacing zinc ion cofactors, hence repairing zinc binding sites (Houben *et al*., 2005; Dudev and Lim, 2003). The zinc finger domain core is composed of a zinc ion, bound to cysteine and histidine residues, creating a hydrophobic core. Although, various zinc finger domains may have the same structural configuration, it does not necessarily mean they all have the same functions or related roles (Krishna *et al*., 2003). Zinc finger domains are classified into eight groups,

depending on their similarities and differences in function implication, which include C_2H_2-like finger, Gag knuckle, treble clef finger, zinc ribbon, zn_2/cys_6, TAZ2 domain like, zinc binding loops and metallothionein (Krishna *et al*., 2003). However, this study will focus on Gag knuckle which is familiarised as zinc knuckle or zinc finger domain.

Gag knuckle is composed of two β-sheets connected to a zinc knuckle followed by either a helix or loop (Holub, 2017). To complete the structural fold, addition of two N-terminal zinc ligands from the zinc knuckle and the other two either from the loop or at both ends of a short helix are present (Holub, 2017). Gag knuckle is a truncated form of C_2H_2 composed of approximately 20 amino acids (Larsen, 2018). There are about three family divisions of Gag knuckle which contain C_2HC zinc fingers from the retroviral gag proteins. The first one is a retroviral Gag knuckle family defined by the presence of one-turn α-helix, followed by a β-hairpin (Larsen, 2018). The structural motif is more familiar in the retroviral nucleocapsid protein from HIV and has also been reported in other related viruses. Gag knuckle plays a role in the recruitment of RNA needed for viral packaging by recognising their specific sequences and binding to single-stranded RNA (Carlson *et al*., 2016). Second, the polymerase Gag knuckle family demonstrates the helix structure that has been replaced by the loop, indicating a possibility of it being involved in RNA-binding functions (Carlson *et al*., 2016). Finally, the reovirus outer capsid protein σ3 Gag knuckle family is composed of a zinc binding motif and mutation of zinc-binding region that has been reported to have negative consequences in its ability to interact with the capsid protein μ1 (Lemay, 2018).

2.4.3. SERINE/ARGININE (SR) DOMAIN

Serine/Arginine domain is the third domain of RBBP6 protein, and like its neighbouring domains, it also has the character of expressing independently of other domains.

Studies have shown that structurally it is composed of an arginine/serine (RS) domain located at the C-terminal, and 1 or 2 RNA recognition motif (RRM) domains located at the N-terminal (Graveley and Maniatis, 1998). Although these domains make up the SR domain, they have individual responsibilities in terms of functioning; the RS domain role is an interactive process, between protein-protein and protein-RNA, whilst RRM is specifically implicated in protein-RNA interactions (Kumar *et al*., 2015). RS and RRM are both involved in splicing; a process of editing pre-mRNA to mature mRNAs by cleaving of introns and ligating exons, but it is noted that RS and RRM work independently of each other (Kumar *et al*., 2015). Activation of serine residues in the RS domain from SR protein is through phosphorylation, by which SR protein functions are mediated. In contrast, up-regulation and down-regulation of phosphorylation has been reported to inhibit the splicing process (Kumar *et al*., 2015). During splicing, SR proteins promote the binding of U1 small nuclear Ribonuclearproteins (snRNP) to 5' sites, by binding to U1-70K followed by U2A^{35} located near 5' and 3' slice sites respectively and simultaneously (Hollander *et al*., 2016). *Saccharomyces cerevisiae,* on the other hand, lacks an easily identifiable RS-rich domain and, in general, the splicing efficiency of different pre-mRNA varies from individual types of SR proteins (Sahebi *et al*., 2016). Splicing factors, together with SR proteins, combine to form a spliceosome, a functional machinery in splicing. For accurate selections of splicing sites, a serine/arginine-rich splicing factor 1(SRSF1) is required (Sahebi *et al*., 2016). SRSF1 is a SR archetypical that plays a role in facilitating transcription while sustaining sTable genome and gene network organization (Jobbins, 2018). Finally, up-regulation of SRSF1 has been observed in human cancers and HIV-infected cells, indicating SRSF1 as a proto-oncogene (Anczuków and Krainer, 2016; Finley, 2015).

2.4.4. RB-BINDING (Retinoblastoma-binding) Domain

pRB domain was the first identified tumour suppressor discovered in retinoblastoma (childhood cancer) by Knudson (1971). A tumour is a neoplasm that develops due to a disorder of cell cycle progression. At the initial stage of tumour formation, G1 phase is caused by an absence of retinoblastoma-binding protein (pRB) (Giacinti and Giordano, 2006). Various human cancer forms are an implication of either a deletion or mutation of the pRB gene, which in turn makes the system highly susceptible to tumour development (Giacinti and Giordano, 2006). Retinoblastoma binding protein is composed of approximately 928 amino acids residue and is located in the middle part of RBBP6 (Thwaites *et al*., 2017). There are about three family members that encompass pRB and collectively they are known as “pocket protein”; a name derived from a conserved binding characteristic of these pocket proteins to viral oncoproteins and transcriptional factors, such as E2F (Vélez-Cruz and Johnson, 2017; Thwaites *et al*., 2017; Giacinti and Giordano, 2006). These pocket proteins include p105, p107 and p130 and, when simultaneously over-expressed, they inhibit cell cycle progression (Giacinti and Giordano, 2006). The p107 is located at human chromosome 20q11.2 area while p105 and p130 are found in chromosome 13q14 and 16p12.2 respectively (Giacinti and Giordano, 2006). Although they are from the same family, study has proved that they cannot perform all the pRB functions (Vélez-Cruz and Johnson, 2017). One of the major functions of pRB is suppression of E2F by binding directly to transactivation factor to avoid the recruitment of transcriptional activators. Additionally, proteins such as histone deacetylase and acetyltransferases are also recruited for transcriptional suppression (Vélez-Cruz and Johnson, 2017).

Tumorigenesis develops when pRB is phosphorylated by cyclin dependent kinases, resulting in a conformational change of pRB releasing E2F, hence restarting cell cycle progression to S-phase. There are eight family members of E2F, however E2F1-3 are the primary targets of pRB and they are transcriptional activators (Vélez-Cruz and Johnson, 2017). Transcriptional repressors E2F4 and E2F5 are targeted by p107 and p130. Lastly, E2F6-8 are also transcriptional repressors but independently of pRB. Besides phosphorylation, pRB is also subjected to other post-translational modifications such as methylation, acetylation, SUMOlatyon and ubiquitylation (Vélez-Cruz and Johnson, 2017). The role of p107 and p130 is to block the recruitment programme of transcriptional co-factors, hence inhibiting E2F-target gene activity. Summing up these pRB activities, the main purpose is genome stability, and failure may result in severe consequences, such as abnormal cell proliferation, loss of cell cycle arrest, defects in apoptotic signalling and hence tumour formation (Vélez-Cruz and Johnson, 2017). This makes pRB a druggable protein against cancer (Vélez-Cruz and Johnson, 2017). Furthermore, it has been elucidated those different kinds of tumour developments, either in retinoblastoma, carcinomas, or osteosarcomas, are as a result of an inactivation in pRB caused by phosphorylation, deletion, mutation, deletion or binding of viral oncoproteins (Nicolay *et al*., 2016; Vélez-Cruz and Johnson, 2017). More so, other duties of pRB are to mediate the differentiation process in various parts of the body, such as the brain, eyes, lens, peripheral nervous system, skin cells, liver, muscles, and hair cells (Eva *et al*., 1992). This was confirmed by a study conducted by Eva *et al*. (1992) where mice with knock-out pRB died 14 days after being conceived due to faults in central nervous system and retarded hematopoietic cell growth.

2.4.5. p53 Binding Domain

8 years after the discovery of pRB, p53 was discovered as a tumour suppressor protein by Arnold Levine, David Lane and William Old (1979), and later Orlova (2006) determined its structure. The p53 tetramer structure in mice was determined using a cryo-electron microscope and it was observed to resemble a shallow, skewed cube configuration stabilised by ATP (Okorokov *et al*., 2006). An overview from the EM microscope revealed that p53 consists of an upper layer and a lower layer that are relative to each other (Okorokov *et al*., 2006). The upper layer has two small nodes and two large nodes forming an apex of a parallelogram, with the similar structure underneath rotating at 70º directly proportional to the upper layer. As shown in Figure 2.3, each side has C-terminal on the left and N-terminal close to each other in a "hand-to-hand" manner (Okorokov *et al*., 2006).

p53 is a 53kDa protein located at chromosome 17p13.1, composed of 393 amino acids and five functional domains shown in Figure 2.4 (George, 2011). These five functional domains include the N-terminal domain, occupying 1-43 amino acid lengths which play a role of transcriptional activator, a proline rich domain which occupied position 63-97 amino acids residue and it is a binding-site for MdM2 (Vassilev and Fry, 2011). More so, the central core domain from length 100-300 also binds DNA and is a location for most oncogenic p53 mutations. An oligomerization domain, which is another motif in p53, occupied position 320-360 residues length which consists of nuclear localization signals and is responsible for tetramerization of p53 (George, 2011; Vassilev and Fry, 2011). The C-terminal domain occupies residue length from 364-393 and its function is to negatively regulate sequence-specific DNA from binding by central core domain (George, 2011).

p53, is an important protein in prevention of cancer which is responsible for the cell cycle progression, apoptosis, DNA damage and signal transduction for abnormal proliferation; depending on types of stress induced (George, 2011). However, the basal level of p53 is required to perform these functions because an excessive level of this protein results in cellular processes retardation, either in cell growth development or apoptosis. It may also cause a significant reduction in p53 polyadenylation (George, 2011), in most cases due to knockdown of PACT which leads to p53-Hdm2 interaction.

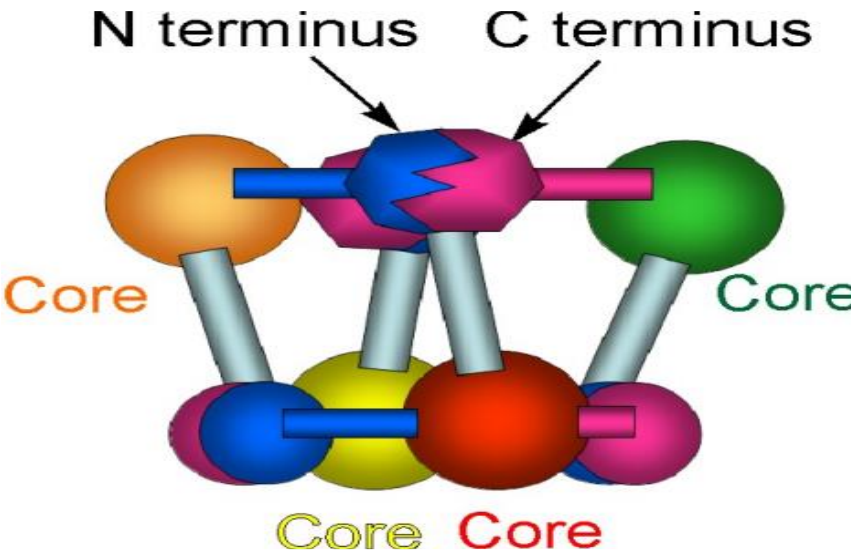


Figure 2.3: **The p53 tetramer structure.** The N-termini are represented by the blue colour and C-termini by magenta, the spheres represent the core nodes of the structure and grey rods are the termini linkers. Figure taken from (Okorokov *et al*., 2006).

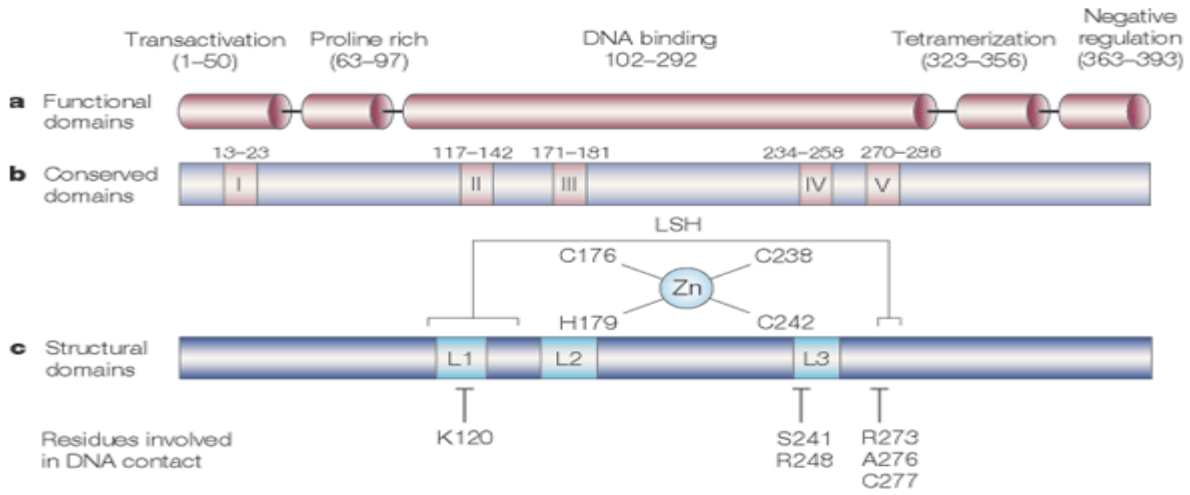


Figure 2.4: p53 functional domains. Schematic diagram representing five p53 functional domains which are, N-terminal (1-50), proline rich (63-97), DNA binding (102-292), tetramerization (323-356) and negative regulation (363-393). Taken from (George, 2011).

Another p53 homologue is p63, which is located at chromosome 3q27-29, and p73, located at chromosome 1p36.2-3. Collectively all three have a highly similar central core domains (60%) and are least similar at the oligomerization domain rating at 30% (Gonzalez *et al*., 2017).

In normal conditions, p53 is latent with no stressors activating it, such as DNA damage, telomere erosion, activated oncogenes, ribonucleotide depletion and/or deficiency of oxygen in tissues (George, 2011). If these stressors are not quickly attended to, they can lead to tumorigenesis in which 50% of human cancer forms. It has been reported that p53 is usually mutated in all cancer forms (George, 2011). The latent form of p53 is accumulated in the cytoplasm, whereas the activated form is located inside the nucleus where it performs its biochemical activities. Moreover, other signalling agents of p53 activated are the oncoproteins such as β-catenin, Myc, Adenovirus E1A and Ras (Qin *et al*., 2018). Therefore, when p53 has been activated, it performs different cellular functions to respond to stressors. These cellular functions are achieved when p53 binds to E2F transcription factor, in order to block cell cycle progression and avoid transcription of damaged DNA from replicating and transmitting from one generation to another, whilst allowing DNA to repair (Komori *et al*., 2018). This is followed by an emergency brake in apoptosis, to eliminate and degrade the DNA beyond repair, and prevent normal cells from developing to cancerous cells. p53 as a tumour suppressor, when activated, suppresses transcription, in turn activating genes tethering cell cycle control. One of these genes is p21 that inhibits the binding of cyclin dependent kinase, allowing p53 to block cell proliferation (Fischer *et al*., 2015; Al-saran *et al*., 2016). The presence of p53, during chemotherapy and radiotherapy, increases its efficiency, but when mutated, the therapy efficiency is reduced. Furthermore, the level of p53 is highly

induced by the presence of protein kinases such as ATM, ATR, DNA-PK, Chk1 and Chk2, preventing binding of MdM2 to p53 through induction of ARF to bind MdM2 (Okorokov *et al*., 2006). These protein kinases mediate this by phosphorylating p53 at residue ser15, thr18 and ser20. Additionally, when the protein has been rectified, cyclin dependent kinase binds to p53, releasing E2F and allowing the interaction of MdM2 with p53, while inhibiting apoptosis and reinitiating cell cycle progression and cell proliferation (Okorokov *et al*., 2006).

2.4.6. REALLY INTERESTING NEW GENE (RING) DOMAIN

In 1991, Freemont and co-workers discovered a zinc-finger characterised by its richness in cysteine residues, known as RING finger domain. Initially it was discovered from a human gene product known as Really Interesting New Gene (*RING-1*) at chromosome 6 position, hence the discovery of its name (Freemont *et al*., 1991). The RING finger domain archetypal structure is composed of 7 cysteines and 1 histidine, in which 4 of those cysteine residues coordinate the second zinc ion, and the remaining residues coordinate the first zinc ion, as shown in Figure 2.5 (Dominguez, 2004). Zinc ion binding contributes to the protein structure stability and is called the tetrahedral configuration, otherwise known as “cross-brace” topology (Miyamoto *et al*., 2017).

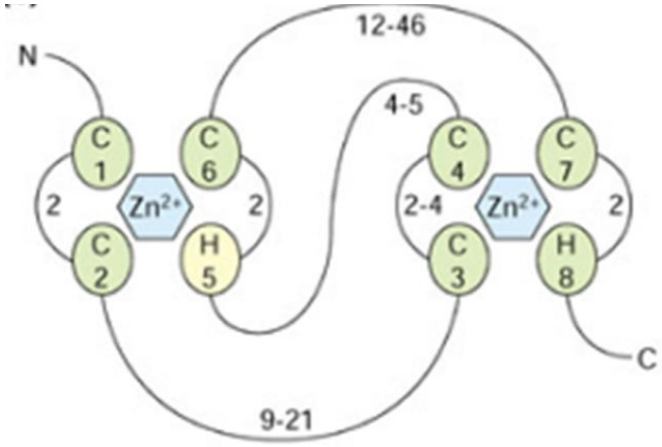


Figure 2.5: Schematic representation of a RING finger domain. The blue hexagon represents the zinc ions, 7 green spheres represent cysteine residues and cream sphere represents histidine residue. Taken from (Kosarev *et al*., 2002)

RING finger domain is highly defined by its consensus sequence Cys_1-Xaa_2-Cys_2-$Xaa_{9–39}$-Cys_3-$Xaa_{1–3}$-His_4-Xaa_{2-3}-Cys/His_5-Xaa_2-Cys_6-$Xaa_{4–48}$-Cys_7-Xaa_2-Cys_8, where X symbolises the amino acids while Cys and His represent cysteine and histidine, respectively (Joazeiro and Weissman, 2000). Cysteine and histidine are highly conserved residues in RING finger domain which form a hydrophobic active site. It is a 9.5kDa protein composed of approximately 70 amino acids located at the N-terminal of the RBBP6 (Callis *et al*., 2005). This gives the RING finger domain a globular configuration (Figure 2.6) which consists of a central α-helix that is involved in active site stability, and variable length loops which are separated by small β-strands, responsible for restricting the unfolding of the protein, and hence, proteolysis (Diog *et al*., 2001; Perczel *et al*., 2005; Kappo, 2009).

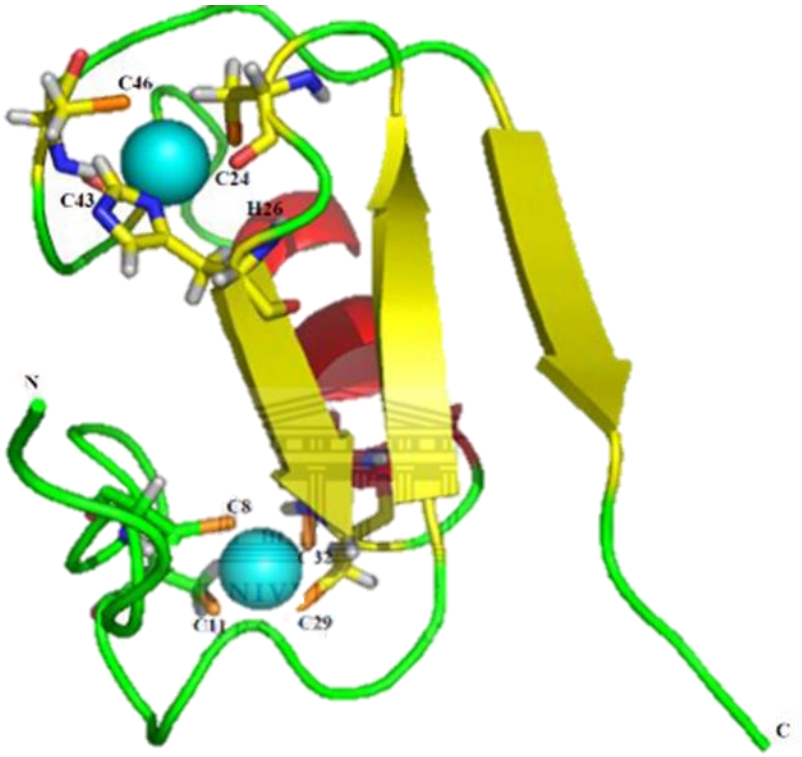


Figure 2.6: Tertiary structure of a C3HC4 RING finger domain. The structure shows the presence of three anti-parallel β- strands located near a central α-helix and variable length loops. The structure is stabilised by two zinc ions which are cyan spheres, the first coordinated zinc is bound to Cys24, Cys43, Cys46 and His26 residues and the second zinc is coordinated by Cys8, Cys11, Cys29 and Cys32 forming a tetrahedral configuration. Figure taken from Kappo 2009.

Up till today, there is no specific biological function(s) that has been linked with RING finger domain, except the various study reports of larger proteins (e.g. RBBP6, TRIM etc.) where a RING finger domain is involved in functional roles, such as an interaction promoter either between protein-protein during ubiquitination, when its binds with E2 conjugating enzyme as an E3 ligase enzyme to facilitate the binding of small ubiquitins to misfolded proteins, which then serve as the signalling factor to 26S proteasome, or in protein-nucleic acids where it promotes the degradation of damaged DNA and mRNA to prevent incorrect transcription (Aldaghmi, 2017; Miyamoto *et al*., 2017; Qi *et al*., 2016). It is also involved in cellular processes such as viral functions, apoptosis, oncogenesis and cell growth (Dominguez, 2004; Joazeiro and Weissman, 2000). A family of TRIM (tripartite motif-containing protein) are RING finger containing proteins and TRIM16 (a TRIM family member) has been postulated to have a vital role in different types of cancer, such as in reducing neuroblastoma cell growth, tumorigenicity and in retinoid-induced differentiation enhancement *in vivo* (Qi *et al*., 2016). This was confirmed by the studies presented by Qi and colleagues in which the downregulation of TRIM in prostate cancer cells increased their invasiveness, affecting other healthy cells. Hence RING finger domain plays an important role as a tumour suppressor. However, the major function, where RING finger domain is involved as single entity, is in proteolysis of modified or damage proteins in a process known as ubiquitination as shown in Figure 2.7 (Aldaghmi, 2017; Miyamoto *et al*., 2017).

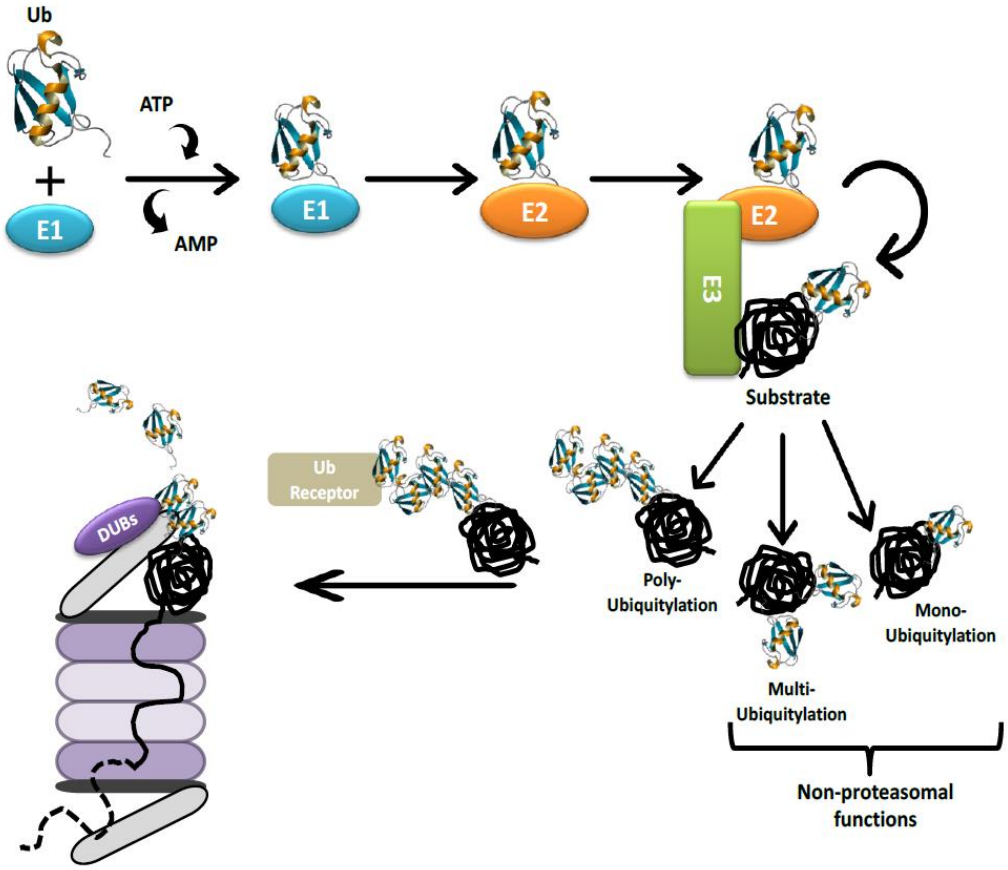


Figure 2.7: The ubiquitination pathway. UB (ubiquitin), E1 (Activating enzyme), E2 (Transfer/Conjugating enzyme), E3 (Ligase enzyme). This is a mechanism that is responsible for the targeting of modified or misfolded proteins which are then directed to a 26S proteasome for degradation back to peptides that will be used to form new proteins. Native ubiquitin is activated by E1 which is then conjugated to E2 in a transthioesterification reaction. The E2-Ub thioester complex is then transferred to E3 ligating enzyme and when E2 detaches; E3 is responsible for recruiting mutated protein substrates and attaching them to polyubiquitination chain that serves as a signalling factor to a 26S proteasome. Taken from Kocaturk, 2018.

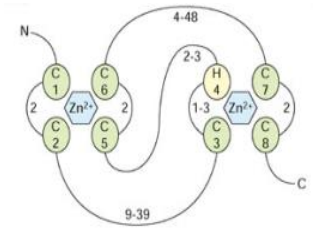
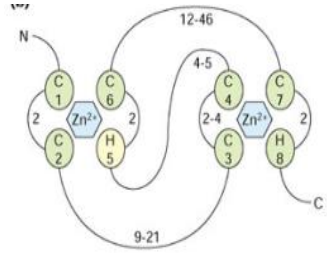
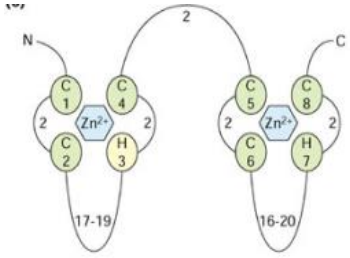
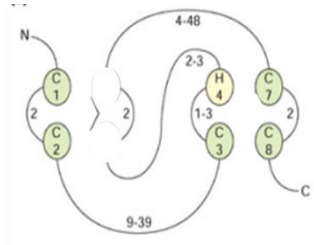
Ubiquitination is a multi-enzymatic irreversible process where modified intracellular proteins are tagged by ubiquitin for degradation. A cascade of three enzymes are involved for the success of this process, the first being E_1 activating enzyme which binds to ubiquitin, in an ATP dependent manner, forming a thioester linkage at position Gly76 (Callis *et al*., 2005). The activated ubiquitin then disassociates from the E_1 active site and binds to the E_2 conjugating enzyme active site in a process known as the transthioesterification reaction (Miyamoto *et al*., 2017). An E_2-Ub thioester complex is then delivered to the E_3 ligating enzyme where E_2 is dislocated, releasing ubiquitin to bind at the lysine residue of the specific modified intracellular protein (Miyamoto *et al*., 2017). The specificity of the modified intracellular protein is regulated by the presence of degrons which serve as the signalling machineries for such protein damages (Guharoy *et al*., 2016). The dislocated E_2 therefore continues to supply more ubiquitin molecules in the same process until 8-10 are joined together, forming a polyubiquitination chain that is recognised by the 26S proteasome and therefore degrades the protein back to its peptide, allowing the reformations of the correct proteins (Pickart and Fushman, 2004).

2.4.6.1. *S. cerevisiae* RING finger domain

Although not much publication in found on *Sc*RING finger domain, however, Finely *et al*., 2012 reported that there about 44 *S. cerevisiae* protein with RING domain E_3 ligase in which two of those also include a U-box domain. The comparison in these protein structure highlights that *Sc*RING finger protein binds a single zinc metal ion for structural stability while the U-box domain depends on the hydrogen bond for structural stability (Finely *et al*., 2012). *Sc*RING finger protein is involved in ubiquitination as an E_3 ligase and unlike other ligases *Sc*RING finger protein E_3 ligase does not form a thioester, rather it facilitates the transfer of ubiquitin delivered by E_2 conjugating enzyme

to the modified intracellular protein (Finely *et al*., 2012). *Sc*RING finger protein portrays a truncated version of a cross-brace topology of a human RBBP6 RING finger domain, whereby, the first zinc binding site is lost (Kappo *et al., 2012*). RING-finger like domains such as PHD, FYVE and U-box also have the cross-brace topology as shown in Table 2.1, but they are different from *Sc*RING finger protein due to spaces between the coordinating residues and functions they facilitate (Dudev and Lim, 2003). However, it still difficult to categorise a *Sc*RING finger protein on whether it is a RING-finger like protein that is involved in ubiquitination like a U-box or a RING finger protein that will be the first to be reported as a single zinc ion binding RING finger domain (Aravind and Koonin, 2000). Additionally, *Sc*RING finger protein unlike the classified RING-like domains, it has the same functions as the human RBBP6 RING finger domain.

Table 2.1: Different types of RING finger-like domains, outlining their locations, functions, and their structural configuration.

Domain name	Structure	Location	Functions	References
FYVE		Found in human proteome and genome.	Involved in vesicle transport, cytoskeleton regulation and signal transduction.	Stenmark and Aasland., 1999
PHD		Found in inhibitor of growth 1 protein at the C-terminal.	Involved in cell cycle regulation, stress signalling, apoptosis, DNA damage repair and senescence.	Coscoy and Ganem, 2003
LIM		Found in LIM homeodomains	Involved in cell lineage specification, cytoskeleton organization and organ development	Dawid *et al*, 1998
U-BOX		Found in yeast and human proteins	Traces of it in ubiquitination. Uses hydrogen bonds for structural stability.	Hatakeyama and Nakayama, 2003

Moreover, due to different configurations of RING finger domain, apart from its canonical structure, there are other “modified” RING finger domains which give rise to the following various subclasses:

2.4.6.2. C3HC4 classical RING finger domain

The C3HC4 classical RING finger domain is also known as the RING-HC (Figure 2.8), a canonical subclass of the RING finger domain (Lian *et al.*, 2009). This domain differs from other domains in that histidine is at position 4 while the other domains have it at position 5 (Lian *et al.*, 2009). The C3HC4-type RING finger genes are abundant in the plant kingdom which contributes to the vast biological functions in the physiological processes of plant life. This was confirmed by the study of C3HC4 in *Arabidopsis,* in which genes were found to possess different functions in cellular processing and signalling transduction (Agarwa and Khurana., 2018). C3HC4 acts as an E3 ubiquitin ligase which plays a role in the downstream ubiquitination and degradation of targeted photomorphogenesis-promoting transcription factors (Chen *et al.,* 2018). OsRHC24 and OsRHC1 are C3HC4 genes that are related to the disease resistance mechanism (Agarwa and Khurana., 2018)

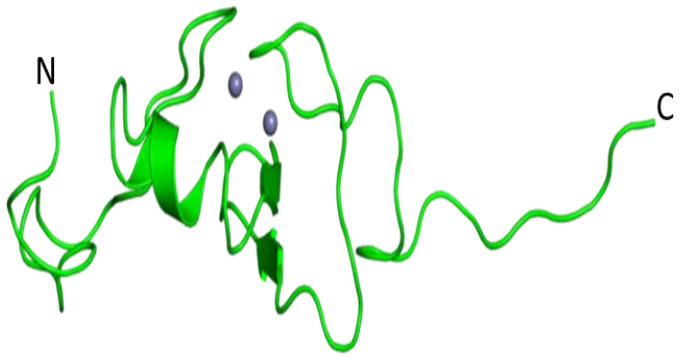


Figure 2.8: A typical C3HC4 RING-type domain. It is composed of two β-strands, one helix and two grey spheres representing zinc ions that is used for structural stability. Taken from (Kellenberger *et al*., 2005).

2.4.6.3. C3HHC3 RING finger domain

This is also known as RING-H2 and the expression of this protein is induced by *N-acetylchitooligosaccharides* in the rice cell (Katoh *et al*., 2003). RING-H2 is provided by EL5, which is structurally related to the *Arabidopsis ATL* protein family (Koiwai *et al.,* 2007). This protein is characterised by a transmembrane domain (I), a basic domain (II), a conserved domain (III) and a RING-H2 finger domain (IV), followed by the C-terminal region that is composed of highly diverse amino acid sequences (Katoh *et al.*, 2003). The solution structure of the *EL5* RING-H2 is stabilized by the tetrahedral coordination of zinc ions (Katoh *et al.*, 2003). The difference between RING-H2 and other RINGs is the replacement of Cys4 by Histidine (Kappo *et al.*, 2009). Katoh and co-workers (2003) determined the structure of the RING-H2 type of RING finger *EL5* by heteronuclear NMR and it was shown to be composed of β-strands and α-helix in a ββα fold (β1, Ala147-Phe149; β2, Gly156-His188; α1, Cys161-Leu166). It was also shown to have a long flexible loop on each side of the ββα structure (N-loop, Val133-Glu146; C-loop, Gly167-Val180) (Katoh *et al.*, 2003). Ubiquitination studies done *in vitro* have also shown that RING-H2 finger domain can bind to E2 (Xie *et al*., 1999; Sun *et al*., 2019).

2.4.6.4. C2H2C4 RING domain

This domain is an example of Hdm2 (Human domain minute 2), which is the human homologue of Mdm2. Hdm2 is a ubiquitin protein ligase which is responsible for the suppression of the tumour suppressor p53 transcription activity and is simultaneously an inducer of its degradation (Kappo, 2009). The solution structure of this domain has shown a novel cross-brace zinc binding topology and a symmetrical dimer. Kostic and co-workers (2006) showed Hdm2 is the only protein that has C2H2C4 zinc coordination. C2H2C4 is composed of the ββαβ fold, a small hydrophobic core and two zinc ions. The

two zinc ions in the C2H2C4 domain structure are important for the maintenance and stability of the structure (Kostic *et al.,* 2006). As earlier mentioned, the C2H2C4 is the novel rearrangement of zinc-binding sites that Hdm2 possesses, in which the tight binding of the two provide E3 ubiquitin ligase activity, thereby creating the isopeptide bond between a lysine of the target protein and a C-terminal glycine of ubiquitin (Metzger *et al.*, 2012).

2.4.6.5. C4C4-type RING finger domain

The C4C4-type RING finger domain is a modified type of RING finger domain that is also known as RING-C2 (Yu *et al.*, 2011). C4C4 domain was identified from *Arabidopsis thaliana* along with other domain types (Yu *et al.*, 2011). Yu and co-workers (2011) showed the analysis of amino acid sequence of the VpRFP1, a protein from grapevines, contains a nuclear localization signal at the N-terminal and the RING finger at the C-terminal. The RING finger on the C-terminal is from the novel RING finger C4C4 with the consensus sequence of Cys-X_2-Cys-X_{13}-Cys-X_1-Cys-X_4-Cys-X_2-Cys-X_{10}-Cys-X_2-Cys, where the X represents any amino acids (Yu *et al.*, 2011). The C4C4 is also very similar to the C3HC4, with only one cysteine residue replacing histidine as shown in Figure 2.9. The C4C4-type RING domain structure has a βββα fold and the chemical nature of this domain is said to be identical, according to Houben and co-workers (2005). The C4C4-type RING domain has significantly different dynamic properties, raising the assumption that this difference might be related to the stability of the domain and perhaps, the functionality of protein-protein interactions that they are involved in (Houben *et al.*, 2005).

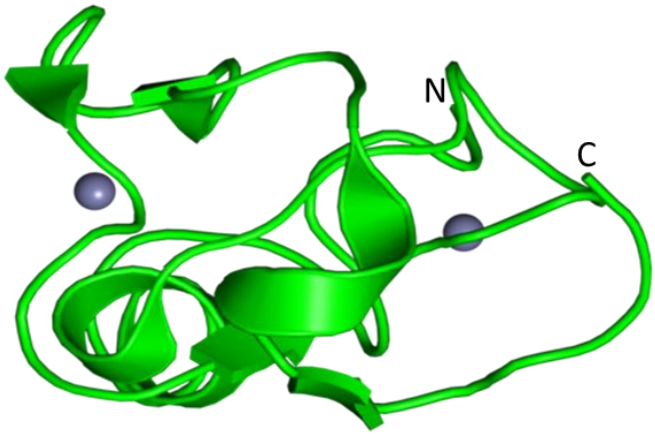


Figure 2.9: The C4C4 RING finger domain. The structure is composed of two β-strands, two α-helices, two 3_{10} helices and two grey spheres representing zinc ions binding. Taken from (Kobashigawa *et al*., 2011).

2.5. Problem Statement

A genome-wide bioinformatics analysis by Kappo and co-workers (2012), using the human RING finger domain from RBBP6 as a searching bait, showed the zinc ion coordination pattern was the same in all but three organisms: *Saccharomyces cerevisiae* and *Pichia pastoris,* where the first binding-site of zinc has been lost completely or weakly bound, and in *Aspergillus niger* where the second coordinating signature cysteine residue has been substituted by an aspartic acid (Figure 2.10). However, it is unclear whether the coordinating zinc ion was lost or was weakly bound. Structural characterization of the loss of the zinc first binding site will add to the body of knowledge of the biological role played by the *Sc*RING finger domains either in ubiquitination, multimerization or in protein quality control. Hence, since the structure of a protein reveals the function of that protein, it is imperative to characterise the structure of the *S. cerevisiae* RING finger domain. Biophysical characterization of this protein will provide the biological basis of high throughput expression and purification strategy to obtain appropriate amount of S. *cerevisiae* RING finger domain protein. It is imperative to ascertain the optical properties of the protein before embarking on the acquisition of structural characterisations. This is because methods employed are mostly based on the light-absorbing properties of the protein.

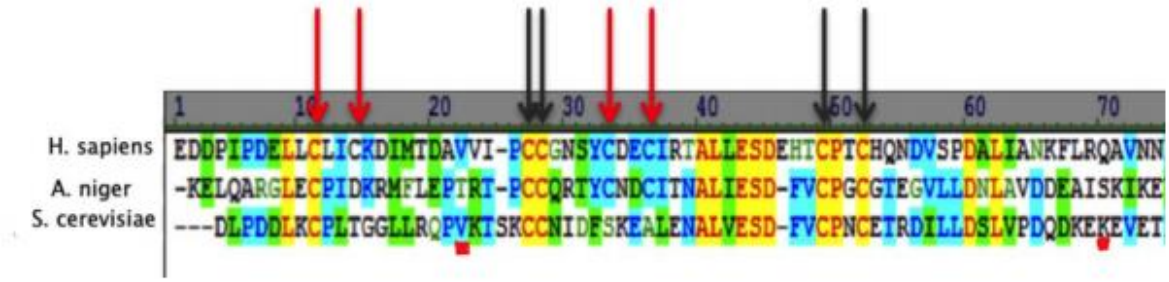


Figure 2.10: Sequential alignment showing Cross-brace configuration for Zn2+ binding.
Sequence analysis of RING finger domains from RBBP6 homologs reveals that *H. sapiens* RING finger binds two Zn2+ions, in which the 1st and 3rd pair of cysteine residues (red arrows) bind one Zn2+ and the 2nd and 4th pairs (black arrows) the other. The alignment reveals a disruption of the first pair of zinc binding site (red arrows) in the *A. niger* RING and *Sc*RING homologs. The alignment reveals that some of the residues coordinating the first Zn2+ ion (red arrows) are missing from *A. niger* and *S. cerevisiae*, raising the possibility that both of these orthologues may bind only one Zn2+. Taken from (Hassem, 2014).

2.6. Aim and Objectives

2.6.1. Aim

The aim of the study was to express, purify and structurally characterise *S. cerevisiae* RING finger domain to postulate the possible biological significance in the loss of the first zinc binding-site. This aim will be achieved via the following objectives:

2.6.2. Objectives

- Overexpress single labelled soluble protein of *S. cerevisiae* RING finger domain in *E. coli*
- Determine *S. cerevisiae* RING finger domain secondary structure using CD
- Determine the folding and oligomeric state of the *S. cerevisiae* RING finger domain (whether monomeric or dimeric in solution)
- Determine the tertiary structure using Tryptophan-UV spectroscopy
- Determine the quaternary structure using SE-HPLC
- Determine the hydrophobic pockets using ANS binding dye
- Bioinformatic structural characterisation of RING-type inhibitors.

CHAPTER 3

MATERIALS AND METHODS

3.0 MATERIALS AND METHODS

3.1 MATERIALS

This chapter outlines the reagents used, the preparation of the stock solutions and the way different methods were used to achieve the results obtained.

3.1.1 General stock solutions, buffers and media (in no particular order)

30% Glycerol: 30ml of glycerol was mixed with 70ml of dH_2O to make up a 100ml of 30% glycerol solution.

Luria Broth: 10.0g Tryptone powder, 5.0g Yeast extract, 5.0g NaCl and 2.0g Glucose were mixed in 1000ml of dH_2O.

0.1M Calcium Chloride: 1.1099g was dissolved in 100ml of dH_2O.

0.1M Magnesium Chloride: 2.033g was dissolved in 100ml of dH_2O.

PBS: 137mM NaCl, 27mM KCl, 4.3mM $Na_2HPO_4.7H_2O$ and 1.4 mM KH_2PO_4 were dissolved in 1 litre of deionized water and autoclaved.

Ampicillin: In 10ml of deionized water, 1.0g of ampicillin was dissolved, filter-sterilized and stored at 4°C until needed.

Agarose gel: 2.0g of Agarose powder was dissolved in 200ml of TAE buffer, boiled and allowed to cool to 55 ºC before adding Ethidium bromide.

10X TAE buffer: 48.4g of Tris, 20ml of 0.5M EDTA and 11.42ml of glacial acetic acid were mixed and made up to 1000ml with deionized water.

IPTG: 1.0g of IPTG was dissolved in 4.2ml of dH_2O and filter-sterilized and then stored at -20°C.

10% SDS: 10.0g of SDS was made to solution by dissolving it in 100ml of deionized water.

SDS Running buffer (pH 8.3): 15.0g of Tris, 72.0g of glycine and 0.5g of SDS were made up to 1000ml with deionized water. The pH was checked but not adjusted.

2X SDS Sample buffer: 4.0ml of 10% SDS, 2.5ml of 0.5M Tris and 1.5ml Glycerol were added to 10ml of deionized water and stored at 4°C.

Destaining Buffer: 100ml of glacial acetic acid and 300ml of ethanol were mixed and topped up with 600ml of deionized water (dH_2O) to make up a 1000ml solution.

Staining Buffer: 0.25g of Coomassie Blue R250, 40% of methanol and 10% Acetic acid was made up to a litre solution with deionized water.

Zinc sulphate: 3.22g of zinc sulphate was dissolved in 200ml of dH_2O, covered with foil and stored in 4°C.

30% Bis-acrylamide solution: 150.0g of Acrylamide and N,N-Methylene-bisacrylamide were mixed together with 300ml of dH_2O. The solution was lightly heated at 37°C until a true solution was formed; thereafter it was topped up with deionized water to make 1000ml, filter-sterilized and stored at 4°C covered in foil.

1.5M Tris (pH 8.8): 36.3g of Tris powder was dissolved in 200ml of dH_2O and pH adjusted to 8.8.

0.5M Tris (pH 6.8): 12.11g of Tris powder was dissolved in 200ml of dH_2O and pH adjusted to 6.8.

Lysis buffer: 100µg/ml Lysozyme, 1% Triton X-100, 1mM PMSF, 50µM $ZnSO_4$ and 1mM DTT were made up to 20ml with 1X PBS buffer.

PMSF stock solution: 1.74g of PMSF was dissolved in 100ml of ethanol and stored at -80°C.

DTT stock solution: 3.9g of DTT was dissolved in sodium acetate, filter-sterilized and stored at -80°C.

Lysozyme stock solution: 1g of Lysozyme powder was dissolved in 10ml of deionized water, filter-sterilized and stored at -80°C.

10% Ammonium sulphate: 0.1g of $(NH_4)_2SO4$ was dissolved in 900µl of dH_2O, filter-sterilized and stored at -80°C.

Equilibration: 1X PBS. The solution was thereafter kept at 4ºC until needed.

Wash buffer: 1X PBS and 1% Triton™ X100. The solution was kept at 4ºC.

Elution buffer: 50mM Tris containing 150mM NaCl pH 8 was used to dissolve 15mM glutathione. The solution was kept at 4ºC.

70% Ethanol: 70ml of ethanol was dissolved in 30ml of deionized water and kept at room temperature.

Separating buffer: 36.3g of Tris was dissolved in 200ml of deionized water and adjusted to pH 6.8, giving a solution with a final concentration of 1.5M.

Stacking buffer: 12.1g of Tris was dissolved in 200ml of deionized water and adjusted to pH 8, giving a solution with a final concentration of 0.5M.

NMR buffer: A 100mM Sodium phosphate buffer at pH 6 containing 50mM NaCl, 5mM DTT and 0.2% of NaN_3.

Sodium azide: 10g of NaN_3 in 100ml dH_2O. Stock solution stored in room temperature.

Chloramphenicol: 0.034g of chloramphenicol in 1ml. Stock solution was stored in -20ºC.

Glutathione: 0.23g of glutathione in 50ml of elution for 15mM glutathione.

Sodium acetate: 4.2g of sodium acetate into 300ml dH_2O. Adjust to pH 4.5 using acetic acid and top up to 500ml to make 0.1M acetate buffer. Solution was stored at room temperature.

Sodium borate: 3.82g of sodium borate into 300ml dH_2O. Adjust to pH 8.5 using boric acid and top up to 500ml to make 0.1M borate buffer. Solution was stored at room temperature.

Cleavage buffer: 50mM Tris-HCl pH 7.0 containing 150mM NaCl, 1mM DTT. Prepared fresh daily. Stock solution was at 4ºC.

Glycine-NaOH: 6.01g of glycine and 2.05g NaOH was added to 800ml. Adjust to pH 10 using HCl or NaOH and top up to 1L to make 80Mm. Stock solution was stored in room temperature.

SE-HPLC buffer: 1X PBS containing 1mM DTT. The solution was gas filtered and stored in room temperature.

ANS: 1.5g of ANS powder was added to 5ml of 1X PBS buffer to make 1mM stock. The solution was stored in -20ºC.

Urea: 3g of Urea powder was added to 50ml of sodium phosphate buffer to make 8M stock. The solution was stored in -20ºC.

3.1.2 Bacterial strains

a. T7 express competent *Escherichia coli: fhuA2 lacZ::T7 gene1 [lon] ompT gal sulA11 R(mcr-73::miniTn10--TetS)2 [dcm] R(zgb-210::Tn10--TetS) endA1 D(mcrC-mrr)114::IS10.*

3.2 METHODOLOGY

3.2.1 Preparation of T7 express competent *Escherichia coli* competent cells

A 5 ml Lb broth in 5 test tubes was inoculated with cells using a sterile inoculation loop and incubated overnight at 37 °C. The following day, one 5 ml culture, that grew best compared to the others, was diluted in a 50ml Lb broth in a flask and incubated at 37 °C until it reached an optical density between 0.3-0.6 Å. The cells were then harvested by centrifugation at 5000 rpm for 10 mins in 50 ml Eppendorf tubes. Harvested cells were resuspended in 5 ml of 0.1 M $MgCl_2$, lightly agitated, and kept on ice for 30 mins. After 30 mins, cells were harvested again by centrifugation, the supernatant was discarded and the pellet was again resuspended in 0.1 M $CaCl_2$, lightly agitated, and kept for 20 mins on ice and again harvested. The pellet was resuspended in 0.1 M $MgCl_2$ for 4 hrs with occasional swirling while kept on ice. Lastly, the suspension was centrifuged to harvest the cells, and 30% glycerol and 0.1 M $MgCl_2$ was added to the pellet, lightly vortexed and kept at -80 °C in storage for future use.

3.2.2 Transformation

100 µl of *E. coli T7* cells were thawed and mixed with 2 µl of *Sc*RING finger domain Amp^R plasmid in a 2 ml Eppendorf microtube and labelled “test” and the other microtube with cells only was “labelled” control. Two set of the experiment was prepared the first one included Ampicillin (100 mg/ml) + chloramphenicol (34 mg/ml) and the second set was with Ampicillin (100 µg/ml) + chloramphenicol (34 µg/ml). Both sets were incubated on ice for 20 mins and immediately exposed to heat in 42 °C heating-block for 45 sec, then returned to ice for 10 mins incubation. 900 µl Lb broth was added to both sets and further incubated for 2 hrs at 37 °C in a shaking incubator. After incubation, 100 µl of cells from both set were plated in ampicillin and chloramphenicol containing Lb agar

plates and the remaining 900 µl was centrifuged at 15000 rpm for 5 mins to harvest the cells. The supernatant was discarded, and the remaining cells were lightly vortexed and plated on ampicillin/chloramphenicol containing Lb agar plates. All eight plates were incubated overnight at 37 °C. From the overnight incubated plates on the second set there was colony contamination, therefore the first set with Ampicillin (100 mg/ml) + chloramphenicol (34 mg/ml) was used to continue with the experiment.

3.2.3 Small-scale expression screening

Fresh colonies were randomly picked from the plates incubated overnight in section 3.2.2. The colonies were inoculated into 4 tubes containing 5 ml of Lb broth with ampicillin/chloramphenicol and incubated for 4 hrs in a shaking incubator at 37 °C. After the incubation period, 2 ml of culture from each 5 ml of cell culture were divided into two Eppendorf microtubes to create two set of cultures from each 5 ml culture. Thereafter, 1 ml from each 5 ml culture was labelled induced and 0.5 Mm of IPTG was added, and the other 1 ml was labelled un-induced and no IPTG was added. Both sets of cultures from each 5 ml culture were incubated for 2 hrs in 37 °C in shaking incubator. The cells from both set of cultures were then harvested by centrifugation and the pellet collected. The pellets were run in a 15% SDS-PAGE with an addition of 2x sample buffer for sample preparation. The SDS-PAGE protein expression profile permitted a selection of the best recombinant protein expressed culture from the 3 ml culture previous set apart to set-up an overnight cell culture at 37 °C for large-scale recombinant protein expression.

3.2.4. Large-scale protein expression

A 100 µl of cell culture, from the method mentioned above, was inoculated in 100 ml of Lb broth containing ampicillin were incubated with shaking overnight at 37 °C.

Chloramphenicol was added to maintain the copy number, however at this stage chloramphenicol was excluded to provide a conducive environment for protein synthesis. The overnight culture was diluted in a 4 L Lb broth divided into 200 ml flasks with ampicillin at a concentration of 100µg/ml and further incubated at 37 °C while periodically monitoring OD until it reached 0.4. Subsequently, a concentration of 1 mM IPTG was added as well as a filter-sterilised 0.5 Mm of $ZnSO_4$ prior to induction. The cell culture was further incubated 6 hrs at 30 °C in a shaking incubator. After incubation, the cells were harvested by centrifugation at 6000 rpm for 15 mins and the pellets were stored at -80 °C for future use.

3.2.5. Extraction and affinity purification of *Sc*RING finger protein

The cell pellets previously stored at -80 °C were thawed on ice. After the cell pellets had thawed, they were re-suspended in a protein lysis buffer by vigorous vortexing. To break the cell wall, the cell pellets were sonicated for 10 min at 5 sec intervals for each Eppendorf tube. The sonicated cell mixture was then centrifuged at 6000 rpm for 20 mins at 4 °C to harvest the total bacterial cell lysate, clear of cell debris. The clear total bacterial lysate was collected into fresh 50 ml Eppendorf tubes. Following this, an affinity purification column was prepared using an overnight swollen glutathione-agarose bead at 4 °C, which was later poured into a plastic column, permitting buffer to flow-through. Thereafter, the packed affinity column was equilibrated with 3 CV of PBS at pH 7.4. After the equilibration buffer had been drained off, the clear total bacterial lysate, containing the GST~*Sc*RING domain protein, was poured into the column and allowed to flow by gravitational force. The flow-through was collected and stored on ice. Thereafter, the affinity column was washed with 5 CV of PBS at pH 7.4. The last three drops of the wash buffer, before the column dried up, was collected and used to

ascertain the purity of the GST-column before eluting the GST-tag *Sc*RING domain protein. A 3 CV protein elution buffer containing 15 mM Glutathione was used to elute the fusion GST-*Sc*RING domain protein off the GST-agarose affinity column. The different fractions from the affinity purification process were electrophoresed on a 15% SDS-PAGE.

3.2.6. SDS-PAGE

A short glass was placed on a spacer plate, while using the green casting frame to hold the plates on a flat surface by clamping the plates. A freshly prepared separating gel was added between the two glasses to about 5 cm before the top. Thereafter, 10% isopropanol was added to help level the top of the gel and allowed to solidify for some time. In the time the separating gel was solidifying, a stacking gel was prepared, which was later added to the top of the separating gel and a comb was inserted into the stacking gel to create wells. The gel was allowed to solidify for about 10 mins. After the gel had solidified, the comb was removed and the gel, which was still inside the casting glasses, was placed in an electrophoresis tank containing 1X SDS-PAGE running buffer, covering the entire gel cast system, before the protein samples were loaded. The first well of the SDS-PAGE was loaded with 2 µl of a pre-mixed protein ladder, while the other wells were loaded with pre-boiled protein fractions. The protein samples were prepared by mixing 20 µl of the protein samples and 10 µl of 2X sample buffer. The mixture was boiled at 95 °C, before 20 µl of the boiled mixture was loaded into the wells. The gel was then run for 1 hr at 120 V. After the dye front of the gel was about to be released into the running buffer, the process was stopped, and gels were carefully taken out of the tank and removed from the caster before they were transferred into a container. An appropriate volume of Coomassie brilliant blue solution was poured onto

the gel and allowed to stain for 60min or overnight. After the gel had been properly stained, a destaining solution was added onto the gel in the container and allowed to shake for as long as possible, until the gel became very clear, and the protein bands could be properly seen.

3.2.7. Far-UV Circular dichroism

The *S. cerevisiae* RING finger protein sample for CD was freshly prepared at 20 ºC before measuring the CD spectra. For accurate results, a 5 µM protein sample, which amounted to 250 µl, was diluted in a filter-sterilized PBS buffer, pH 7.4. The CD Jasco-1500 spectropolarimeter was set using the following parameters: bandwidth 5 nm, data pitch 1 nm, quartz cuvette 1 mm and the wavelength started from 250 nm and ended at 180 nm. The spectral readings accumulated at the average of 3 at a scanning speed of 200 nm/min in triplicate were recorded. Firstly, the CD spectrometer was blanked with PBS buffer in which the protein sample was dissolved. This was to ascertain that the buffer components do not yield excessive noise or any other unnecessary defects in the spectra. Following this, O_2 from the lamp housing and sample compartments in the CD spectrometer was removed, by flushing with N_2 gas. This was done to keep the measurements below 200 nm and prevent the formation of ozone and damage of the optical system. Furthermore, the absorbance results obtained from the CD experiment were analysed and expressed as mean residue molar ellipticity [ΘMRE] using this equation ($\Theta = 100 \times \theta / c \times l \times n$), where θ is for measured ellipticity (mdeg), *c* is protein concentration (mM), *n* for the number of protein amino acids and *l* for pathlength (mm). Lastly, the same method was used for denatured protein, where 8M urea was used as a denaturing agent to the previous prepared native protein sample. Beta Structure Selection (BeStSel) is a method that uses CD spectrum for secondary structure

elements determined by quantitative analysis and fold recognition using CATH classification (Micsonai *et al*., 2018). This is a speedy and reliable structure analysis especially when x-ray crystallography and NMR spectroscopy is not available (Micsonai *et al*., 2018).

3.2.8. Intrinsic Tryptophan Fluorescence

Intrinsic tryptophan fluorescence is one of the important instruments in biochemical research because of its highly sensitive, robust and non-invasive analysis of proteins (Amar *et al*., 2014). It makes use of fluorophores (aromatic residues), that are readily available in native proteins, which absorb light in different wavelengths (Amar *et al*., 2014). This is done to study the dynamics and conformational changes of the protein tertiary structure (Lewkowicz, 2014). A beam with a specific wavelength passes through a protein sample in a cuvette; thereafter, both the excitation spectrum (light emitted) and the emission spectrum (light absorbed) by the sample are measured from the angle (Lewkowicz, 2014). A Jasco FP-6300 spectrofluorometer was used for fluorescence measurements at a temperature of 20 ºC for accuracy. Parameters were set, with excitation and emission bandwidth of 5 mm, and 1mm size cuvettes were used. Protein samples were prepared in triplicate for both native and denatured samples. Native samples were prepared as 5 µM in PBS buffer pH 7.4 while 8 M urea was added for denatured samples. All spectra were measured and recorded at an average of 3 accumulations in a scanning speed of 200 nm/min. A graph was plotted for analysis.

3.2.9. Size Exclusion High Pressure Liquid Chromatography

Size exclusion high pressure liquid chromatography (SE-HPLC), better known as gel-filtration chromatography, is a vital tool used in science to characterize and separate proteins according to different sizes, hence estimation of protein size (Tayyab *et al*.,

2002). A sieving medium, used to separate proteins, is made of a porous gel which may either be Sephadex, Sephacryl, Bio-gel, controlled pore glass beads or Sepharose, where smaller molecules have the highest probability to penetrate compared to the bigger molecules (Tayyab *et al*., 2002). SE-HPLC can also be used for large scale separation, however due to time consuming and expensive media required to fill proteins, this method is usually applied as the final step of protein purification (Tayyab *et al*., 2002).

A 20 µl *Sc*RING protein sample concentrate (20 µM) was injected into a Shimadzu UFLC SPD-20A SE-HPLC system in a TSK gel super SW2000 size exclusion column (resolution of 5-150 kDa) for analysis, which was limited to the separation of protein within this range. The protein absorbance was monitored at 280nm with a flow rate of 0.25ml/min. The SE-HPLC 1X PBS buffer was used as a mobile phase which contained 500mM and 1mM DTT at pH 7. The recorded absorbance was used to calculate the estimated size and concentration of the *Sc*RING protein, as well as ascertaining the purity of the protein.

3.2.10. 8-Anilino-1-naphthalene sulfonate (ANS) Fluorescence

ANS is a widely used fluorescent dye for protein characterization of binding pockets. It has been trusted to bind strongly through ion pair formation in cationic proteins and poly-amino acids (Gasymov and Glasgow, 2007). Although ANS dye can bind external binding sites (those exposed to the aqueous phase) in proteins, the high intensity of the ANS fluorescence spectrum is observed in internal binding sites (hydrophobic phase), buried in the core of the protein, which indicates the restriction of the ANS dye during binding (Gasymov and Glasgow, 2007). ANS ability to bind both external and internal

binding sites has made it a vital tool in science for the analysis of different ligands, including druggable molecules (Gasymov and Glasgow, 2007). The excitation wavelength of ANS is 390 nm and when not bound, its energy is emitted at 540 nm.

A 200 µM ANS stock solution was prepared by dissolving the powdered dye in 20 mM PBS buffer containing 0.02% sodium azide. The concentration was spectrophotometrically confirmed at 350nm where the extinction coefficient was 4950 M-1 cm-1. A 5 µM *Sc*RING protein sample mixed with 35 µM ANS solution in 2 ml Eppendorf vials was prepared in triplicate and 35 µM ANS alone was used as blank. The Jasco FP-6300 spectrofluorometer was used to measure the fluorescence spectrum of ANS at 20 ºC in a 10 mm path length cuvette. The excitation and emission bandwidth were 5 nm and ANS was excited at 390 nm wavelength. Spectra readings were recorded at an average of 3 accumulation at a scanning speed of 200 nm/min. Lastly, recorded spectra were used to plot a graph for analysis.

CHAPTER 4

RESULTS AND DISCUSSION

4.0. *Sc*RING PROTEIN PRODUCTION

*Sc*RING finger protein is a piece of a fragmented sequence isolated from Mpe-1 that has a characteristic of independent expression and replication. An interest towards this protein was because of the possibility that it might perform the same function as that of a human RING finger protein. Hence, in this study we provide evidence to already known characteristics. This section entails results from the cloning of *Sc*RING gene to pGEX-6P-2, bacterial transformation, recombinant expression, purification, and structural characterisation of *S. cerevisiae* RING finger protein.

4.1. Recombinant expression, purification of GST-tagged *Sc*RING finger protein

4.1.1. pGEX-6P-2*Sc*RING finger domain p53728 cloning

A fusion gene pGEX-6P-2*Sc*RING finger domain p53728, encoding residues 180-197 of the full length Mpe-1 tagged with GST, was constructed at the University of Western Cape (Biotechnology department). The *Sc*RING finger domain sequence was cloned between BamH I and Xho I restriction enzyme sites as shown in Figure 4.1. The construct had a *tac* promoter that was inducible by IPTG in which the 3C prescission protease recognition site (Gln↓Gly) was encoded in a fusion protein as a linker region between GST and *Sc*RING sequence. After the protease had cleaved, some of the amino acids were left attached to the protein and therefore the sequence of the protein was as follows: **GPLS**HFKDLPDDLKCPLTGGLLRQPVKTSKCCNIDFSKEALENAL VESDFVCPNCETRDILLDSLVPDQDKEKEVETFLKKQEELHGS, where the highlighted amino acids were the remainders after cleavage. The construct was then introduced into *E. coli* T7 cells to incorporate it with the plasmid. Judging from the work presented on a MSc thesis (Mathenjwa, 2018) on protein expression screening, it was evident that

not enough protein yield was obtained, and it was only obvious to try other types of cells. Hence, T7 express competent *Escherichia coli* cells were chosen. *E. coli T7* cells were used for bacterial transformation because of their known features of being a high efficiency expression system for recombinant proteins compared to other protein expression cells (Fathi-Roudsari *et al*., 2016). Additionally, it was important that the cells we chose were resistant to chloramphenicol as well to avoid culture cells from releasing the plasmid during protein expression, i.e., to choose cells with one copy number. The contamination from the first set of bacterial transformation, it was assumed it might be because of low concentrations of chloramphenicol and ampicillin Figure 4.2 shows the transformants that were able to accept *Sc*RING finger protein Amp^R plasmid through their resistant to ampicillin antibiotic plates and growing into distinct colonies.

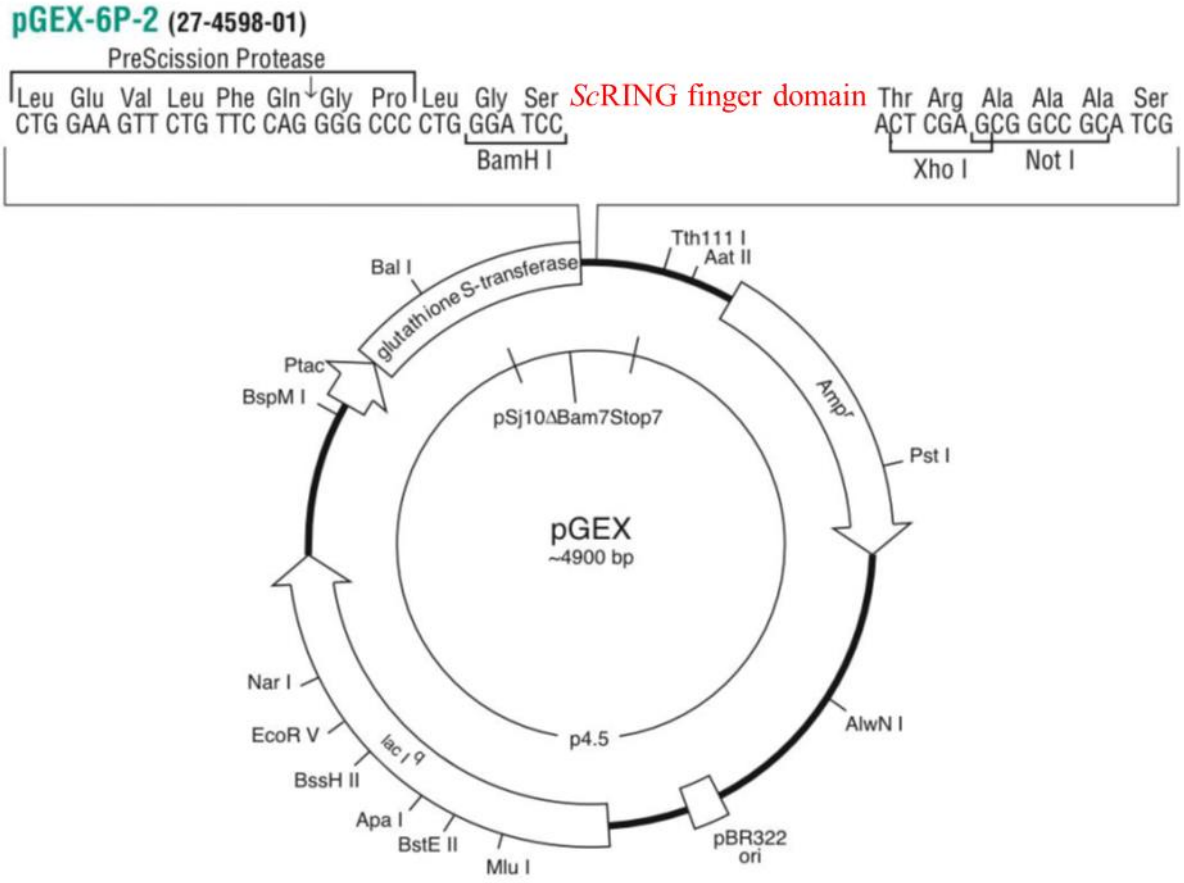


Figure 4.1: A representative diagram illustrating a pGEX-6p-2~*Sc*RING domain construct. *Sc*RING domain insertion was done between BamH I and Xho I sites in a 4900bp C-terminal pGEX-6P-2 GST vector.

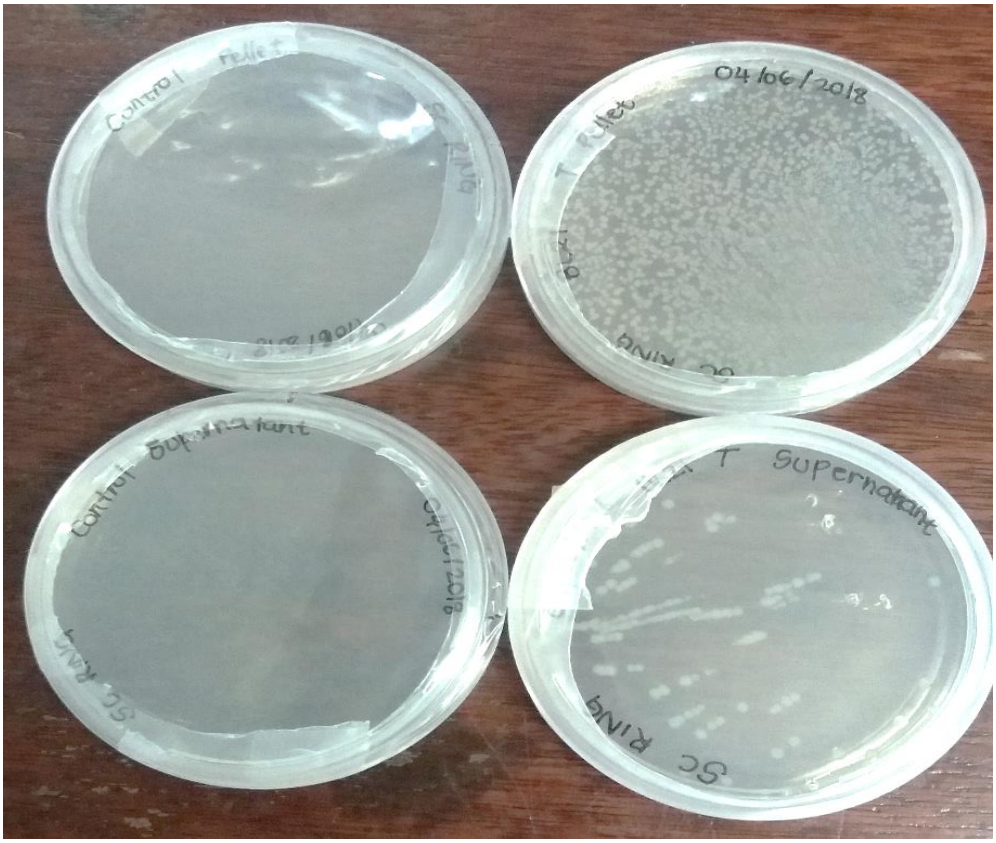


Figure 4.2: **A representative diagram showing the *S. cerevisiae* RING finger protein transformants**. Colonies on the test plates shows that the incorporation of the *Sc*RING plasmid was successful and the plates without colonies were used as control for cells without *Sc*RING plasmid.

4.1.2. *Sc*RING finger protein screening and large-scale expression

Successful transformants were later used for small-scale expression to scan for a colony that was able to produce more protein compared to other colonies. SDS-PAGE gel in Figure 4.3 illustrates the *Sc*RING finger protein bands from the randomly picked transformants showing the difference between the IPTG induced and un-induced colony culture in small-scale expression. Colony in lane 3 was chosen to recombinantly express the protein in large quantities due to the assumption that the band size corresponds to protein yield. Initially, the large-scale protein expression was prepared in a single 4 L flask overnight, however, that resulted in the release of protein when incubated for a long hour. This was then followed by a "cold-drunk protein expression" whereby the cell culture was expressed in a 4 L Lb broth overnight at 4 °C with the addition of 2% ethanol. The idea was that difficult protein produce better protein yield under cold condition, however, this experiment failed as well, and it was assumed that the cells were less competent at lower temperatures and possibly ethanol was not an ideal solvent to this process as it added to the incompetency of cells. Therefore, optimization methods were made, where instead of a single 4 L flask protein expression; the cell culture was expressed in 200 ml broth in 500 ml flasks for 6 hrs instead of 24 hrs at 30 °C instead of 37 °C and 4 °C. This was done to prevent the formation of inclusion bodies that arise from either a high concentration of an inducer or from a strong promoter that causes high protein translation rate (Singh *et al*., 2015). Additionally, 30 °C was a temperature that was neither too cold nor too warm for the T7 bacteria cells to produce enough protein. A bacterial lysate was recovered which was a pool mixture of different proteins and results in Figure 4.4 exhibits the large-scale expression of *Sc*RING finger protein and the band appeared to be twice the size as the one shown in small-scale expression. This proved that the optimization method was a

success. A GST tag is approximately 26 kDa which makes the total molecular weight of the *Sc*RING finger fusion protein to be 37 kDa. Hence, *Sc*RING finger protein on the SDS-PAGE gel ran in pairs at 37 kDa, the top band was a GST-tag and below was the protein of interest.

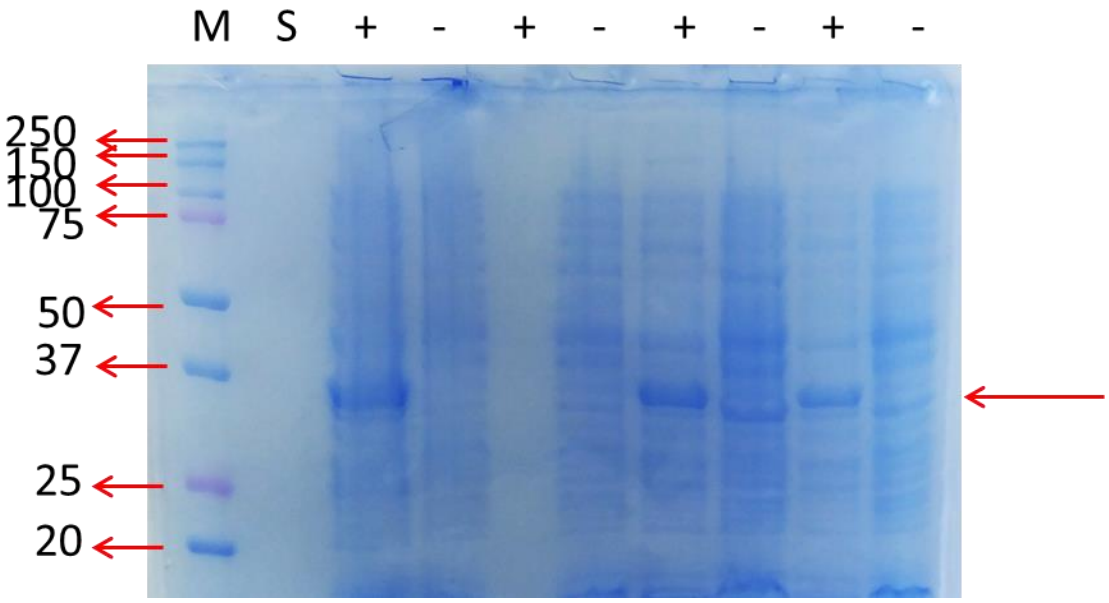


Figure 4.3: **A representative diagram showing expression screening of the induced and un-induced transformants.** The cells were incubated at 37 °C in a 5 ml broth after inducing them with 1 mM IPTG. Lane 1 is a protein marker, lane 2 is the supernant recovered after harvesting the cells. Lane 3 shows the induced colonies and lane 4 is the un-induced; the same thing applies on the following lanes.

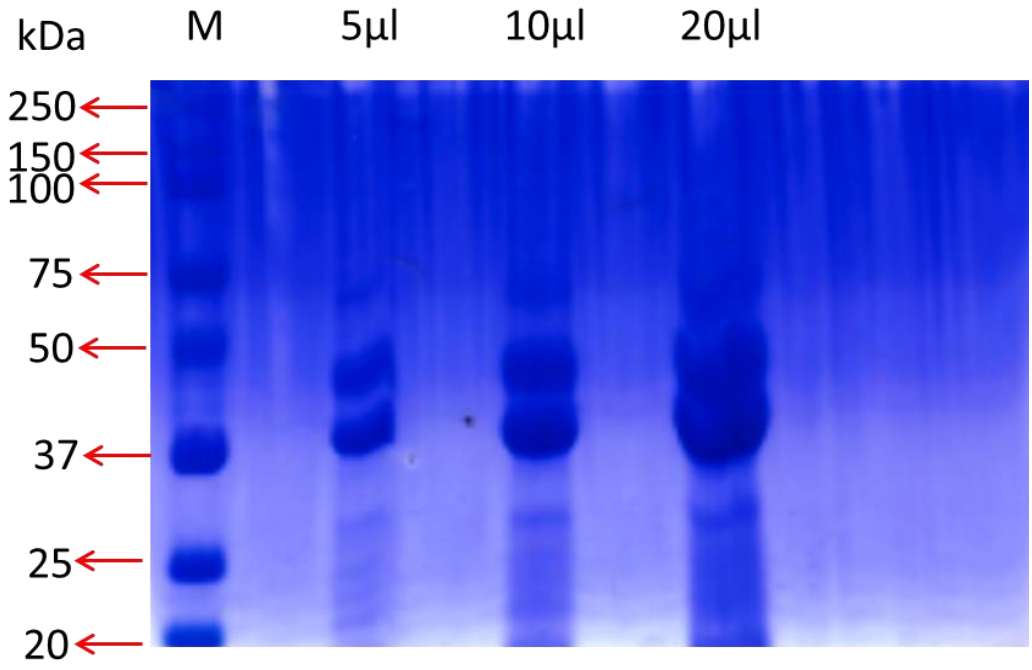


Figure 4.4: A diagram showing a Large-scale recombinant expression of *S. cerevisiae* RING finger protein. Lane 1 is a protein marker, lane 2 is a 5 µl of sample, lane 3 is a 10 µl of sample and lane 3 is a 20 µl of sample. Judging from the size of the band, the protein shows production increase in large scale.

4.1.3. *Sc*RING finger protein cleavage and purification

To separate the protein of interest from the protein cocktail, a glutathione sepharose affinity column that has the higher affinity to GST was used to isolate the *Sc*RING finger GST-fusion protein. At this point the protein was not usable for structural characterisation because it was still fused with a GST-tag. Therefore, Figure 4.6 illustrates the results obtained after introducing an HRV 3C prescission protease to cleave off the tag after 4 °C incubation overnight; a free GST-tag was running at 26 kDa and *Sc*RING protein was running at 10 kDa. A distinct distance between the two protein bands in Figure 4.6, was an indicative of the perfectly cleaved protein. HRV 3C prescission protease was chosen because of its high cleaving specificity, compared to thrombin and factor Xa that might risk the fragmentation of the target protein as well. To separate the GST-tag and *Sc*RING finger protein in a solution, the mixture was returned to the column and a GST-free protein was recovered as shown in Figure 4.7 (A and B) where SDS-PAGE confirmed the separation of the GST-tag and the protein of interest. The purified protein was stored in 4 °C for later use in protein characterization methods.

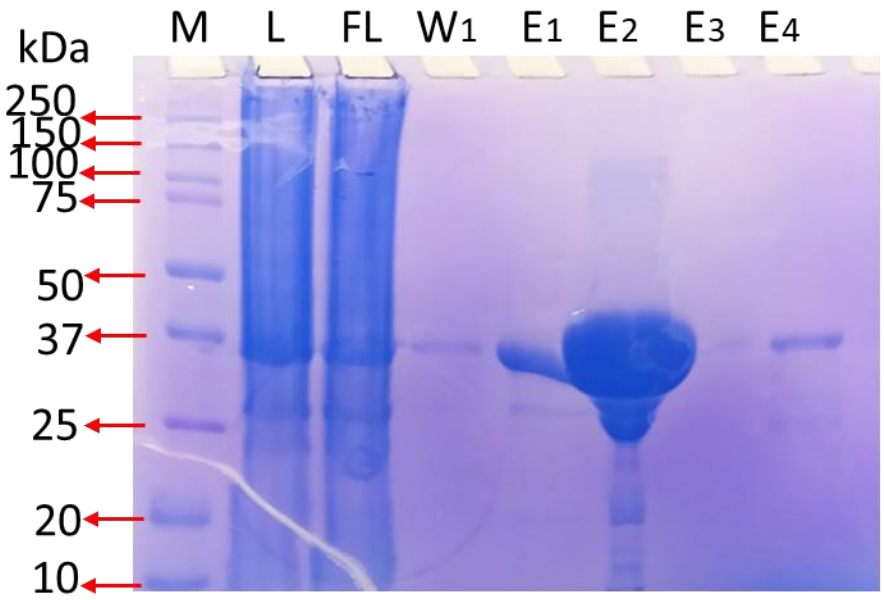


Figure 4.5: **A diagram showing a column affinity purification of *S. cerevisae* RING finger protein.** Lane 1 is a protein marker, lane 2 is a bacterial lysate, lane 3 is a flow-through, lane 4 is wash 1, lane 5-8 are protein eluates.

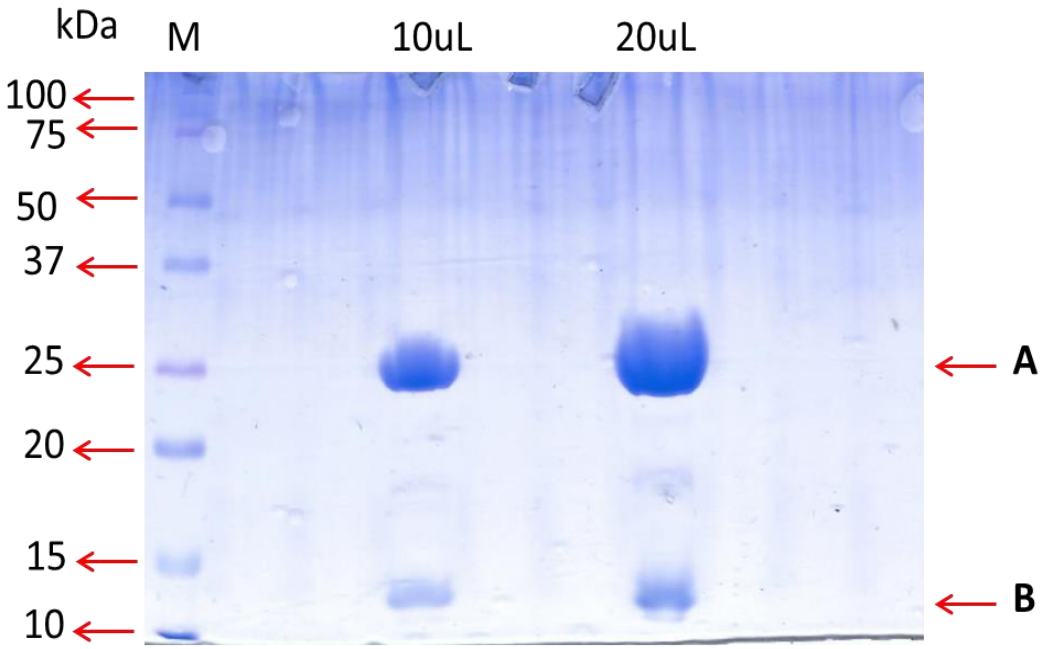


Figure 4.6: **A diagram showing cleavage of a GST-tag from the *S. cerevisae* RING finger domain.** In lane 1 is a marker and lane 2-3 are the eluates of the *S. cerevisiae* cleaved protein. To cleave the protein Prescission protease was used.

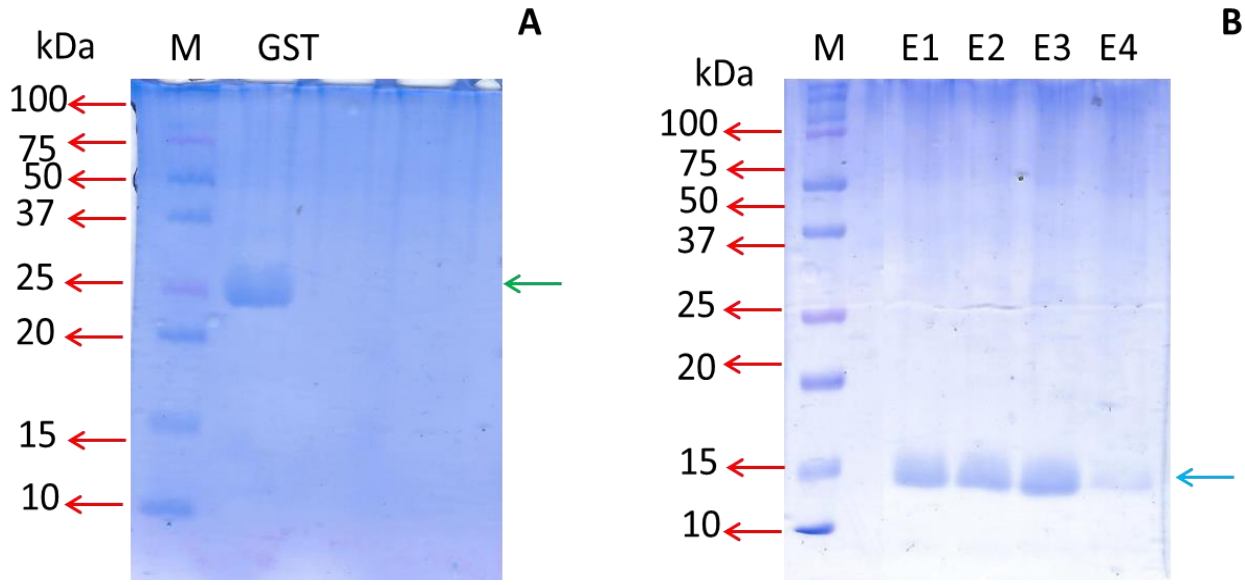


Figure 4.7: **Diagrams showing the complete separation of GST-tag (A) and ScRING protein (B).** In diagram (A) a green arrow shows a floating GST-tag running at 25 kDa and diagram (B) a blue arrow shows a pure *Sc*RING protein running at approximately 10 kDa.

4.2. *Sc*RING finger protein structural characterisations

According to previous structural predictions done on the ExPASy PROTOPARAM online server (Mathenjwa, 2018), it was postulated that the *Sc*RING finger domain lacks the hallmark aromatic residues (i.e tyrosine and tryptophan), and these residues are known for their high fluorescence intensity (Ghisaidoobe and Chung, 2014). Therefore, even though phenylalanine has the lowest quantum yield, the proteins still depend on it for absorbance and emission of light. *Sc*RING finger domain is predominated by cysteine residues which also play a role as a chromophore, provided it is reduced (Middendorf *et al*., 2000). Hence, these *Sc*RING finger protein structure characterisations demonstrated low fluorescence intensity absorbance but distinctive information for structural characterisation was deduced.

4.2.1 Proteins Concentration determination

Figure 4.8 is the first imperative characterisation of the protein that uses a NanoDrop as a tool to determine protein concentration, which was 0.15 mg/ml, equivalent to 20 µM. This was unfortunate after the optimization methods that resulted in a bigger protein band size in large-scale expression SDS-PAGE as shown in Figure 4.5. It was then assumed that perhaps the optimization method favoured the expression of GST rather than the protein of interest and due to tedious purification step required for GST-tag protein, it was possible that lot of protein was lost. However, for structural characterisation that were to be done in this study, this was acceptable due to that the required protein concentration range required was between 0.005 mg/ml to 5 mg/ml. *Sc*RING finger protein was 260/280 ratio of 1.39 which indicated that the protein might be lightly contaminated by nucleic acids (RNA and DNA) however, that did not have any effect for further characterisation of the protein (Dejardins and Conklin, 2010) as the

acceptable ratio for a pure protein should vary between 1.8-2.1. The protein concentration calculations were essential to perform structural characterisations.

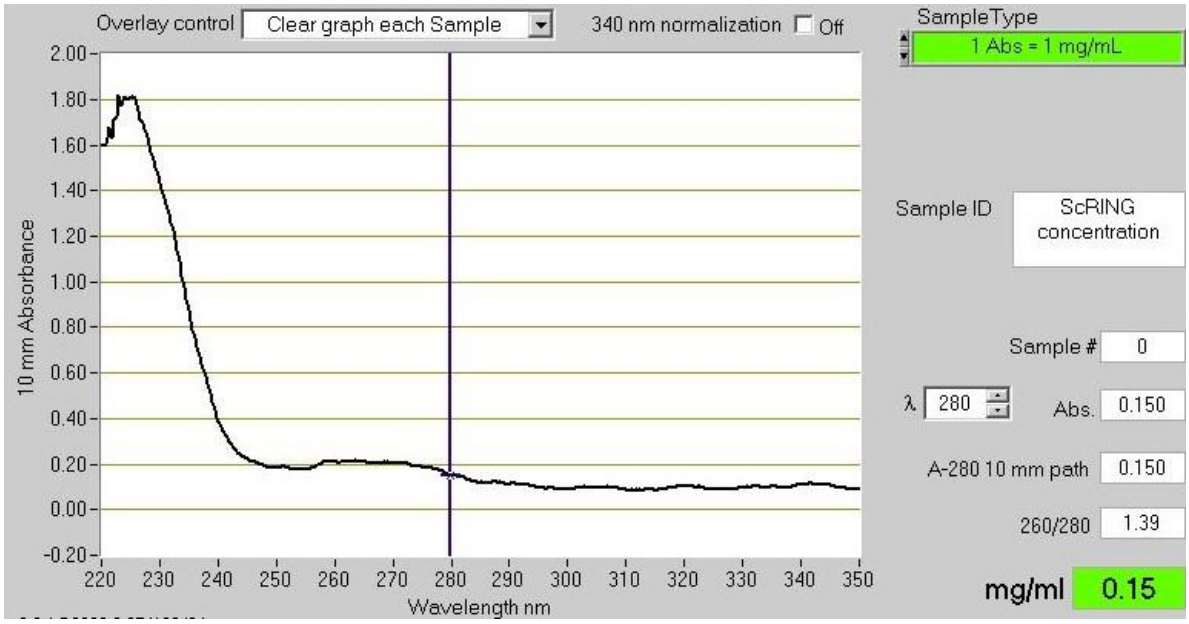


Figure 4.8: Diagrammatic presentation of a *Sc*RING finger protein concentration. Protein concentration was 0.15 mg/ml which is equivalent to 20 µM. Peak absorbance is observed between 260 and 280 nm.

4.2.2. Secondary structure determination using Far-UV Circular Dichroism

Secondary structure elements were determined using Far-UV CD spectrometry. According to the results obtained in Figure 4.9, the comparison between the denatured protein (unfolded) and native protein (folded) showed a significant difference in that the denatured *Sc*RING finger protein exposed more β-strands observed at 195 nm, while the native *Sc*RING finger protein exposes more of the α-helices at around 208-222 nm (Greenfield, 2006). During the denaturation of the protein, through the addition of 8 M urea as a denaturant, it is believed that it migrates towards the core of the protein where it forms hydrogen bonds, resulting in the linear protein structure or change in configuration that inversely affects the orientation of the amino acids, hence the dis-positioning of the secondary structure elements (Dunbar *et al*., 1997). Chemical denaturing of the protein destroyed all the phases of the protein structure except for the primary structure of the protein, hence the *Sc*RING finger protein switches from globular configuration to linear peptide.

Now, as aforementioned, shifting the focus to the denatured state of the *Sc*RING finger protein, more β-sheets and β-turns structures were exposed. This is an indication of *Sc*RING finger protein stability. β-sheets form a hydrophobic core of *Sc*RING finger protein and contribute about 92.6%, which is divided into antiparallel (31.2%) and parallel (61.2%). The antiparallel β-sheets are subdivided into twisted (15.2) and relaxed β-strands, and the remaining 7.4% is occupied by β-turns. β-sheets are formed through the joining of β-strand by hydrogen bonds with an extended repeat of two residues per unit making a 7° pitch (Perczel *et al*., 2005). It has been noted that proteins with abundant β-sheets are less likely to proteolyse and the biomolecular recognition, such as DNA base pairing attached at the edges of hydrogen-bonding between β-strands,

prevent protein aggregation, enhance protein-protein interaction, and contribute towards the formation of the quaternary structure. Hence, cancer and protein dysfunction-related diseases are prevented (Nowick, 2008). On that note, β-turns form the most important part of the protein for the formation of a quaternary structure and functionality (Lewis *et al*., 1973). They are formed during the overall reverse direction at approximately 180° of a polypeptide chain and, unlike the β-sheets and α-helices that comprise ordered and repetitive structure, β-turns only have ordered structures that are classified based on the torsion angles (phi and psi) (Lewis *et at*., 1973). They play a crucial role in the flexibility of the peptide chain to bend during the formation of secondary and quaternary structure and the more β-turns presence, the more stable the protein structure. Hence, the presence of β-turns in the *Sc*RING finger protein confirms the ability of the protein to fold and its stability.

Helical structures were observed as a secondary element in *Sc*RING protein. The α-helices are formed by a 3.6 residue/turn which is defined as a length of 36 amino acids required to form 10 turns. This is a repetitive structure caused by the repetition of torsion angles (psi and phi) that are formed when carbonyl groups are attached to amino groups to form hydrogen bonds. Formation of hydrogen bonds in this manner results in the formation of helical structures which contribute to the molecular stability of the protein. It should be noted that not all amino acids are responsible for the formation of α-helices but only methionine, alanine, lysine, leucine, and glutamate. The *Sc*RING protein amino acid composition includes alanine (2.4%), lysine (10.7%), leucine (14.3%), glutamate (10.7%) and lacks methionine. In total these amino acids make up 38.1%, which amounts to 32 amino acids, resulting in an incomplete formation of an α-

helix. It can then be assumed that the missing residues required to form a complete α-helix are the ones that were expected to form the first zinc binding site.

T

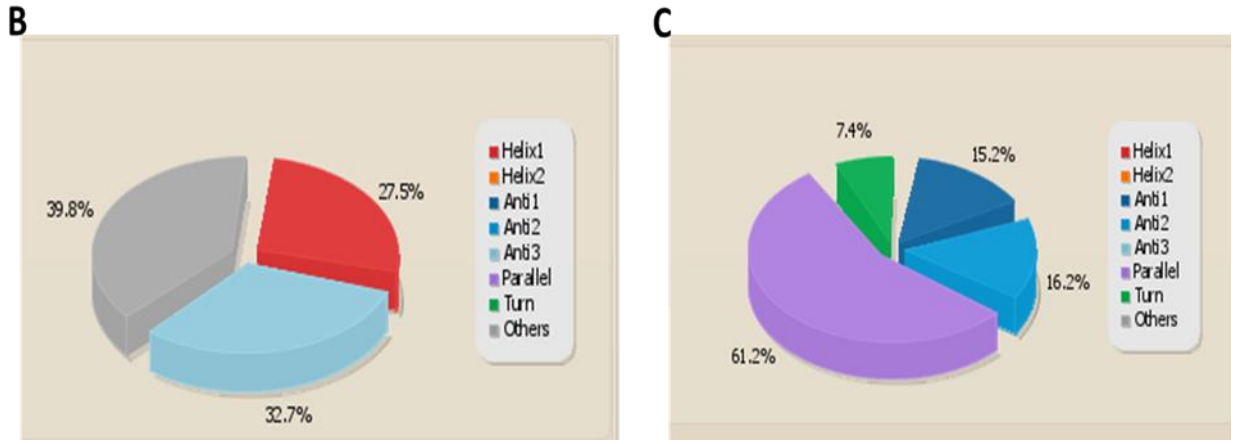


Figure 4.9: A diagram showing Far-UV circular dichroism of *Sc*RING finger protein. Negative bands of α-helices were observed 208-222 nm and positive bands around 193 nm in both native (red) and denatured (purple) *Sc*RING finger protein **(A)**. BeStSel web server was used for secondary structure elements quantification for native **(B)** and denatured **(C)** *Sc*RING protein. From the general view of the results, it shows that different secondary elements are exposed in different state of the protein in which native shows more of α-helical structures while denatured shows more of β-strands.

4.2.3. Intrinsic fluorescence emission spectrum

Intrinsic fluorescence emission spectrum tool was used to study *Sc*RING finger protein conformational changes. In this study, the native state of the *Sc*RING finger protein was compared to the denatured state, where 8 M urea was used as the denaturing agent. Significant difference was observed where the denatured *Sc*RING finger protein was shown to have a higher pick compared to the native *Sc*RING finger protein. It was assumed that this may be caused by the exposure of amino acids that absorb light, which may have been buried in the core during the native state of the protein. This indicated the conformational charges of the *Sc*RING finger protein during denaturation. In diagnostics, this tool could therefore be useful to study protein behaviour when denatured which may be a lead to protein aggregation resulting in various diseases, including cancer. According to Jpred4 server for secondary structure prediction shown in Figure 4.11, in conjunction with intrinsic fluorescence in Figure 4.10, it can be deduced that before *Sc*RING finger protein denaturation, two phenylalanine residues at position 32 and 74 are buried in the core of the protein, hence a lower emission wavelength was observed (Drozdetskiy *et al*., 2015). In contrast, with the denatured *Sc*RING finger protein, all the phenylalanine residues exposed yielding a higher emission wavelength.

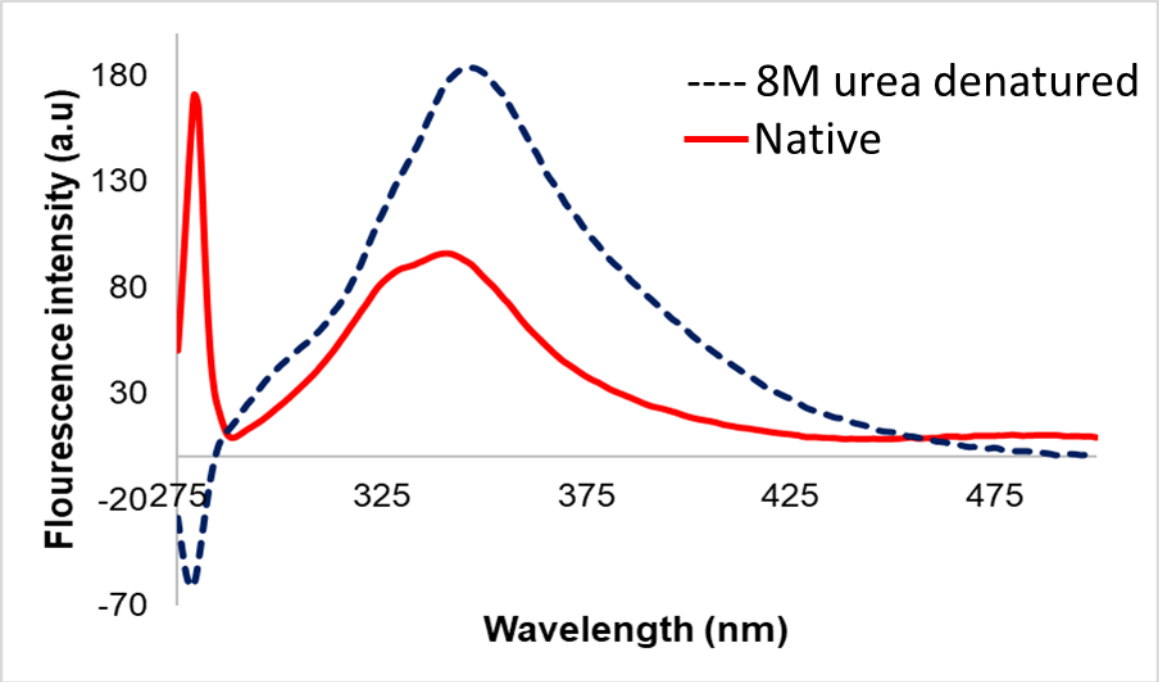


Figure 4.10: A diagram showing intrinsic fluorescence emission spectrum. A 5 µM native *Sc*RING finger protein (red) sample was prepared in PBS pH 7.4 and 8M urea was used as the denaturing agent to unfold the proteins samples (dark blue). Fluorescence spectra was recorded at 275 nm excitation wavelength at 20 ºC.

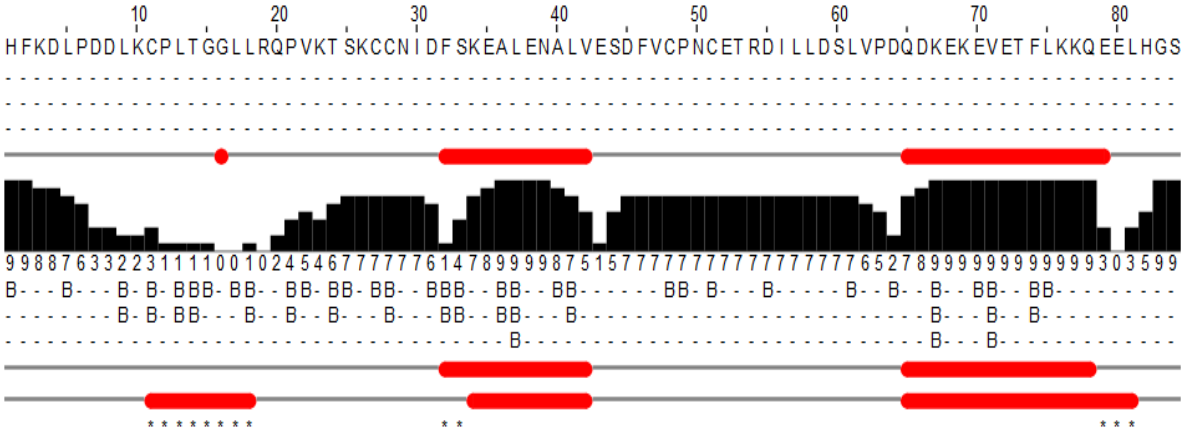
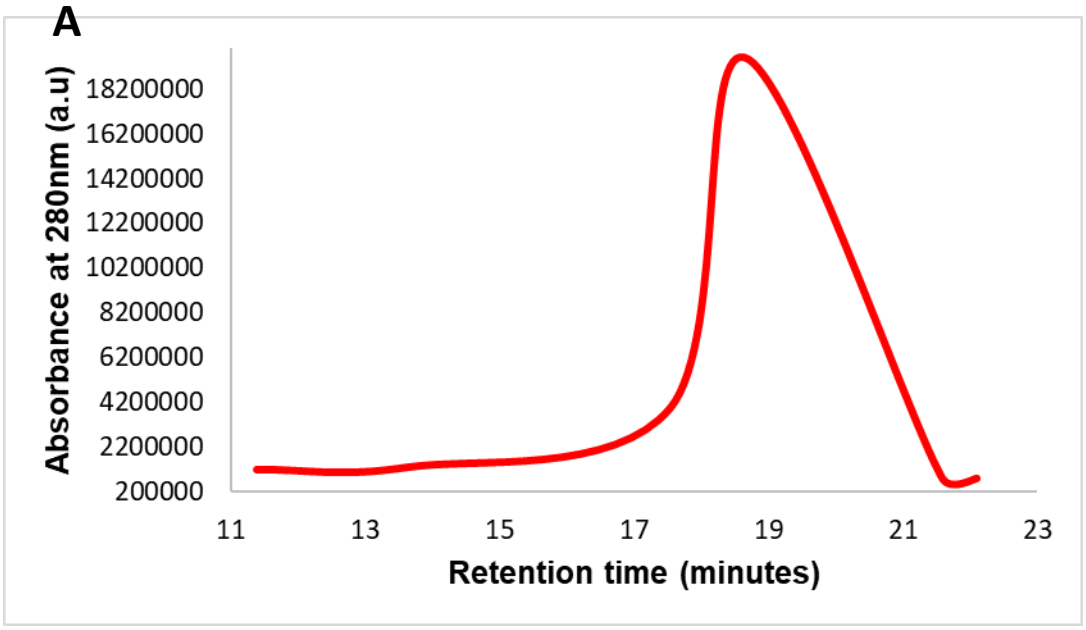


Figure 4.11: Secondary structure prediction. The diagram shows the predicted amino acids composition of *Sc*RING finger protein, where **"B"** represents the buried amino acids and **"-"** representing the exposed amino acids. The number on the top of the abbreviated amino acids (see appendix A) represents the amino acids position numbers.

4.2.4. Size exclusion high performance liquid chromatography

Determination of the tertiary structure elements of *Sc*RING finger protein brought us to an interest in also understanding the quaternary structure of the protein; that is whether we are looking at the monomeric, dimeric, trimeric etc. subunit conformation of the protein. Reflecting on Figure 4.12 (A), the retention time of the *Sc*RING finger protein peak started from 17.5-21.5 mins which was approximately an elution of about 5 mins. It is important for E3 ligases to be homo-dimeric for their activity, hence, structural studies on isolated RBBP6 RING finger protein proves that it is a weak homodimer. Therefore, results obtained from SE-HPLC confirmed that the ScRING finger protein as a possible weak homodimer. The other peaks were described because of a high salts solutes in a buffer solution. The highest retention time of the *Sc*RING finger protein peak of 18.6 mins is almost the same as Myoglobin at 18.9 mins, as shown in a standard calibration curve in Figure 4.12 (B) in which the *Sc*RING finger protein had a molecular weight of 10 kDa and Myoglobin at 17 kDa. Observation of these results gave confidence regarding the molecular weight of the *Sc*RING finger protein, taking note of their light difference in retention time, which might be a result of nonideal adsorption in a matrix that tends to affect the accuracy of the molecular weight. According to literature Hong *et al*. (2012) stated that high concentrations of salt in the SE-HPLC buffer may be one of the solutions to prevent nonideal adsorption of proteins. However, a high close-up evaluation of the protein behaviour in a high concentration of salts is needed to prevent aggregation. Hence, 750 mM NaCl2 was added to the SE-HPLC buffer, coupled with 10 mM NaH2PO4 at pH 7.4, as an optimization step towards protein size determination accuracy and quaternary structure determination.



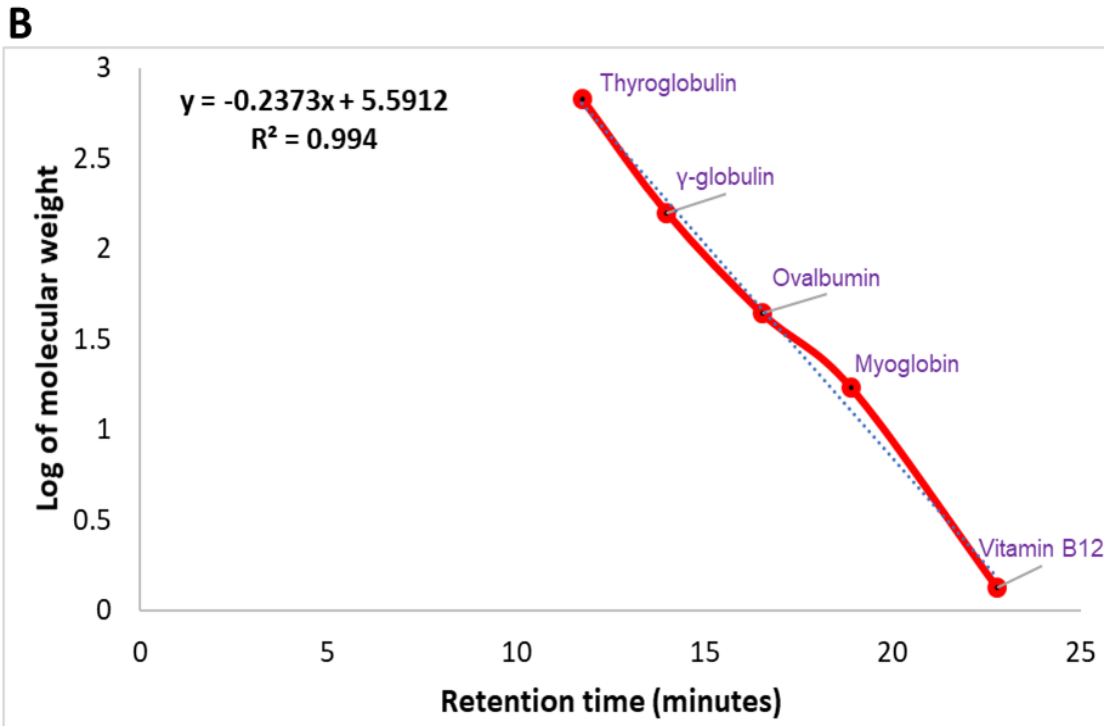


Figure 4.12: **A diagram showing a standard curve (B) and a chromatogram of *Sc*RING protein in size exclusion high performance liquid chromatography (A).** A 10 mM sodium phosphate pH 4.7 containing 750 mM NaCl was used as the mobile phase for the elution profile for proteins at 280 nm absorbance using a Shimadzu HPLC system. The retention time for *Sc*RING protein was 18.6 mins at 10 kDa with the flowrate of 0.25 ml/min. **(B)** Calibration curve was plotted using the following standards: Vitamin B12 (1.35 kDa), myoglobin (17 kDa), ovalbumin (44 kDa), γ-globulin (154 kDa) and thyroglobulin (670 kDa); the correlation coefficient was 0.994.

4.2.5 *8*-Anilino-*1*-naphthalene sulfonate (ANS) Fluorescence

One of the characteristics of a *Sc*RING finger protein is a protein to ligand interactions and to achieve this, a protein needs to have hydrophobic pockets known as binding sites. Therefore, ANS fluorescence was used to determine the presence of hydrophobic pockets for *Sc*RING finger protein. ANS dye is known to bind positively charged amino acid side chains, namely arginine and lysine. Histidine also has the positive side chain even though it has a weak affinity for H^+ and only in neutral pH. ANS binding is through electrostatic interaction forces of the protein in which the excess of the ANS leads to protein aggregation, while moderate concentrations stabilise the protein.

Secondary structure predictions in Figure 4.11 show two amino acids, arginine and lysine at position 1 and 69 respectively, buried in the core of the protein, while the other 9 amino acids are exposed to solvent. The sulfonate group in an ANS dye binds to positively charged side chains of arginine, lysine and histidine which reduce the intermolecular charged-electron transfer rate leading to increased ANS fluorescence. In Figure 4.13, a comparison between the denatured and native state of the *Sc*RING finger protein ANS binding shows a denatured *Sc*RING finger protein to have a higher ANS fluorescence compared to the native *Sc*RING finger protein. This observation could be explained through literature where it is said that higher concentrations of positively charged side chains are directly proportional to the ANS florescence. Therefore, is safe to assume that higher fluorescence in denatured *Sc*RING finger protein were as the result of exposed arginine and lysine residues that are usually buried in native state of the protein. However, out of the three; arginine has the highest ANS fluorescence, followed by lysine then histidine. This is because of the differences in their side chain pK values, which are 12.5, 10.5 and 6.0 respectively. In contrast, although higher ANS

fluorescence maybe be desirable, however too much of it maybe be an indicative of protein aggregation which is why in this study higher values were observed in denatured *Sc*RING finger protein.

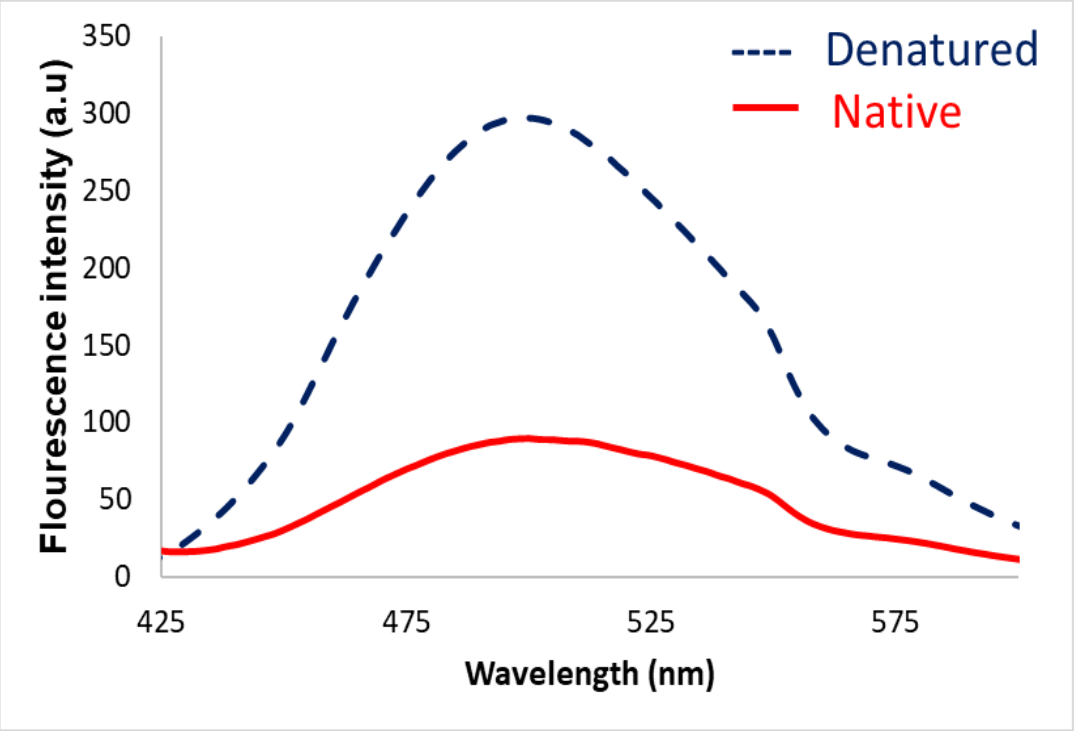


Figure 4.13. Diagram illustration of ANS binding fluorescence spectra. *Sc*RING finger protein samples were prepared in the presence of 35 µM ANS dye and PBS buffer for native (red) in which 8 M urea was added for denatured sample (dark blue). The excitation wavelength was 390 nm.

4.3. Conclusion

An interest in the *Sc*RING finger protein was triggered by the possibility that it might have the same functions as the human RING finger domain. Bioinformatic studies conducted on this protein have shown that the protein lacks two major aromatic residues, namely tyrosine and tryptophan. This leaves phenylalanine and reduced cysteine residues for the absorption and emission of light. The concentration of the *Sc*RING finger protein was 0.5 mg/ml and it was acceptable for structural characterisations conducted in this study. However, for future purpose that would include advanced structural methods such as NMR and/or crystallography, a minimum of 1 mg/ml would be required. Far-UV CD spectroscopy indicated that the secondary structure of the *Sc*RING finger protein was dominated by β-sheets, which contributes to the rigidity, stability, and protein-protein interaction enhancement. The denatured *Sc*RING finger protein compared to native was shown to have more exposed β-sheets, assumed to have been buried in the core during the native state. However, in future it would be interesting to also conduct the same study in different conditions that include variation in temperatures and during the present and absence of zinc metal ions. This would bring more insight in terms of *Sc*RING finger protein behavioural conformational changes. Furthermore, a same trend was observed in the intrinsic fluorescence emission spectrum between native and denatured *Sc*RING finger protein, in which the denatured *Sc*RING finger protein had higher light absorption compared to the native *Sc*RING finger protein. This may also be due to the additional exposure of light absorbing residues. JPred4 server prediction indicated phenylalanine as buried residues during the native state of the *Sc*RING finger protein, which is why when the *Sc*RING protein structure opens up during denaturation more light is absorbed. The validity of the result reported here could also be confirmed through other methods such

NMR and crystallography, in which it was unfortunate that it could not be possible due to costs and the pandemic (COVID-19). The SE-HPLC chromatography showed that the *Sc*RING finger protein was a possible weak homodimer subunit tertiary structure with the molecular weight of 10 kDa when traced through Myoglobin on a standard curve. Lastly, it was of high importance to determine the binding sites of the *Sc*RING finger protein to confirm if the protein has the interaction properties as known in the human RING finger protein, which was an important knowledge needed in drug design. Therefore, in ANS fluorescence result it was interesting to notice that the native *Sc*RING finger protein had a lower binding affinity compared to the denatured *Sc*RING finger protein. This may indicate the severity of a denatured *Sc*RING finger protein, as it may be a major contributing factor in the protein aggregation leading to tumorigenesis. However, lower binding affinity of the *Sc*RING finger protein could be advantageous to maintain protein balance and protein-protein interactions, especially during ubiquitination. Thus far, the characteristics of the *Sc*RING finger protein seem to be almost the same as that of the human RING finger protein, which was assumed that it due to the N-terminal identical score of 29% reported above in literature. The next chapter, therefore, enlightens on the ligand binding properties of a RING finger protein containing inhibitors (Mdm2 and MdmX) of p53 tumour suppressor protein, as an extension of protein ligands that can be useful in cancer therapy. The chapter highlights the two inhibitors function to be regulated by RING finger domain (Wienken *et al*., 2016).

CHAPTER 5

RING FINGER DOMAIN CONTAINING INHIBITORS (MDM2 AND MDMX) OF p53: STRUCTURAL CHARACTERISATION USING BIOINFORMATICS

5.0. Introduction

Human RBBP6 RING finger domain activity is mainly associated with degradation of either mutated protein or dysfunctional cells that later form a tumour because of the coagulation of hydrophobic patches exposed on the surface, hence leading to cancer ((Chasapis *et al.,* 2010; Dominguez, 2004)). It is for this reason that researchers have developed interest in further studies on the malignant tumour biological formation with the aim to lead them towards the prevention of tumorigenesis.

Murine double minute 2 (Mdm2) is a RING finger containing inhibitor protein structurally composed of 6 motifs that contributes to its functionality, namely p53 binding domain, nuclear localization signal, nuclear export signal, zinc finger domain followed by nucleolar localization signal and a RING finger domain which is responsible for all Mdm2 activities (Lwakuma and Lozano, 2003). It was discovered initially as a protein that promotes cell transformation and later showed that increased levels of Mdm2 in human sarcomas was due to Mdm2 inhibitory tendencies on p53, a tumour suppressor protein, via protein-protein interaction (Cahilly-Snyder *et al*., 1987 and Fakharzadeh *et al*., 1991). Mdm2 has an ability to ubiquitinate itself and facilitate a monomeric type of ubiquitination for p53, which is a function adopted from RING finger domain that is essential for inhibitory effects. Nevertheless, Mdm2 can act jointly with p300/CBP for p53 polyubiquitination, preventing its activity, although recent studies argued that interaction of p300/CBP with SIRT7 stabilizes p53 protein activity (Lu *et al*., 2020; Lwakuma and Lozano, 2003). In contrast, although an over-expression of Mdm2 may lead to cell tumorigenesis, a complete deletion has been experimentally shown to be lethal, resulting in haematopoiesis defects and body mass decrease in mice (Oliner *et al*., 2016; Tackmann and Zhang, 2017). This is an indication of the crucial role of Mdm2

as a titrator of p53 to maintain a counterpoise between Mdm2 and p53 (Tackmann and Zhang, 2017). It is also important to note that Mdm2 activity is incomplete without the presence of Mdm4, better known as MdmX (Venkatesh *et al*., 2020). MdmX also presents inhibitory activity for p53 but not similarly as Mdm2 (Venkatesh *et al*., 2020). Besides its increased levels in human tumour cell lines and glioblastomas, it also plays an essential role in cell proliferation (Qi and Cobrinik, 2017). Although Mdm2 inhibitors can bind to MdmX, due to the structural differences in the binding site of p53 protein, the Mdm2 inhibitors have a low affinity for MdmX, resulting in them being incompetent in cancerous cells whereby MdmX is regulated (Sanz *et al*., 2019). The interaction and stability of MdmX to Mdm2 is mediated by RING finger domain in which MdmX prevents Mdm2 degradation and p53 translocation (Wienken *et al*., 2016). The complete deletion or mutation of a RING finger domain in Mdm2 has the same effect as the complete deletion of Mdm2 (Wienken *et al*., 2016). In contrast, it is important to note that Mdm2 and MdmX are not the inhibitors of RING finger domain, however, they are RING finger containing inhibitors of p53. MdmX is not necessarily ascribed to E3 ligase activity in ubiquitination, rather its interaction with Mdm2 inhibits p53 activity by masking the transcription activation domain (Jiang and Zawacka-Pankau, 2020).

This chapter therefore covers computational studies in which RO-2443 was nominated as a control inhibitor of Mdm2/MdmX E3 RING-type ligases and Hit, formally known as ZINC16046951, as an experimental ligand. This is an interesting approach for cancer therapy resulting in increased levels of p53 activity.

5.1 *In silico* methods

5.1.1. Virtual Screening

Virtual screening was performed using the ZINCPharmer online tool (Koes and Camacho, 2012), whereby RO-2443 compound was loaded as a query to seek compounds with its structural similarity. The screened compounds were filtered with respect to Lipinski's rule of five (i.e Molecular Weight < 500, H-bond acceptor ≤ 10, H-bond donor ≤ 5, Log P ≤ 5) and then subjected to molecular docking.

5.1.2. Preparations of Molecular docking

X-ray crystal structures of Mdm2 and MdmX complexed with RO-2443 ligand were obtained from the RCSB Protein Bank with PDB ID of 3VBG and 3U15 (Graves *et al*., 2012) respectively. Protein structures were separated from the ligand, undesirable co-crystalized molecules deleted, and the protein was inspected for presence of hydrogen atoms using UCSF Chimera software (Pettersen *et al*., 2004). Ligands were prepared using Avogadro software, whereby bond order correctness was inspected, and hydrogen atoms were removed.

5.1.3 Molecular docking

Autodock Vina (Pettersen *et al*., 2004; Trott and Olson, 2010) was used to conduct molecular docking. The grid box parameters for the dimensions were determined as x = 24 Å, y = 28 Å and z = 26 Å; and the centres were recorded as x = -6.803 Å, y = 42.011 Å and z = 2.824 Å. The spacing was = 8 which covered the entire region occupied by the ZINC16046951 ligand in the binding site of the receptors. Different docking conformations were obtained in a descending order of a docking scores using the Lamarckian genetic algorithm (Lazauskas *et al*., 2017). The best docking conformation scores were selected for molecular dynamic simulations. Visualizations of the

complexed structures was with the use of the UCSF chimera. Avogadro software was utilized to prepare ZINC16046951 ligand and receptors (Hanwell *et al*., 2012).

5.1.4. Molecular dynamics simulations

Molecular dynamics simulations set-up was done on an AMBER 18 software package running on a Graphic Processing Unit (GPU) (Song *et al*., 2019). Receptors and ligands were optimised using the AMBER 18 modules, Antechamber and LEaP to ascertain the presence of all the parameters needed for molecular docking simulations. The Amber force field FF14SB was used to define proteins parameters (Kang *et al*., 2018). The systems were thereafter encased in a TIP3P water box with the distance between the protein surface and the box boundary set at 10 Å (Ong and Liow 2019). Energy minimization, heating and equilibration steps were performed, and topology files were generated (Song *et al*., 2019). Molecular dynamics simulations were extended for 100ns. The generated trajectory from each system simulations was analysed per 1ps intervals. Analysis for RMSF, RMSD and RoG were performed using the CPPTRAJ and PTRAJ modules of the AMBER 18 unit. The graphical representation of the results was plotted using the qtGrace data (Winter, 2011) analysis tool and UCSF chimera was used for image visualization.

5.1.5. Binding energy calculations

Molecular mechanics/Poisson-Boltzmann or Generalised born and surface area (MM/PBSA and MM/GBSA) are widely used methods for computation calculations of total binding free energy affinities of ligands to receptors (proteins) (Pandey *et al*., 2018; Xue *et al*., 2019). It provides a sensible enlightenment into the interaction complex of ligand-protein and the isolated entities of these macromolecules. The affinity strength of a ligand bound to a receptor is determined by ΔG_{bind}. MM/PBSA was originally defined

by Massova and Kollman (2000) from the AMBER 14 software package running on a Graphical Processing Unit (GPU) however, in this study an updated version of AMBER 14, AMBER18, was used. The MM/PBSA objective with its complementary MM/GBPSA is to calculate the binding free energy difference between solvated molecules (Genheden and Ryde, 2015). MM/GBSA approach computation is based on the following equations:

$$\Delta G_{bind} = G_{complex} - [G_{receptor} + G_{ligand}]$$

$$\Delta G_{bind} = E_{MM} + G_{sol} - T\Delta S$$

$$E_{MM} = E_{vdw} + E_{ele}$$

$$G_{sol} = G_{polar} + G_{nonpolar}$$

where ΔG_{bind} signify the free energy composition of protein-ligand complex, which is the sum of molecular mechanics energy (E_{MM}), solvation free energy (Gsol) and the entropy ($-T\Delta S$). The E_{ele} and E_{vdw} demonstrate the electrostatic and van der Waals interactions. G_{sol} represents the solvation free energy, which was decomposed to G_{polar} (polar) and $G_{nonpolar}$ (non-polar) contributions. The GB equation is used to determined G_{GB} indicating a polar solvation contribution, whereas the nonpolar contribution state was estimated from the solvent accessible surface area (SASA). *T* and *S* denotes the temperature and total solute entropy, respectively (Ndagi *et al*., 2018). The total binding free energy contributed by each amino acid in the binding site was estimated via a MM/GBSA method using equilibrated trajectories (Zhang *et al*., 2017).

5.2 Results and Discussion

5.2.1 Virtual Screening and Molecular docking

Ligand-based virtual screening resulted in one hit, ZINC16046951 as the only small molecule structurally similar to RO-2443 as shown in Figure 5.1. ZINC16046951 was docked into MdmX and Mdm2, in the RO-2443 binding site, with a docking score of –9.2 kcal/mol and –10.1 kcal/mol, respectively. These docking scores meant ZINC16046951 had more binding affinity for Mdm2 rather than MdmX, which confirms that inhibitors that binds Mdm2 have low binding affinity for MdmX, even though they may have a similar structure (Sanz *et al*., 2019). RO-2443 was re-fitted into its binding site on MdmX and Mdm2 as shown in Figure 5.2 and Figure 5.3; needless of being subjected to molecular docking, hence there were no docking scores for these complexes. Looking at these complexes it is easy to identify them as the same structural complexes because of the high similarities between Mdm2 and MdmX molecular structures. Hence, docking scores and interaction forces used were recorded as a fingerprint to differentiate them.

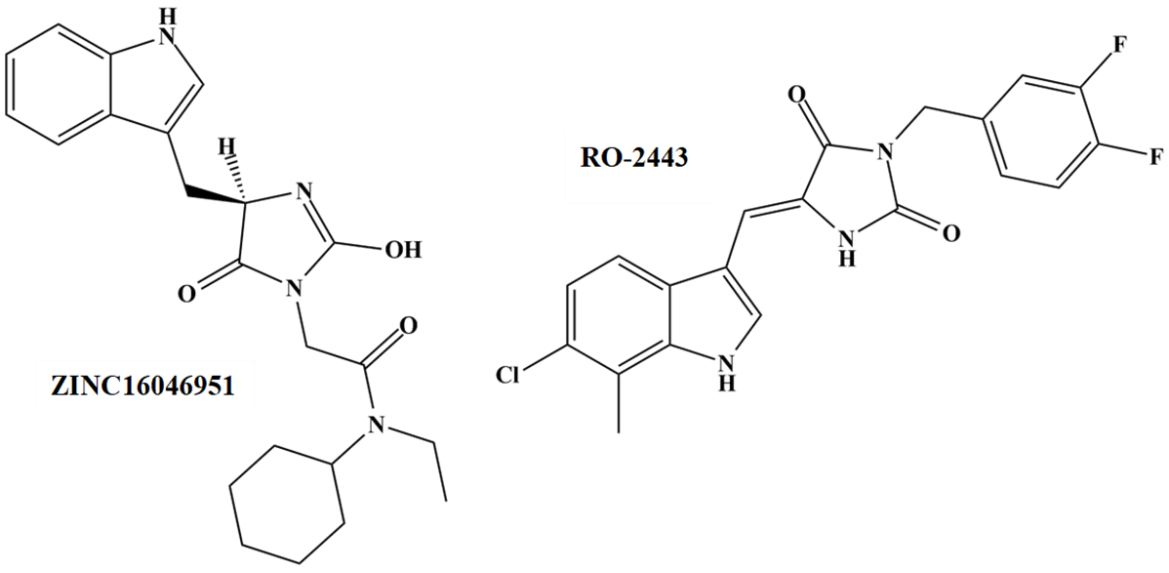


Figure 5.1: A 2D structure of a hit ZINC16046951 and reference compound RO-2443. A ZINC16046951 was the only ligand found to be structurally similar to RO-2443, which was used as the control in this study.

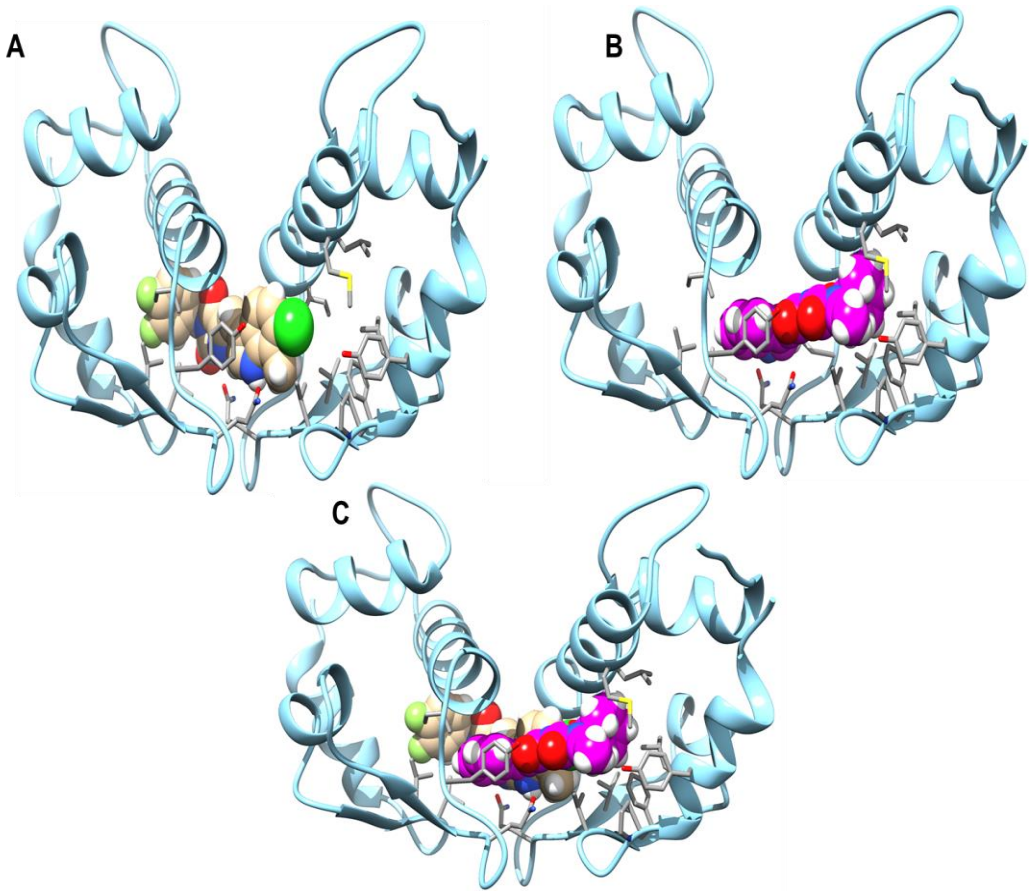


Figure 5.2: Docking poses of ZINC16046951 and RO-2443 in Mdm2 binding site. A RO-2443 ligand is identified by light brown spheres (**A**) while the ZINC16046951 ligand is represented by purple spheres (**B**). Here, the illustration (**C**) confirms that both ligands bind on the same Mdm2 binding site.

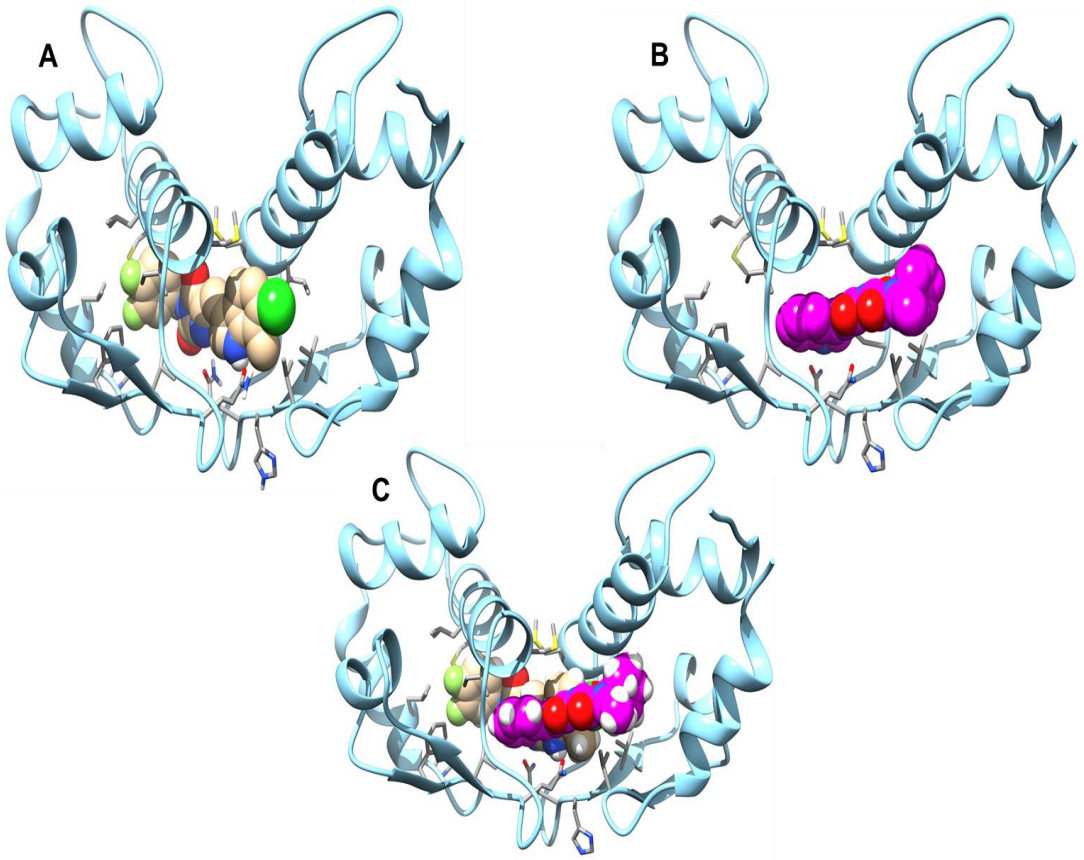


Figure 5.3: Docking poses of ZINC16046951 and RO-2443 in MdmX binding site. A RO-2443 ligand is identified by light brown spheres (**A**) while the ZINC16046951 ligand is represented by purple spheres (**B**). Here, the illustration (**C**) confirms that both ligands bind on the same MdmX binding site.

5.3 Molecular dynamic (MD) simulations

5.3.1 Root Mean Square Deviation (RMSD)

The structural stability of Mdm2 and MdmX in the complex, with the ZINC16046951 ligand and R0-2443 as the control, were computed using RMSD, which measures the distance between the receptor protein atoms and ligands in MD simulations (Kufareva and Abagyan, 2011). Usually, it is the backbone heavy atoms, such as carbon, nitrogen, oxygen and alpha-carbons (Cα) only, that are measured (Damm and Carlson, 2006). According to the graphical representation in Figure 5.1(A), the Mdm2+ZINC16046951 complex shows an accepTable stability at an average of 1.7 Å, starting from the beginning of the MD simulation until a small instability was observed around 32518 ps, as well as between 67962 ps and 81570 ps, which was not more than 2.8 Å. In comparison with the Mdm2+RO-2443 control complex, the system had an average of 1.75 Å, which was slightly different from the experimental complex. It exhibited stability from the beginning of the MD simulations until 35281 ps, while the complex was mostly unsTable up to 3.5 Å; the same pattern was also observed between 88278 ps and 98858 ps. Contrastingly, the RMSD results of the MdmX+ZINC16046951 complex in Figure 5.1(B) showed the system demonstrated stability throughout the MD simulation, at an average of 1.07 Å, with small instability traces observed around 47962 ps but still not above 2.5 Å. This is an indication of rigidity in the complex during the binding of a ZINC16046951 ligand. The RMSD results of the control MdmX+RO-2443 complex was averaged at 1.30 Å, showing accepTable stability from the beginning of the MD simulation but, unlike the MdmX+ZINC16046951 complex, due to the observation of a highest instability around 70435 ps and 83433 ps, which goes up to 3.2 Å. The instabilities of Mdm2+ZINC16046951 and Mdm2+RO-2443 demonstrated were presumably due to conformational changes during the MD simulations. The receptor

protein was assumed to have an ability to open and close during the binding of both ligands, which may have resulted in a change in system complexes atomic distances. Such postulations are supported by the work of Tackmann and Zhang (2017), where it was shown that even though over-expression of Mdm2 leads to tumorigenesis, complete deletion is also lethal, resulting in haematopoiesis, therefore indicating Mdm2 as a vital titrator of p53 creating a balance between Mdm2 and p53. In contrast, restricted atomic movements in the MdmX system complexes, with either Hit or RO-2443 ligands, might be due to a high demand to maintain an undisturbed *Sc*RING finger domain mediated activity of Mdm2 or p53 interactions in order to prevent Mdm2 degradation and p53 translocation (Wienken *et al*., 2016).

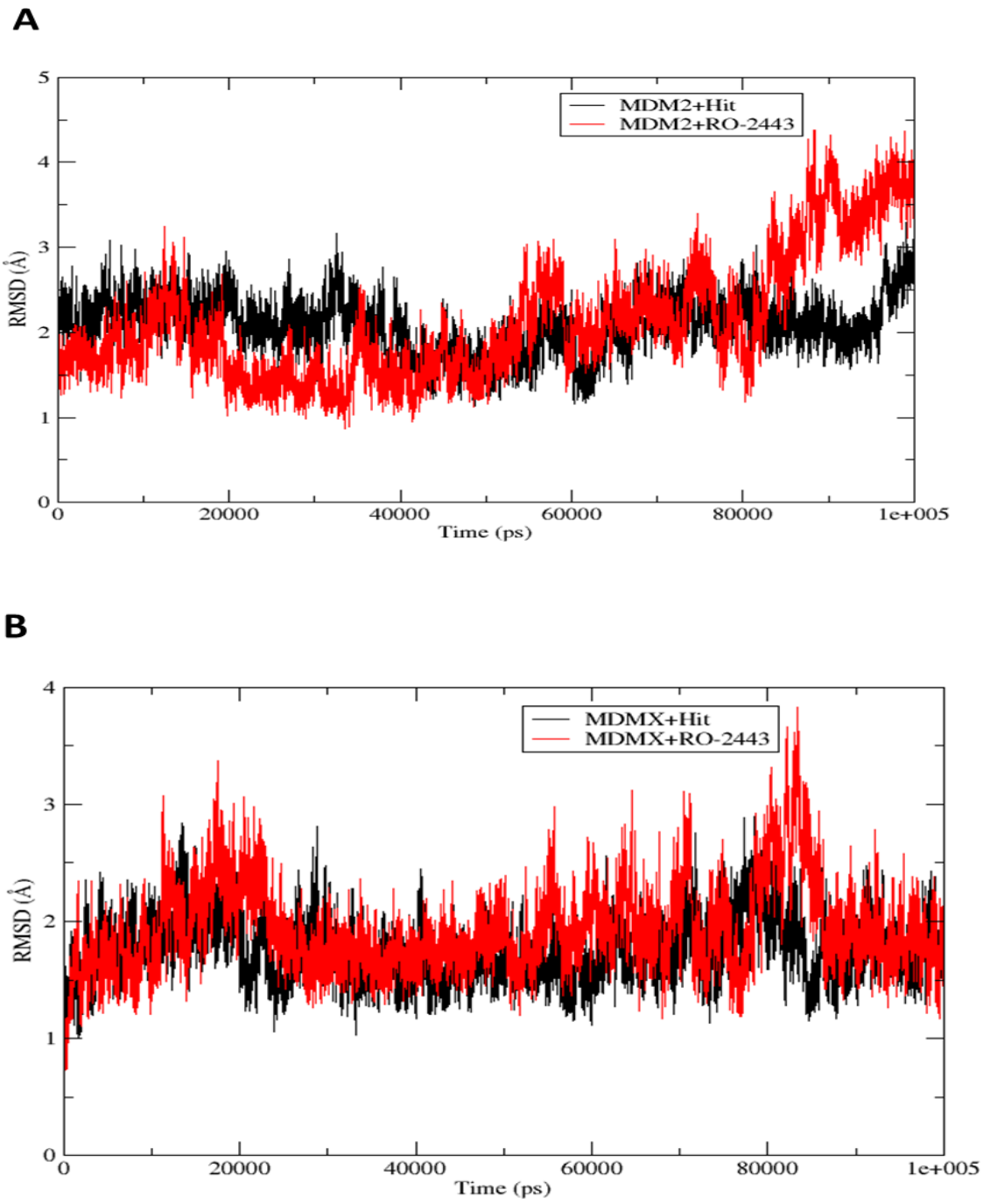


Figure 5.4: RMSD results for Mdm2 (A) and MdmX (B) bound to a ZINC16046951 as an experimental ligand and RO-2443 as a control ligand. The system complexes RMSD computed averages of Mdm2+ZINC16046951, Mdm2+RO-2443, MdmX+ZINC16046951 and MdmX+RO-2443 were 1.70 Å, 1.75 Å, 1.07 Å and 1.30 Å, respectively.

5.3.2 Root Mean Square Fluctuation (RMSF)

Protein functionality does not only depend on the protein's structure, protein flexibility and the fluctuation of amino acids is also considered. Therefore, RMS fluctuation was used to measure amino acid fluctuation in the binding site' protein complexes where ΔRMSF values >0.5 Å were considered a significant indicator of amino acid movement or fluctuation (Dong *et al*., 2018). Although, fluctuation of amino acids may be desirable for protein functionality, too much fluctuation may be a dangerous indication of protein unfolding (Khezri *et al*., 2018).

RMS fluctuation was calculated for all system complexes where the Mdm2 amino acid composition was 169 and MdmX was 164. In Figure 5.5 (A), a comparison between Mdm2+ZINC16046951 and Mdm2+RO-2443 exhibited a significant amino acid fluctuation in which Mdm2+ZINC16046951 showed less fluctuation compared to Mdm2+RO-2443. The most fluctuated amino acids for the Mmd2+ZINC16046951 complex were at position 1, 19, 72, 85, 104, 130 and 169, which are at the water exposed surface of the Mdm2 protein with RMSF values of 3.72 Å, 3.51 Å, 2.64 Å, 3.57 Å, 2.66 Å, 2.48 Å and 3.39 Å, respectively. The binding site amino acids RMS fluctuation contributions were Gly33 (0.86Å), Ile36 (0.74 Å), Met37 (0.95 Å), Arg40 (1.11 Å), Try42 (1.17 Å), Leu114 (0.95 Å), Leu177 (0.77 Å), Gly118 (0.85 Å), Ile121 (1.02 Å), Met122 (1.25 Å), Try127 (1.40 Å), Phe151 (1.39 Å) and Val153 (1.20 Å), where amino acid's number indicate the position. The Mdm2+ZINC16046951 complex was averaged to 1.64 Å which was less than Mdm2-RO2443 at 1.79 Å. A similar trend in RMS fluctuation was observed on the same amino acids at position 19 and 72, which is acceptable since the ligands were not 100% identical in structure. Additionally, RMS fluctuation measurements for the MdmX+ZINC16046951 complex, graphically

presented in Figure 5.5 (B), showed less fluctuation with the amino acid composition RMSF of not >2.5 Å. The MdmX amino acids responsible for the binding of a ZINC16046951 ligand, namely Met28, Leu31, Gly32, Ile35, Try41, Gln46, Phe65, Val67, Leu73, Ile117 and Gln128 (where a number indicate residue position), demonstrated RMS fluctuation of <2 Å elaborating that even though there was movement, it was not too much to risk the release of the ZINC16046951 ligand, instead necessary for the MdmX-ZINC16046951 complex activity. In comparison with MdmX+RO-2443, a small degree of RMS fluctuation >2.5 Å but <3 Å was observed in amino acids at position 83 and 164, which was assumed to be due to the high exposure to water molecules at the surface of the protein during 3D conformational folding. However, the amino acid responsible for the binding of the RO-2443 ligand also exhibited low RMS fluctuation, with values below 2 Å, which was lower than that of Mdm2 complexed with ligands. This corresponds to the results presented by RMSD that explains the rigidity and stability of the binding site's amino acids in which caution may be taken during it binding rigidity to avoid restricted flexibility of the ligand to protein interactions.

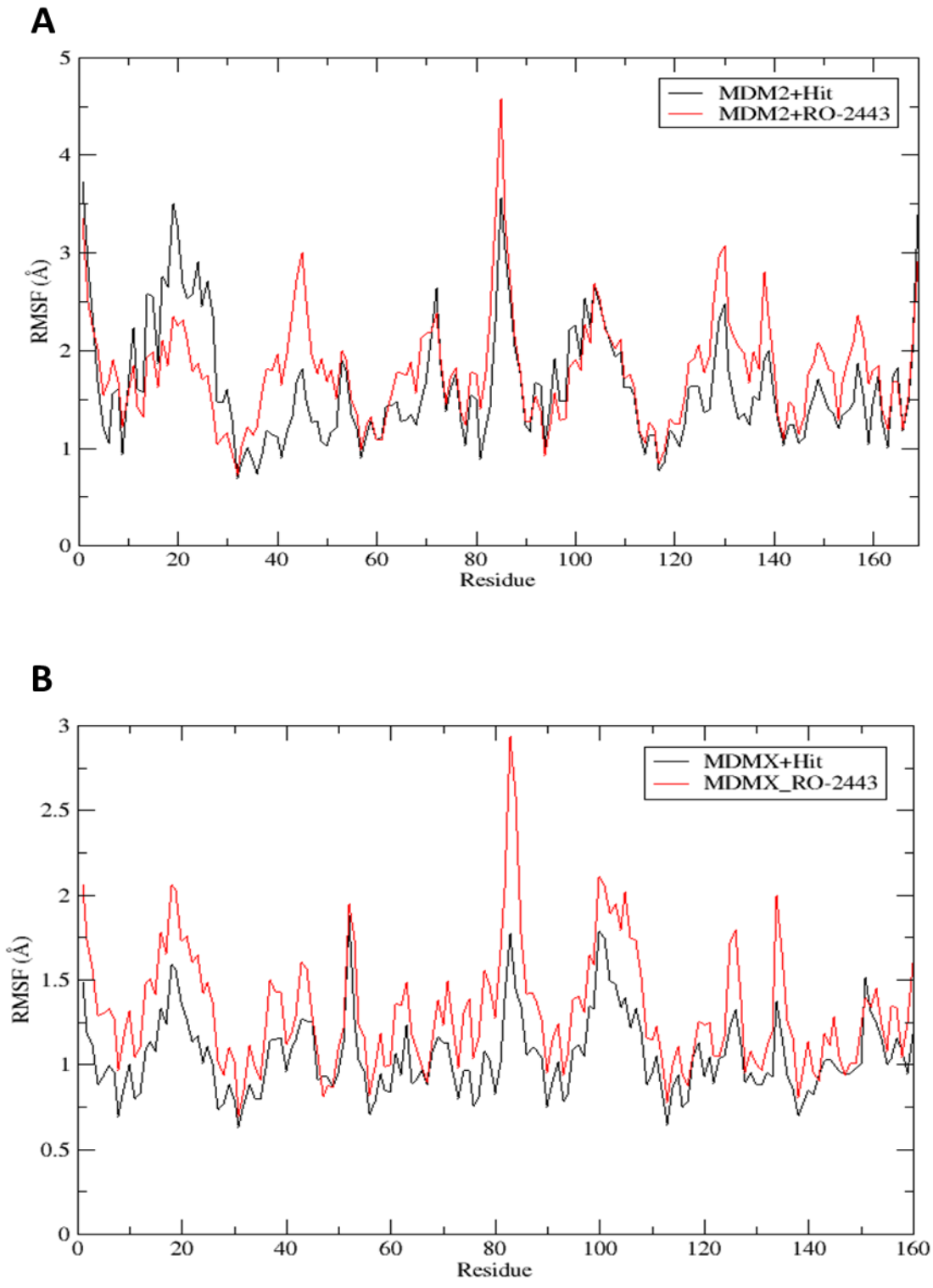


Figure 5.5: RMSF results for Mdm2 (A) and MdmX (B) bound to a ZINC16046951 as an experimental ligand and RO-2443 as a control ligand. Threshold value of >0.5 Å in RMSF was considered as a significant value signifying fluctuation or movement in Cα of the receptor protein's amino acids (Dong *et al*., 2016). The average of RMSF of Mdm2+ZINC16046951, Mdm2+RO-2443, MdmX+ZINC16046951 and MdmX+RO-2443 were computed to be 1.64 Å, 1.79 Å, 1.06 Å and 1.33 Å, respectively.

5.3.3. Radius of Gyration (RoG)

Radius of gyration is defined as a moment of inertia of Cα atoms measuring a weight-root mean square distance between the atoms of the complexes and their centre to compute the activity and compactness of the molecular structure (Khezri *et al*., 2018). Higher radius of gyration values corresponds to less compactness and activity of the complexes structure (Arnittali *et al*., 2019). Figure 5.6 (A) presents a comparison of structural compactness between Mdm2+ZINC16046951 and MdmX+RO-2443. There was a slight difference in structure compactness for both complexes in which Mdm2+ZINC16046951 had an average of 17.32 Å, which is lower than that of Mdm2+RO-2443 at 17.66 Å. Mdm2+ZINC16046951 shows an acceptable stable complex conformation and less fluctuation upon the ZINC16046951 ligand binding with a rotation point and maximum energy transfer 18 Å>17 Å. In contrast, Mdm2+RO-2443 demonstrated more fluctuation in structural compactness across the simulation time, which was expected as it had high RSMF values, which indicate less tightly packed Cα atoms (Dong *et al*., 2018; Arnittali *et al*., 2019). A similar trend is observed in MdmX+ZINC16046951 compared to MdmX+RO-2443 in Figure 5.6 (B), where MdmX+ZINC16046951 has a lower average of radius of gyration at 17.11 Å, while MdmX+RO-2443 has an average of 17.25Å. These results correspond to the RMSD and RMSF where more stable and less fluctuation in the structure was observed in the MdmX complexes compared to the Mdm2 complexes, insinuating more structural compactness in MdmX complexes than Mdm2 complexes (Lobanov *et al*., 2008).

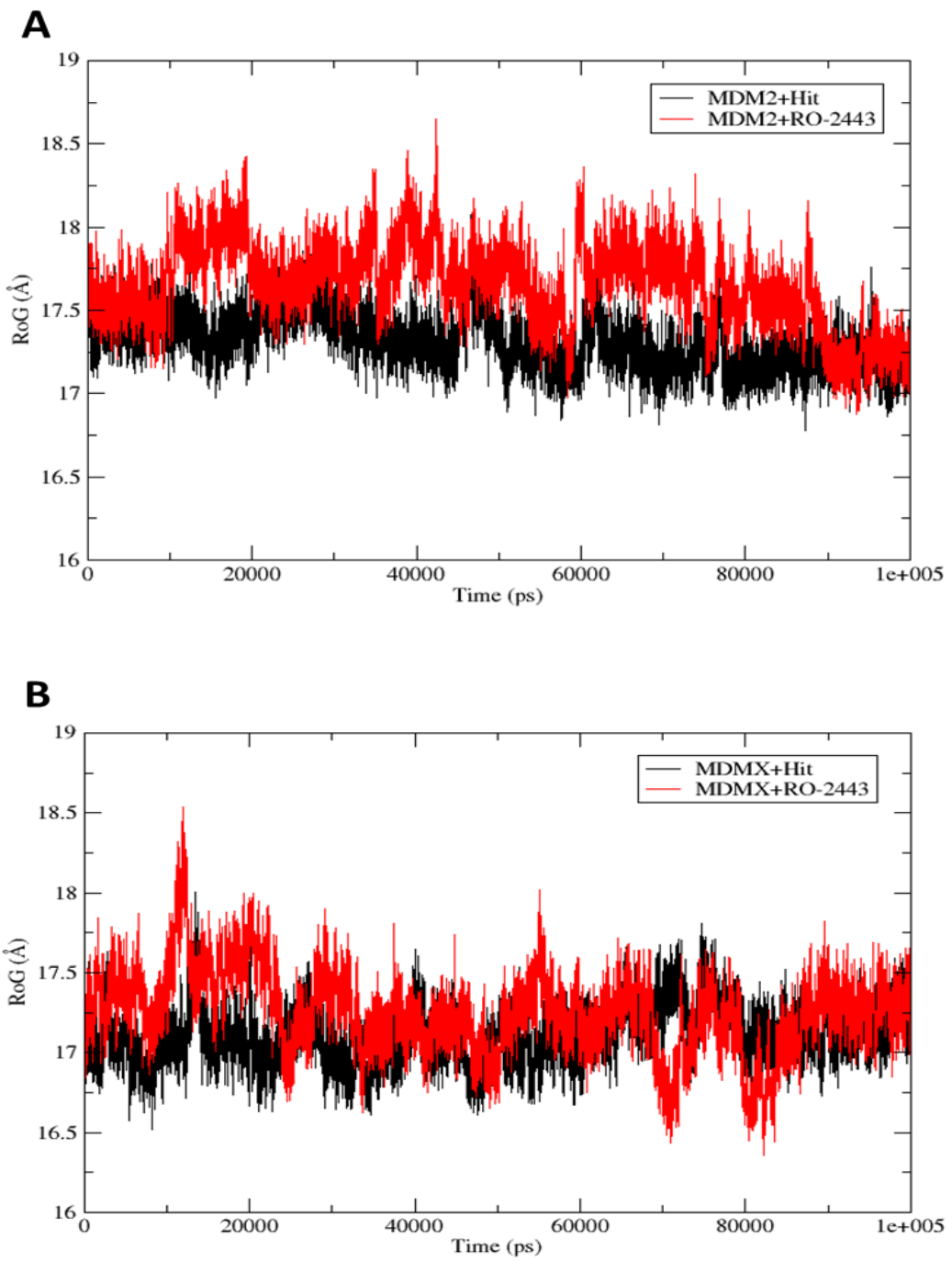


Figure 5.6: RoG results for Mdm2 (A) and MdmX (B) bound to a ZINC16046951 as an experimental ligand and RO-2443 as a control ligand. Higher values in radius of gyration are directly proportional to less compactness in protein complexes Cα atoms. Mdm2+ZINC16046951, Mdm2+RO-2443, MdmX+ZINC16046951 and MdmX+RO-2443 has an average of 17.32 Å, 17.66 Å, 17.11 Å, 17.25 Å.

5.3.4. Binding free-energy calculations

MM/PBSA and MM/GBSA approaches were used to estimate the total binding free energy contributed by decomposed complexes individual amino acids for a better understanding of a ligand binding at an atomistic level (Pandey *et al*., 2018; Xue *et al*., 2019). Figure 5.7 (A-D) illustrates van der waals (Evdw), electrostatics (Eele), polar and non-polar interactions, which were computed for an elaborated view on specific amino acids with most impact in energy component contributed to the total binding energy. The comparison between Mdm2+ZINC16046951 and Mmd2+RO-2443 shows that the most energy contributing amino acids in Mdm2+RO-2443 are Ile36(Evdw -1.48 Kcal/mol), MET37(Evdw -1.68 Kcal/mol), TYR42(Evdw -3.77 Kcal/mol), GLN47(Eele -4.6 Kcal/mol) and VAL153(Evdw -1.87 Kcal/mol), in which TYR42 and GLN47 were large contributors. MDM2+ZINC16046951 contributing amino acids include ARG40(Eele -5.43 Kcal/mol), TYR42(Evdw -2.74 Kcal/mol), TYR127 (Evdw -2.49 Kcal/mol) and VAL153(Evdw -2 Kcal/mol), in which ARG40 and TYR42 were the large contributors. Moreover, MdmX+ZINC16046951 and MdmX+RO-2443 binding energy amino acid contributors were GLN46(Eele -5.48 Kcal/mol), VAL67(Evdw -2.07 Kcal/mol), TYR123(Evdw -4.13 Kcal/mol) and GLN128(Eele -5.33 Kcal/mol) for MdmX+ZINC16046951, while MdmX+RO-2443 has TYR41(Evdw -2.56 Kcal/mol), GLN46(Eele -3.54 Kcal/mol), GLN128 (Evdw -2.4 Kcal/mol) and VAL149(Evdw -2.19 Kcal/mol). The high binding energy contributors for MDMX+ZINC16046951 were GLN46, GLN128 and TYR123 and for MdmX+ RO-2443, GLN46 was the highest contributor. From results stipulated in this report van der Waals interactions were shown to be a major interactive force contributor distributed among the highest amino acids contributors. Table 5.1 presents the total binding free energy for Mdm2+ZINC16046951, Mdm2+RO-2443, MdmX+ZINC16046951 and MdmX+RO-2443 to be -47.47 Kcal/mol, -46.28 Kcal/mol,

42.71 Kcal/mol and -45.65 Kcal/mol, respectively. Mdm2+ZINC16046951 has a higher binding free energy accounting for 47.47 Kcal/mol contributed by electrostatic (-21.47 Kcal/mol) and van der Waal (-56.45 Kcal/mol). Mdm2+RO-2443 has -46.28 Kcal/mol binding free energy contributed by electrostatic (-17.07 Kcal/mol) and van der Waals (-51.26 Kcal/mol). Additionally, MdmX+ZINC16046951 has a total binding free energy of -42.71 Kcal/mol contributed by electrostatic (-35.60 Kcal/mol) and van der Waals (-51.12 Kcal/mol). MdmX+RO-2443 binding free energy was -45.65 Kcal/mol, which is contributed by electrostatics and van der Waals at -20.97 Kcal/mol and -51.07 Kcal/mol, respectively. While Mdm2+ZINC16046951 had a higher binding free energy compared to Mdm2+RO-2443, MdmX+ZINC16046951 had lower binding free energy compared to MdmX+RO-2443. Moreover, Mdm2+ZINC16046951 showed a significantly higher ligand binding energy than MdmX+ZINC16046951, with a difference of -1.82 Kcal/mol. Major electrostatic component contributors for Mdm2+ZINC16046951 and MdmX+ZINC16046951 were ARG40(Eele -5.43 Kcal/mol) for Mdm2+ZINC16046951 and GLN46(Eele -5.48 Kcal/mol) together with GLN128(Eele -5.33 Kcal/mol) for MdmX+ZINC16046951. More details on the final decomposition binding energy for all complexes are in appendix iii.

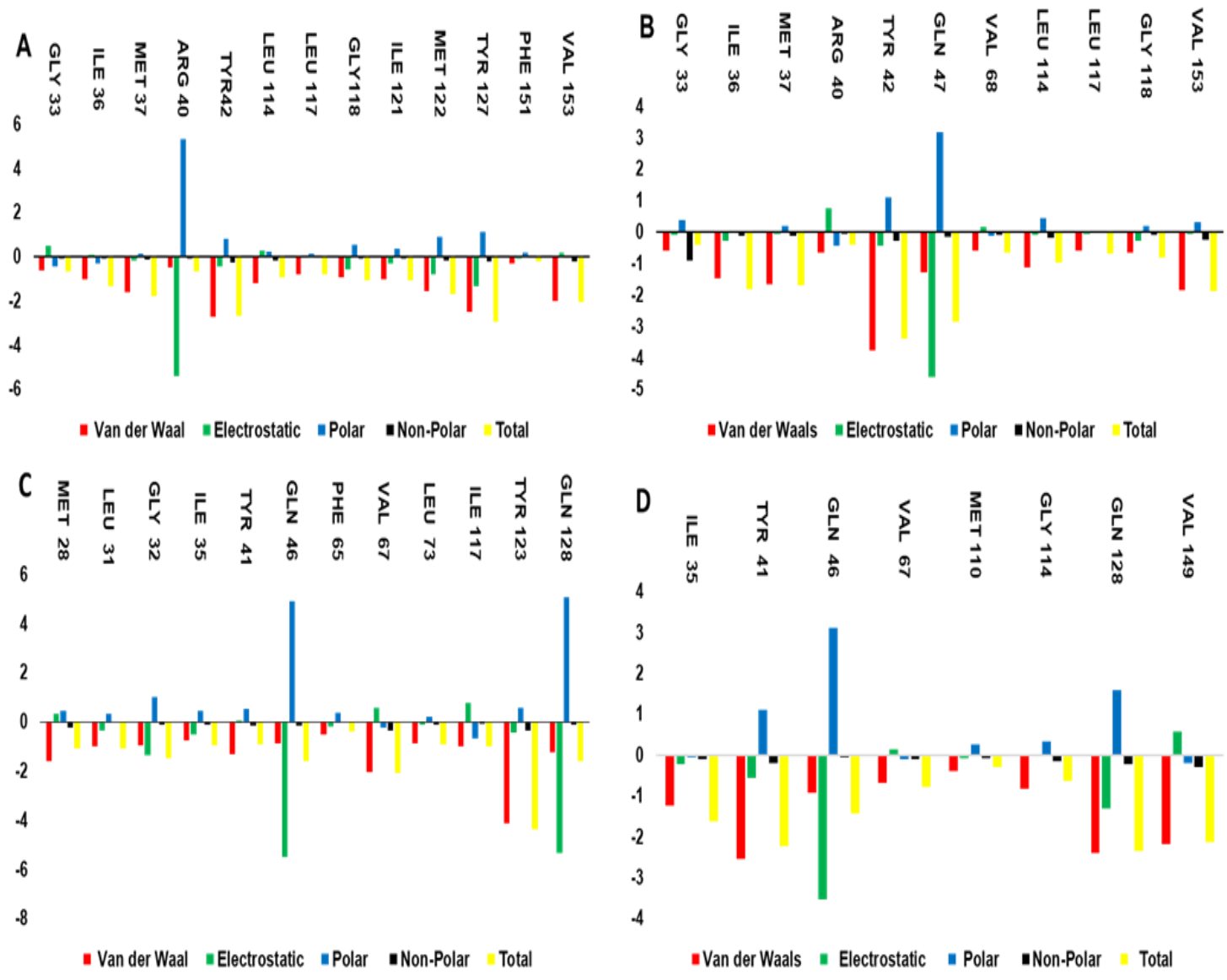


Figure 5.7: Graphical representation of per-residue free energy decomposition in Mdm2+ZINC16046951 (B), Mdm2+RO-2443(B), MdmX+ZINC16046951 (C) and MDMX+ RO-2443(D). MM/PBSA and MM/GBSA approach was used to compute per-residue free energy contributions in total binding energy from individual complexes.

Table 5.1: MM/PBSA and MM/GBSA total binding free-energy contributions for Mmd2+ZINC16046951, Mdm2+RO-2443, MdmX+ZINC16046951 and MdmX+ RO-2443.

	Energy Components (kcal/mol)				
Complexes	**Δ Evdw**	**Δ Eelec**	**Δ Ggas**	**Δ Gsolv**	**Δ Gbind**
MDM2+ ZINC16046951	-56.45 ± 3.39	-21.47 ±6.92	-77.92 ± 8.00	30.45 ± 5.20	-47.47 ± 4.52
MDM2+RO-2443	-51.26 ± 3.12	-17.07 ± 4.22	-68.33 ± 5.47	22.05 ± 2.71	-46.28 ± 4.22
MDMX+ ZINC16046951	-51.12 ± 2.92	-35.60 ± 6.85	-86.73 ± 7.10	44.02 ± 5.81	-42.71 ± 3.38
MDMX+RO-2443	-51.07 ± 3.45	-20.97 ± 4.66	-72.03 ± 6.00	26.39 ± 3.31	-45.65 ± 3.82

Δ Evdw: van der Waals, **Δ** Eelec: electrostatic, **Δ** Ggas: gas phase interaction, and **Δ** Gsolv: solvation free energy **Δ** Gbind: calculated total free binding energy.

5.4 Conclusion

RING finger containing inhibitors Mdm2 and MdmX activity depends on the presence of RING finger domain. Crystal structures of Mdm2 and MdmX complex individually, with RO-2443 for control and ZINC16046951 as the experimental ligand, were subjected to *in silico* molecular dynamics, to acquire information regarding the distance between the receptor and ligand using RMSD, complexes flexibility and fluctuation through RMSF, and structural compactness and activity through RoG. RMSD results show MdmX+ZINC16046951 to be more stable than Mdm2+ZINC16046951; this occurs to maintain an undisturbed interaction of RING finger domain in the MdmX structure. The control RO-2443 ligand complex was less stable than MdmX+ZINC16046951 but more stable than Mdm2+ZINC16046951. This was assumed to be due to the significance of the Mdm2 protein to open and close during the binding of a ZINC16046951 ligand to avoid permanent inhibition of p53 activity. Following this, RMSF was used to determine protein functionality by measuring the complexes flexibility and fluctuation of amino acids in the binding site. Values greater that 0.5 Å were considered significant to signify amino acid fluctuation. Moreover, greater values of RMSF were observed in Mdm2+RO-2443 compared to Mdm2+ZINC16046951. However, it was not clear whether this was due to protein activity such as protein unfolding. MdmX+ZINC16046951 showed less structural flexibility compared to MdmX+RO-2443. This denotes that the stability during the interaction of the MdmX+ZINC16046951 complex is biologically necessary to allow p53 activity in preventing tumorigenesis. Radius of gyration was used to measure the inertia of Cα atoms in the complexes to determine the structural compactness and activity. There was a small difference of 0.34 Å for structural compactness and activity between Mdm2+ZINC16046951 and Mdm2+RO-2443. MdmX+RO-2443 had a higher structural compactness and activity compared to MdmX+ZINC16046951, with a

difference of 0.11 Å. This interesting observation led to the conclusion that even though protein-ligand interaction may be rigid, it still allows for flexibility during the moment of inertia initiation, between Cα atoms necessary for Mdm2+ZINC16046951 and MdmX+ZINC16046951 activity. Therefore, MM/GBSA and MM/PBSA depict binding free-energy contributions for complexes, confirming the ZINC16046951 ligand as the strongest binding inhibitor for Mdm2 compared to RO-2443, while Mdm2 provides an adequate flexibility of the structure to open and close during ligand binding. This is a desirable relationship required between a receptor and ligand during drug discovery.

CHAPTER 6

STUDY'S LIMITATIONS AND FUTUREWORK

6.1. Limitations of this study

Although more clarity could have been provided on structural characterization of the protein through NMR studies, such studies could not be persuaded because of the small yield of the protein obtained. The financial constraints and the negative effects that the COVID-19 pandemic also contributed to the inability to go into NMR studies.

6.2. Future work

Structural characterisations of a *Sc*RING finger protein presented in this study contributed a stable background for further studies that will provide clarity on whether this protein is a RING finger domain or belongs to the class of RING finger-like proteins. Although SE-HPLC quaternary structure determination gave an insight of a possible homo-dimeric subunit. However, more clarity will be provided with the use of X-ray crystallography, NMR spectroscopy and/or electron microscopy. CD spectroscopy results showed the protein structure behavioural changes of the *Sc*RING finger protein as there were more secondary elements that were dictated on a denatured protein compared to the native protein. However, for future studies further investigation on the protein behaviour using CD, in the absence of a zinc metal ion (the structural stability cofactor) will give an insight on whether the protein maintains it structure or not, if the structure is still maintained, what are other interactions bonds that contributes to structural stability.

REFERENCES

Agarwal P and Khurana P (2018). Characterization of a novel zinc finger transcription factor (TaZnF) from wheat conferring heat stress tolerance in Arabidopsis. *Cell Stress and Chaperones* **23**(2): 253-267.

Aldaghmi AS (2017). The Role of CREG1 as a Master Regulator of Liver Function. *Physical Review Journals* **27**: 19

Al-saran N, Subash-Babu P, Al-Nouri DM, Alfawaz HA and Alshatwi AA (2016). Zinc enhances CDKN2A, pRb1 expression and regulates functional apoptosis via upregulation of p53 and p21 expression in human breast cancer MCF-7 cell. *Environmental Toxicology and Pharmacology* **47**: 19-27.

Amar B, Ghisaidoobe Sang T and Chung J (2014). Intrinsic Tryptophan Fluorescence in the Detection and Analysis of Proteins: A Focus on Förster Resonance Energy Transfer Techniques*. International Journal of Molecular Sciences* **15**: 22518-22538.

Anczuków O and Krainer AR (2016). Splicing-factor alterations in cancers. *RNA* **22**(9): 1285-1301.

Aravind L and Koonin EV (2000). The U box is a modified RING finger—a common domain in ubiquitination. *Current Biology* **10**(4): 132-134.

Arnittali M, Rissanou AN and Harmandaris V (2019). Structure of biomolecules through molecular dynamics simulations. *Procedia Computer Science* **156**: 69-78.

Baskar R, Lee KA, Yeo R and Yeoh KW (2012). Cancer and radiation therapy: current advances and future directions. *International Journal of Medical Sciences* **9**(3): 193.

Cahilly-Snyder L, Yang-Feng T, Francke and George DL (1987). Molecular analysis and chromosomal mapping of amplified genes isolated from a transformed mouse 3T3 cell line. *Somatic Cell and Molecular Genetics* **13**(3): 235-244.

Cajee UF, Hull R and Ntwasa M (2012). Modification by ubiquitin-like proteins: significance in apoptosis and autophagy pathways. *International Journal of Molecular Sciences* **13**(9): 11804-11831.

Callis J, Stone SL, Hauksdottir H, Troy A, Herschleb J and Kraft E (2005). Functional analysis of the RING-type ubiquitin ligase family of *Arabidopsis*. *Plant Physiology* **137**: 13-30.

Capili AD, Schultz DC, Rauscher FJ and Borden KL (2001). Solution structure of the PHD domain from the KAP-1 corepressor: structural determinants for PHD, RING and LIM zinc-binding domains. *The EMBO journal* **20**(1-2): 165-177.

Carlson LA, Bai Y, Keane SC, Doudna JA and Hurley JH (2016). Reconstitution of selective HIV-1 RNA packaging in vitro by membrane-bound Gag assemblies. *Elife* **5**: 14663.

Chasapis CT, Loutsidou AK, Orkoula MG and Spyroulias GA (2010). Zinc binding properties of engineered RING finger domain of Arkadia E3 ubiquitin ligase. *Bioinorganic Chemistry and Applications.* doi:10.1155/2010/323152.

Chatr-Aryamontri A van der Sloot A and Tyers M (2018). At long last, a C-terminal bookend for the ubiquitin code. *Molecular Cell* **70**(4): 568-571.

Chen J, Li Y, Li F, Wu Q, Jiang Y and Yuan D (2018). Banana MaABI5 is involved in ABA-induced cold tolerance through interaction with a RING E3 ubiquitin ligase, MaC3HC4-1. *Scientia Horticulturae* **237**: 239-246.

Coscoy L and Ganem D (2003). PHD domains and E3 ubiquitin ligases: viruses make the connection. *Trends in Cell Biology* **13**(1): 7-12.

Damm KL and Carlson HA (2006). Gaussian-weighted RMSD superposition of proteins: a structural comparison for flexible proteins and predicted protein structures. *Biophysical Journal* **90**(12): 4558-4573.

Darmon SK and Lutz CS (2012). mRNA 3′ end processing factors: a phylogenetic comparison. *Comparative and Functional Genomics* ***2012***. doi:10.1155/2012/876893

Dawid IG, Breen JJ and Toyama R (1998). LIM domains: multiple roles as adapters and functional modifiers in protein interactions. *Trends in Genetics* **14**(4).

Desjardins P and Conklin D (2010). NanoDrop microvolume quantitation of nucleic acids. *Journal of Visualized Experiments* **45**: 2565.

Di Giammartino DC, Li W, Ogami K, Yashinskie JJ, Hoque M, Tian B and Manley JL (2014). RBBP6 isoforms regulate the human polyadenylation machinery and modulate expression of mRNAs with AU-rich 3′ UTRs. *Genes & development* **28**(20): 2248-2260.

Doig AJ, Andrew CD, Cochran DA, Hughes E, Penel S, Sun JK, Stapley BJ, Clarke DT and Jones GR (2001). Structure, stability and folding of the α-helix. *Biochemical Society Symposia* **68**: 95-110.

Dominguez C, Folkers GE and Boelens R (2004). Biological introduction: RING domain proteins. *Handbook of Metalloproteins* **3**: 338-351.

Dong YW, Liao ML, Meng XL and Somero GN (2018). Structural flexibility and protein adaptation to temperature: Molecular dynamics analysis of malate dehydrogenases of marine molluscs. *Proceedings of the National Academy of Sciences* **115**(6): 1274-1279.

Drozdetskiy A, Cole C, Proter J, Barton GJ (2015). JPred4: A protein secondary structure prediction server. *Nucleic Acids Research* **1**(43): 389-394.

Dudev T and Lim C (2003). Metal binding and selectivity in zinc proteins. *Journal of the Chinese Chemical Society* **50**(5): 1093-1102.

Dunbar J, Yennawar HP, Banerjee S, Luo J and Farber GK (1997). The effect of denaturants on protein structure. *Protein Science* **6**(8): 1727-1733.

Eitzen G, Smithers CC, Murray AG and Overduin M (2019). Structure and function of the Fgd family of divergent FYVE domain proteins. *Biochemistry and Cell Biology* **97**(3): 257-264.

Eva YHL, Chang CY, Hu N, Wang YCJ, Lai CC, Herrup K, Lee WH and Bradley A (1992). Mice deficient for Rb are nonviable and show defects in neurogenesis and haematopoiesis. *Nature* **359**(6393): 288-294.

Fakharzadeh SS, Trusko SP and George DL (1991). Tumorigenic potential associated with enhanced expression of a gene that is amplified in a mouse tumour cell line. *EMBO Journal* **10**(6): 1565-1569.

Fathi-Roudsari M, Akhavian-Tehrani A and Maghsoudi N (2016). Comparison of three Escherichia coli strains in recombinant production of reteplase. *Avicenna Journal of Medical Biotechnology* **8**(1): 16.

Finley D, Ulrich HD, Sommer T and Kaiser P (2012). The ubiquitin–proteasome system of Saccharomyces cerevisiae. *Genetics* **192**(2): 319-360.

Finley J (2015). Reactivation of latently infected HIV-1 viral reservoirs and correction of aberrant alternative splicing in the LMNA gene via AMPK activation: Common mechanism of action linking HIV-1 latency and Hutchinson–Gilford progeria syndrome. *Medical Hypotheses* **85**(3): 320-332.

Fischer M, Quaas M, Nickel A and Engeland K (2015). Indirect p53-dependent transcriptional repression of Survivin, CDC25C, and PLK1 genes requires the cyclin-dependent kinase inhibitor p21/CDKN1A and CDE/CHR promoter sites binding the DREAM complex. *Oncotarget* **6**(39): 1402.

Freemont PS, Hanson IM and Trowsdale J (1991). A novel cysteine-rich sequence motif. *Cell* **64**(3): 483-484.

Gao S, Witte MM and Scott RE (2002). P2P-R protein localizes to the nucleolus of interphase cells and the periphery of chromosomes in mitotic cells which show maximum P2P-R immunoreactivity. *Journal of Cellular Physiology* **191**(2): 145-154.

Gasymov OK and Glasgow BJ (2007). ANS fluorescence: potential to augment the identification of the external binding sites of proteins. *Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics* **1774**(3): 403-411.

Genheden Sand Ryde U (2015). The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities. *Expert opinion on drug discovery* **10**(5): 449-461.

George P (2011). p53 how crucial is its role in cancer. *International Journal of Current Pharmaceutical Research* **3**(2): 19-25.

Ghisaidoobe AB and Chung SJ (2014). Intrinsic tryptophan fluorescence in the detection and analysis of proteins: a focus on Förster resonance energy transfer techniques. *International Journal of Molecular Sciences* **15**(12): 22518-22538.

Giacinti C and Giordano A (2006). RB and cell cycle progression. *Oncogene* **25**(38): 5220-5227.

Gonzalez F, Loidi L and Abalo-Lojo JM (2017). Novel variant in the TP63 gene associated to ankyloblepharon-ectodermal dysplasia-cleft lip/palate (AEC) syndrome. *Ophthalmic Genetics* **38**(3): 277-280.

Goyal L, McCall K, Agapite J, Hartwieg E and Steller H (2000). Induction of apoptosis by Drosophila reaper, hid and grim through inhibition of IAP function. *EMBO Journal* **19**(4): 589-597.

Graveley BR and Maniatis T (1998). Arginine/serine-rich domains of SR proteins can function as activators of pre-mRNA splicing. *Molecular Cell* **1**(5): 765-771.

Graves B, Thompson T, Xia M Janson C, Lukacs C, Deo D, Di Lello P, Fry D, Garvie C, Huang KS and Gao L (2012). Activation of the p53 pathway by small-molecule-induced MDM2 and MDMX dimerization. Proceedings of the National Academy of Sciences 109 (29): 11788-11793.

Green JA, Kirwan JM, Tierney JF, Symonds P, Fresco L, Collingwood M and Williams CJ (2001). Survival and recurrence after concomitant chemotherapy and radiotherapy for cancer of the uterine cervix: a systematic review and meta-analysis. *Lancet* **358**(9284): 781-786.

Greenfield NJ (2006). Using circular dichroisim spectra to estimate protein secondary structure. *Nature Protocols* **1**(6): 2876-2860.

Guharoy M, Bhowmick P, Sallam M and Tompa P (2016). Tripartite degrons confer diversity and specificity on regulated protein degradation in the ubiquitin-proteasome system. *Nature Communications* **7**: 10239

Hanwell MD, Curtis DE, Lonie DC, Vandermeersch T, Zurek E and Hutchison GR (2012). Avogadro: an advanced semantic chemical editor, visualization, and analysis platform. *Journal of cheminformatics* **4**(1): 1-17.

Hassem FA (Masters' thesis) 2014. Investigation of protein-metal ion and protein-protein interactions using mass spectrometry and nuclear magnetic spectroscopy. University of Western cape. http://hdl.handle.net/11394/3369. Accessed 16/02/2020

Hatakeyama S and Nakayama KI (2003). U-box proteins as a new family of ubiquitin ligases. *Biochemical and Biophysical Research Communications* **302**: 635-645.

Hollander D, Naftelberg S, Lev-Maor G, Kornblihtt AR and Ast G (2016). How are short exons flanked by long introns defined and committed to splicing? *Trends in Genetics* **32**(10): 596-606.

Holub JM (2017). Small scaffolds, big potential: developing miniature proteins as therapeutic agents. *Drug Development Research* **78**(6): 268-282.

Hong H, Luo Y, Zhou Z and Shen H (2012). Effects of low concentration of salt and sucrose on the quality of bighead carp (Aristichthys nobilis) fillets stored at 4°C. *Food Chemistry* **133**(1): 102-107.

Houben K, Wasielewski E, Dominguez C, Kellenberger E, Atkinson RA, Timmers HTM, Kieffer B and Boelens R (2005). Dynamics and metal exchange properties of C4C4 RING domains from CNOT4 and the p44 subunit of TFIIH. *Journal of Molecular Biology* **349**(3): 621-637.

Hu D, Zhang S, Zhao Y, Wang S, Wang Q, Song X, Lu D, Mao Y and Chen H (2014). Association of genetic variants in the retinoblastoma binding protein 6 gene with the risk of glioma: a case-control study in a Chinese Han population. *Journal of Neurosurgery* **121**(5): 1209-1218.

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. *PLoS ONE* **8**(3): e58736.

Hull R, Oosthuysen B, Cajee UF, Mokgohloa L, Nweke E, Antunes RJ, Coetzer TH and Ntwasa M (2015). The Drosophila Retinoblastoma Binding Protein 6 Family Member Has Two Isoforms and Is Potentially Involved in Embryonic Patterning. *International Journal of Molecular Sciences* **16**(5): 10242-10266.

Jiang L and Zawacka-Pankau J (2020). The p53/MDM2/MDMX-targeted therapies—a clinical synopsis. *Cell Death & Disease* **11**(4): 1-4.

Joazeiro CA and Weissman AM (2000). RING finger proteins: mediators of ubiquitin ligase activity. *Cell* **102**(5): 549-552.

Jobbins A M (PhD Thesis) 2018: The Involvement of SRSF1 in pre-mRNA splicing. University of Leicester. https://hdl.handle.net/2381/42848. Accessed 20/01/2021.

Kang W, Jiang F and Wu YD (2018). Universal Implementation of a Residue-Specific Force Field Based on CMAP Potentials and Free Energy Decomposition. *Journal of Chemical Theory and Computation* **14**(8): 4474-4486.

Kappo MA (PhD thesis) 2009. Solution structure of the RING finger domain from the human splicing-associated protein RBBP6 using heteronuclear Nuclear Magnetic Resonance (NMR) Spectroscopy. http://hdl.handle.net/11394/3231. Accessed 20\01\2021.

Kappo MA, AB E, Hassem F, Atkinson RA, Faro A, Muleya V, Mulaudzi T, Poole JO, McKenzie JM, Chibi M, Moolman-Smook JC, DJG Rees and DJR Pugh (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.

Katoh S, Hong C, Tsunoda Y, Murata K, Takai R, Minami E, Yamazaki T and Katoh E (2003). High precision NMR structure and the function of the RING-H2 finger domain of EL5, a rice protein whose expression is increased upon exposure to pathogen-derived oligosaccharides. *Journal of Biological Chemistry* **278**: 15341-15348.

Kellenberger E, Dominguez C, Friborug S, Wasielewski E, Moras D, Poterszman A, Boelens R and Kieffer B (2005). Solution structure of the carboxy-terminal of human TFIIH P_{44} subunit. *Journal of Biological Chemistry* **349**(3): 621-637.

Khezri A, Karimi A, Yazdian F, Jokar M, Mofradnia SR, Rashedi H and Tavakoli, Z., (2018). Molecular dynamic of curcumin/chitosan interaction using a computational molecular approach: Emphasis on biofilm reduction. *International Journal of Biological Macromolecules* **114**: 972-978.

Knudson AG (1971). Mutation and cancer: statistical study of retinoblastoma. *Proceedings of the National Academy of Sciences* **68**(4): 820-823.

Kobashigawa Y, Tomitaka A, Kumeta H, Noda NN, Yamaguchi M and Inagaki F (2011). Autoinhibition and phosphorylation-induced activation mechanisms of human cancer and autoimmune disease-related E3 protein Cbl-b. *Proceedings of the National Academy of Sciences* **108**(51): 20579-20584.

Kocatürk NM (Doctoral thesis) 2018. Cross-link between the ubiquitin-proteasome system (UPS) and autophagy in the regulation of mitophagy. Sabanci University. https://research.sabanciuniv.edu/36991/1/10206918_NurMehpareKocaturk.pdf. Accessed 20/01/2021.

Koes DR and Camacho CJ (2012). ZINCPharmer: pharmacophore search of the ZINC database. Nucleic Acids Research 40 (1): 409-414.

Koiwai H, Tagiri A, Katoh S, Katoh E, Ichikawa H, Minami E and Nishizawa Y (2007). RING-H2 type ubiquitin ligase EL5 is involved in root development through the maintenance of cell viability in rice. *Plant Journal* **51**(1): 92-104.

Komori H, Iwanaga R, Bradford AP, Araki K and Ohtani K (2018). Distinct E2F-Mediated Transcriptional Mechanisms in Cell Proliferation, Endoreplication and Apoptosis. *In Gene Regulation*. DOI: 10.5772/intechopen.82448.

Kosarev P, Mayer KF and Hardtke CS (2002). Evaluation and classification of RING-finger domains encoded by the Arabidopsis genome. *Genome Biology* **3**(4): 16-1.

Kostic M, Matt T, Matinez-Yamout MA, Dyson HJ and Wright PE (2006). Solution structure of the Hdm2 C2H2C4 RING, a domain critical for ubiquitination. *Journal of Molecular Biological* **363**: 433-450.

Krishna SS, Majumdar I and Grishin NV (2003). Structural classification of zinc fingers: survey and summary. *Nucleic Acids Research* **31**(2): 532-550.

Kufareva I, Abagyan R (2011). Methods of Protein Structure Comparison. *Homology Modeling Methods in Molecular Biology* **857**. https://doi.org/10.1007/978-1-61779-588-6_10. Accessed 20/01/2021.

Kumar K, Muthamilarasan M, Bonthala VS, Roy R and Prasad M (2015). Unraveling 14-3-3 proteins in C4 panicoids with emphasis on model plant Setaria italica reveals phosphorylation-dependent subcellular localization of RS splicing factor. *PLoS ONE* **10**(4): 0123236.

Larsen EC (2018). Interaction of Zinc Finger Proteins 146 & 507 with the Transposable Element LINE-1. https://digitalcommons.wpi.edu/mqp-all/2502. Accessed 19/01/2021.

Lazauskas T, Sokol AA and Woodley SM (2017). An efficient genetic algorithm for structure prediction at the nanoscale. *Nanoscale* **9**(11): 3850-3864.

Lee G, Sehgal R, Wang Z, Nair S, Kikuno K, Chen CH, Hay B and Park JH (2013). Essential role of grim-led programmed cell death for the establishment of corazonin-producing peptidergic nervous system during embryogenesis and metamorphosis in Drosophila melanogaster. *Biology Open* **2**(3): 283-294.

Lee SD and Moore CL (2014). Efficient mRNA polyadenylation requires a ubiquitin-like domain, a zinc knuckle, and a RING finger domain, all contained in the Mpe1 protein. *Molecular and Cellular Biology* **34**(21): 3955-3967.

Lemay G (2018). Synthesis and translation of viral mRNA in reovirus-infected cells: Progress and remaining questions. *Viruses* **10**(12): 671.

Lewis PN, Momany FA and Scheraga HA (1973). Chain reversals in proteins. *Biochemica Biophysica et Acta* **303**: 211-229.

Lewkowicz DJ (2014). Early experience and multisensory perceptual narrowing. *Developmental psychobiology* **56**(2): 292-315.

Li L, Deng B, Xing G, Teng Y, Tian C, Cheng X, Yin X, Yang J, Gao X, Zhu Y and Sun, Q (2007). PACT is a negative regulator of p53 and essential for cell growth and embryonic development. *Proceedings of the National Academy of Sciences* **104**(19): 7951-7956.

Lian X Ma K, Xiao J, Li X, Zhang Q and (2009). Sequence and expression analysis of the C3HC4-type RING finger gene family in rice. *Gene* **444**(1): 33-45.

Lobanov MY, Bogatyreva NS and Galzitskaya OV (2008). Radius of gyration as an indicator of protein structure compactness. *Molecular Biology* **42**(4): 623-628.

Lu YF, Xu XP, Lu XP, Zhu Q, Liu G, Bao YT, Wen H, Li YL, Gu W and Zhu WG (2020). SIRT7 activates p53 by enhancing PCAF-mediated MDM2 degradation to arrest the cell cycle. *Oncogene* **39**: 4650–4665.

Lwakuma T and Lozano G (2003). MDM2, an introduction. *Molecular Cancer Research* **1**(14): 993-1000.

Mandel CR, Bai Y and Tong L (2008). Protein factors in pre-mRNA 3′-end processing. *Cellular and Molecular Life Sciences* **65**(7): 1099-1122.

Massova I and Kollman PA (2000). Combined molecular mechanical and continuum solvent approach (MM-PBSA/GBSA) to predict ligand binding. *Perspectives in drug discovery and design* **18**(1): 113-135.

Mathenjwa MS (Master's thesis) 2018. Recombinant expression, purification and structural characterization of Saccharomyces cerevisiae RING finger domain using Nuclear Magnetic Resonance (NMR) spectroscopy. University of Zululand. http://hdl.handle.net/10530/1662. Accessed 20/01/2021.

Metzger MB, Hristova VA and Weissman AM (2012). HECT and RING finger families of E3 ubiquitin ligases at a glance. *Journal of Cellular Sciences* **125**(3): 531-7.

Micsonai A, Wien F, Bulyáki É, Kun J, Moussong É, Lee YH, Goto Y, Réfrégiers M and Kardos J (2018). BeStSel: a web server for accurate protein secondary structure prediction and fold recognition from the circular dichroism spectra. *Nucleic acids research* **46**(1): 315-322.

Middendorf TR, Aldrich RW and Baylor DA (2000). Modification of cyclic nucleotide–gated ion channels by ultraviolet light. *Journal of General Physiology* **116**(2): 227-252.

Miyamoto K, Uechi A and Saito K (2017). The zinc finger domain of RING finger protein 141 reveals a unique RING fold. *Protein Science* **26**(8): 1681-1686.

Moela P, Choene MM and Motadi LR (2014). Silencing RBBP6 (Retinoblastoma Binding Protein 6) sensitises breast cancer cells MCF7 to staurosporine and camptothecin-induced cell death. *Immunobiology* **219**(8): 593-601.

Motadi LR, Bhoola KD and Dlamini Z (2011). Expression and function of retinoblastoma binding protein 6 (RBBP6) in human lung cancer. *Immunobiology* **216**: 1065-1073.

Ndabambi N (Doctoral thesis) 2004. Recombinant expression of the pRb-and p53-interacting domains from the human RBBP6 protein for in vitro binding studies. University of the Western Cape. http://hdl.handle.net/11394/1685. Accessed 20/01/2021.

Ndagi U, Mhlongo NN and Soliman ME (2018). Emergence of a promising lead compound in the treatment of triple negative breast cancer: an insight into conformational features and ligand binding landscape of c-Src protein with UM-164. *Applied Biochemistry and Biotechnology* **185**(3): 655-675.

Nicolay BN, Danielian PS, Kottakis F, Lapek JD, Sanidas I, Miles WO, Dehnad M, Tschöp K, Gierut JJ, Manning AL and Morris R (2016). Abstract IA15: The consequences of pRb inactivation: insights from a proteomic analysis of Rb loss. *Cancer Cell Cycle - Tumor Progression and Therapeutic Response* **14**(11).

Nowick JS (2008). Exploring β-sheet structure and interactions with chemical model systems. *Accounts of Chemical Research* **41**(10): 1319-1330.

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

Okorokov AL, Sherman MB, Plisson C, Grinkevich V, Sigmundsson K, Selivanova G, Milner J and Orlova EV (2006). The structure of p53 tumour suppressor protein reveals the basis for its functional plasticity. *EMBO Journal* **25**(21): 5191-5200.

Oliner JD, Saiki AY and Caenepeel S (2016). The role of MDM2 amplification and overexpression in tumorigenesis. *Cold Spring Harbor Perspectives in Medicine* **6**(6): a026336.

Ong EE and Liow JL (2019). The temperature-dependent structure, hydrogen bonding and other related dynamic properties of the standard TIP3P and CHARMM-modified TIP3P water models. *Fluid Phase Equilibria* **481**: 55-65.

Pandey P, Srivastava R and Bandyopadhyay P (2018). Comparison of molecular mechanics-Poisson-Boltzmann surface area (MM-PBSA) and molecular mechanics-three-dimensional reference interaction site model (MM-3D-RISM) method to calculate the binding free energy of protein-ligand complexes: Effect of metal ion and advance statistical test. *Chemical Physics Letters* **695**: 69-78.

Parveen R, Rahman SS, Sultana SA and Habib ZH (2015). Cancer Types and Treatment Modalities in Patients Attending at Delta Medical College Hospital. *Delta Medical College Journal* **3**(2): 57-62.

Perczel A, Gáspári Z and Csizmadia IG (2005). Structure and stability of β-pleated sheets. *Journal of Computational Chemistry* **26**(11): 1155-1168.

Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC and Ferrin TE (2004). UCSF Chimera—A visualization system for exploratory research and analysis. *Journal of Computational Chemistry* **25**: 1605–1612.

Pickart CM and Fushman D (2004). Polyubiquitin chains: polymeric protein signals. *Current Opinions in Chemical Biology* **8**(6): 610-616.

Poursheikhali Asgary M, Jahandideh S, Abdolmaleki P and Kazemnejad A (2007). Analysis and identification of β-turn types using multinomial logistic regression and artificial neural network. *Bioinformatics* **23**(23): 3125-3130.

Pugh DJR, AB E, Faro A, Lutya PT, Hoffmann E and Rees DJG (2006). DWNN, a novel ubiquitin-like domain, implicates RBBP6 in mRNA processing and ubiquitin-like pathways. *BioMed Central Structural Biology* **6**: 1.

Qi DL and Cobrinik D (2017). MDM2 but not MDM4 promotes retinoblastoma cell proliferation through p53-independent regulation of MYCN translation. *Oncogene* **36**(13): 760-1769.

Qi L, Lu Z, Sun YH, Song HT and Xu WK (2016). TRIM16 suppresses the progression of prostate tumors by inhibiting the Snail signalling pathway. *International Journal of Molecular Medicine* **38**(6): 1734-1742.

Qin JJ, Wang W, Li X, Deokar H, Buolamwini JK and Zhang R (2018). Inhibiting β-Catenin by β-Carboline-Type MDM2 Inhibitor for Pancreatic Cancer Therapy. *Frontiers in Pharmacology* **9**: 5.

Sahebi M, Hanafi MM, van Wijnen AJ, Azizi P, Abiri R, Ashkani S and Taheri S (2016). Towards understanding pre-mRNA splicing mechanisms and the role of SR proteins. *Gene* **587**(2): 107-119.

Saijo M, Sakai Y, Kishino T, Nikana N, Mastuura Y, Morino K, Tamai K and Taya Y (1995). Molecular cloning of a human protein that 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.

Sanz G, Singh M, Peuget S and Selivanova G (2019). Inhibition of p53 inhibitors: progress, challenges and perspectives. *Journal of molecular cell biology* **11**(7): 586-599.

Scott RE, 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* **90**(1): 6-12.

Scott RE, White-Grindley E, Ruley, HE, Chesler EJ and Williams RW (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**(1): 99-105.

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.

Singh A, Upadhyay V, Upadhyay AK, Singh SM and Panda AK (2015). Protein recovery from inclusion bodies of Escherichia coli using mild solubilization process. *Microbial Cell Factories* **14**(1): 41.

Skepu A, Dlamini Z, Pretorius A, Liebrich W, Meyer M, Pugh DJR. and Rees DJG (2000). Evaluation of the mammalian DWNN gene product in apoptosis and CTL lysis. *Biochemical Society Transactions* **28**(5): 382-382.

Song LF, Lee TS, Zhu C, York DM and Merz Jr KM (2019). Using AMBER18 for Relative Free Energy Calculations. *Journal of Chemical Information and Modeling* **59**(7): 3128-3135.

Stenmark H and Aasland R (1999). FYVE-finger proteins-effectors of an inositol lipid. *Journal of Science* **112**: 4175-4183.

Sun J, Sun Y, Ahmed RI and Ren A (2019). Research Progress on Plant RING-Finger Proteins. *Genes* **10**(12): 973.

Tackmann NR and Zhang Y (2017). Mouse modelling of the MDM2/MDMX− p53 signalling axis. *Journal of Molecular Cell Biology* **9**(1): 34-44.

Tayyab S, Ahmad B, Kumar Y and Khan MM (2002). Salt-induced refolding in different domains of partially folded bovine serum albumin. *International Journal of Biological Macromolecules* **30**(1): 17-22.

Thwaites MJ, Cecchini MJ, Talluri S, Passos DT, Carnevale J and Dick FA (2017). Multiple molecular interactions redundantly contribute to RB-mediated cell cycle control. *Cell Division* **12**(1): 3.

Trott O and Olson AJ (2010). AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, effiecient optimization, and multithreading. *Journal of Computational Chemistry* **31**: 455–461.

Van Hooser AA, Yuh P and Heald R (2005). The perichromosomal layer. *Chromosoma* **114**(6): 377-388.

Vanneman M and Dranoff G (2012). Combining immunotherapy and targeted therapies in cancer treatment. *Nature Reviews Cancer* **12**(4): 237-251.

Vassilev L and Fry D editors (2011). Small-molecule inhibitors of protein-protein interactions *Springer Science & Business Media* **348**. Accessed 20/01/2021.

Vélez-Cruz R and Johnson DG (2017). The Retinoblastoma (RB) Tumor Suppressor: Pushing Back against Genome Instability on Multiple Fronts. *International Journal of Molecular Sciences* **18**(8): 1776.

Venkatesh D, O'Brien NA, Zandkarimi F, Tong DR, Stokes ME, Dunn DE, Kengmana ES, Aron AT, Klein AM, Csuka JM and Moon SH (2020). MDM2 and MDMX promote ferroptosis by PPARα-mediated lipid remodeling. *Genes and Development* **34**(7-8): 526-543.

Vo LTA, Minet M, Schmitter JM, Lacroute F and Wyers F (2001). Mpe1, a Zinc knuckle protein, is an essential component of yeast cleavage and polyadenylation. *Molecular and Cellular Biology* **21**(24): 8346-8356.

Wienken M, Dickmanns A, Nemajerova A, Kramer D, Najafova Z, Weiss M, Karpiuk O, Kassem M, Zhang Y, Lozano G and Johnsen SA (2016). MDM2 associates with polycomb repressor complex 2 and enhances stemness-promoting chromatin modifications independent of p53. *Molecular Cell* **61**(1): 68-83.

Winter A (2011). qtGrace. https://sourceforge.net/projects/qtgrace. Accessed 11 March 2021.

Witte MM and Scott RE (1997). The proliferation potential protein-related (P2P-R) gene with domains encoding heterogeneous nuclear ribonucleoprotein association and Rb1 binding shows repressed expression during terminal differentiation. *Proceedings of the National Academy of Sciences*. **94** (4): 1212-1217.

Xie Y and Varshavsky A (1999). The E2–E3 interaction in the N-end rule pathway: the RING-H2 finger of E3 is required for the synthesis of multiubiquitin chain. *EMBO Journal* **18**(23): 6832-6844.

Xue J, Huang X and Zhu Y (2019). Using molecular dynamics simulations to evaluate active designs of cephradine hydrolase by molecular mechanics/Poisson–Boltzmann surface area and molecular mechanics/generalized Born surface area methods. *RSC Advances* **9**(24): 13868-13877.

Yoshitake Y, Nakatsura T and Monji M (2004). Proliferation potential-related protein, an ideal oesophageal cancer antigen for immunotherapy, identified using complementary DNA microarray analysis. *Clinical Cancer Revised* **10**: 6437–6448.

Yu Y, Xu W, Wang S, Xu Y, Li H, Wang Y and Li S (2011). VpRFP1, a novel C4C4-type RING finger protein gene from Chinese wild Vitis pseudoreticulata, functions as a transcriptional activator in defence response of grapevine. *Journal of Experimental Botany* **62**(15): 5671-5682.

Zhang X, Perez-Sanchez H and C Lightstone F (2017). A comprehensive docking and MM/GBSA rescoring study of ligand recognition upon binding antithrombin. *Current Topics in Medicinal Chemistry* **17**(14): 1631-1639.

APPENDIX I

GENERAL CHEMICALS AND ENZYMES

Chemicals	Company
Glutathione	Merck
Ammonium chloride (NH_4Cl)	Merck
Sodium hydrogen phosphate (Na_2HPO_4)	Merck
Magnesium sulphate ($MgSO_4$)	Merck
Sodium azide (NaN_3)	Merck
Isopropanol	Merck
Tryptone	Merck
Yeast	Merck
Glycerol	Merck
Sodium chloride (NaCl)	Merck
Calcium chloride ($CaCl_2$)	Merck
Magnesium chloride ($MgCl_2$)	Merck
Potassium chloride (KCl)	Merck
Sodium phosphate ($NaHPO_4$)	Merck
Potassium dihydrogen phosphate (KH_2PO_4)	Merck

Tris[hydroxymethyl]amino ethane (Tris)	Merck
Glucose	Merck
Glycine	Merck
Coomasie Brillant Blue R-250	Merck
Ampicillin	Calbiochem
Phenyl Methyl Sulphonyl Fluoride (PMSF)	Calbiochem
Agarose gel powder	Calbiochem
Glutathione agarose beads	Sigma
Ethidium bromide	Sigma
N,N,N',N'-Tetramethylethylenediamine (TEMED)	Sigma
iso-octylphenoxypolyethoxyethanol (Triton X-100)	Sigma
BamHI	Thermo Scientific
HindIII	Thermo Scientific
DNA marker	Thermo Scientific
Ethylene diamine tetra Acetic Acid (EDTA)	Associated chemical enterprises
Glacial acetic acid	Associated chemical enterprises
Zinc sulphate ($ZnSO_4$)	Associated chemical enterprises
Ethanol	Associated chemical enterprises

Isopropylβ-D-thiogalactopyranoside (IPTG)	Melford
Sodium Dodecyl Sulphate (SDS) powder	Promega
Bis-acrylamide	Bio-rad
Ammonium sulphate (APS)	Bio-rad
Protein marker	Bio-rad
Lysozyme	Amersco
Dithiothreitol (DTT)	Amersco
Methanol	Radchem laboratory

APPENDIX II

AMINO ACIDS AND THEIR ABBREVIATIONS

Amino acid	3-letter abbreviation	1-letter abbreviation
Alanine	Ala	A
Arginine	Arg	R
Asparagine	Asn	N
Aspartic acid	Asp	D
Cysteine	Cys	C
Glutamic acid	Glu	E
Glutamine	Gln	Q
Glycine	Gly	G
Histidine	His	H
Isoleucine	Ile	I
Leucine	Leu	L
Lysine	Lys	K
Methionine	Met	M
Phenylalanine	Phe	F
Proline	Pro	P
Serine	Ser	S
Threonine	Thr	T
Tryptophan	Trp	W
Tyrosine	Tyr	Y
Valine	Val	V

APPENDIX III

Final decomposition binding free energy

| Run on Tue Sep 22 17:44:39 2020

|| GB non-polar solvation energies calculated with gbsa=2
idecomp = 1: Per-residue decomp adding 1-4 interactions to Internal.

Energy Decomposition Analysis (All unit's kcal/mol): Generalized Born solvent

Mdm2+Hit

DELTAS:

Total Energy Decomposition:

Residue

	Van der Waal	Electrostatic	Polar	Non-Polar	Total
GLY 33	-0.628	0.478	-0.451	-0.084	-0.685
ILE 36	-1.035	0.105	-0.3	-0.099	-1.329
MET 37	-1.613	-0.187	0.135	-0.122	-1.788
ARG 40	-0.49	-5.429	5.299	-0.075	-0.695
TYR42	-2.738	-0.459	0.788	-0.264	-2.672
LEU 114	-1.227	0.259	0.237	-0.193	-0.924
LEU 117	-0.821	-0.059	0.139	-0.049	-0.791
GLY118	-0.928	-0.6	0.547	-0.094	-1.075
ILE 121	-1.026	-0.317	0.358	-0.096	-1.081
MET 122	-1.586	-0.802	0.878	-0.198	-1.709
TYR 127	-2.487	-1.339	1.126	-0.237	-2.938
PHE 151	-0.305	-0.076	0.186	-0.018	-0.213
VAL 153	-2	0.162	0.004	-0.231	-2.065

Mdm2+RO-2443

DELTAS:

Total Energy Decomposition:

Residue

	Van der Waals	Electrostatic	Polar	Non-Polar	Total
GLY 33	-0.592	-0.109	0.371	-0.91	-0.42
ILE 36	-1.481	-0.269	0.045	-0.113	-1.818
MET 37	-1.677	-0.068	0.19	-0.131	-1.685
ARG 40	-0.652	0.766	-0.437	-0.074	-0.397
TYR 42	-3.766	-0.428	1.095	-0.283	-3.381
GLN 47	-1.288	-4.606	3.177	-0.146	-2.863
VAL 68	-0.584	0.161	-0.13	-0.111	-0.665
LEU 114	-1.116	-0.103	0.428	-0.194	-0.985
LEU 117	-0.606	-0.078	0.04	-0.038	-0.683
GLY 118	-0.645	-0.276	0.186	-0.084	-0.819
VAL 153	-1.867	-0.063	0.299	-0.242	-1.873

MdmX+Hit

DELTAS:
Total Energy
Decomposition:

Residue

	van der Waal	Electrostatic	Polar	Non-Polar	Total
MET 28	-1.611	0.325	0.437	-0.236	-1.086
LEU 31	-0.988	-0.356	0.309	-0.035	-1.07
GLY 32	-0.978	-1.371	1.008	-0.134	-1.475
ILE 35	-0.759	-0.518	0.434	-0.106	-0.949
TYR 41	-1.341	0.06	0.536	-0.17	-0.915
GLN 46	-0.875	-5.483	4.882	-0.139	-1.615
PHE 65	-0.538	-0.192	0.363	-0.046	-0.413
VAL 67	-2.064	0.58	-0.236	-0.347	-2.068
LEU 73	-0.885	-0.122	0.186	-0.115	-0.937
ILE 117	-1.016	0.777	-0.687	-0.082	-1.008
TYR 123	-4.132	-0.45	0.545	-0.35	-4.388
GLN 128	-1.236	-5.326	5.062	-0.12	-1.62

MdmX+RO-2443

DELTAS:

Total Energy Decomposition:

Residue

	Van der Waals	Electrostatic	Polar	Non-Polar	Total
ILE 35	-1.24	-0.227	-0.052	-0.102	-1.62
TYR 41	-2.557	-0.562	1.097	-0.201	-2.224
GLN 46	-0.928	-3.543	3.098	-0.068	-1.44
VAL 67	-0.685	0.13	-0.113	-0.115	-0.783
MET 110	-0.385	-0.086	0.259	-0.094	-0.306
GLY 114	-0.823	0.006	0.316	-0.144	-0.645
GLN 128	-2.403	-1.323	1.595	-0.23	-2.361
VAL 149	-2.188	0.566	-0.201	-0.31	-2.134