University of Zululand

Analysis of the biophysical and *in silico* exploration of the interaction between Small ribonucleoprotein G and the RING finger domain of Retinoblastoma binding protein 6 for the identification of potential anti-cancer compounds

by

Lloyd MABONGA (201639937)

B.Sc. (Hons) Food science and Nutrition (MSU), MTECH, Biotechnology (VUT)

Thesis submitted in fulfilment of the requirement for the degree of

Doctor of Philosophy (PhD) in Biochemistry

Department of Biochemistry and Microbiology, Faculty of Science and Agriculture,

UNIVERSITY OF ZULULAND

Supervisor: Prof. Abidemi Paul KAPPO

Co-Supervisor: Prof. Albertus Kotze BASSON

November 2021

DECLARATION

I declare that ***"Analysis of the biophysical and in silico exploration of the interaction between Small ribonucleoprotein G and the RING finger domain of Retinoblastoma binding protein 6 for the identification of potential anti-cancer compounds"*** is the product of my own work and effort. I have acknowledged all sources of information in line with normal academic conventions to the best of my knowledge. I further affirm that the research is original, and that the material submitted for examination is submitted for the degree of Doctor of Philosophy at the University of Zululand and has not previously been submitted for a degree at this or any other higher institution of learning.

Candidate: Lloyd Mabonga

Signature: …….

Date: …….

Supervisor: Prof. Abidemi Paul Kappo

Signature: …….

Date: …….

Co-Supervisor:Prof. Albertus Kotze Basson

Signature: …….

Date: …….

DEDICATION

This work is dedicated to my late sister Memory Mabonga who died of Cancer on the 2nd of July 2019. She was a pillar to my family. Losing her left us heartbroken. We will forever miss her. Tears may dry but MEMORIES will forever remain. We thank God for the life we shared.

May her soul rest in peace.

ACKNOWLEDGEMENTS

I wish to express my sincere gratitude to my supervisor, Prof A.P. Kappo for his excellent supervision and support. I am thankful for his academic guidance, research support, ceaseless motivation, constructive criticism, brilliant ideas, write-up and in-depth proof reading of this manuscript. Prof, thank you for being that 'one person' who always believed in my capabilities. Since the first day we met, you have been central to my life; a part of history that can never be wished away. I am sincerely grateful for the time we shared throughout the course of my studies. Thank you for your words of wisdom and counsel, and above all, for giving hope when I was at times hopeless.

Special mention is due to my better half, my darling, my friend, my wife; Naume Chinomwe Mabonga. You are a truly a virtuous woman, priced above rubies. Thanks for your care, encouragement, fellowship, love and understanding. Thanks for adding value in the past 8 years of my life. Above all, thanks for believing that this day would eventually come, and for holding together the home front while I was away. Your support is all a Father would need. I am greatly indebted to you Sweetheart. "Many daughters have done well but you excel them all" (Proverbs 31: 29). I cannot forget the two 'special ones', our two 'brilliant' boys (Divine Munashe Mabonga & Bethel Munyaradzi Mabonga). You are the best gift God has given me in this world. Kudos!

My heartfelt gratitude also goes to my parents, Mr E. Mabonga and my late mother Mrs C. Mabonga. Thanks for always being a source of reason and for always cheering me up to keep dreaming big. You have a place in my heart. I will never forget what you have sacrificed for me to be what I am today. I extend my sincere appreciation to my siblings; Slandiwe Mabonga,

Lindani Mabonga, Liberty Mabonga, Priscilla Mabonga, Melody Mabonga, Rudo Mabonga, Nomvelo Mabonga, Edmond Mabonga, Nyasha Mabonga and Remember Mabonga. Not forgetting the late Memory Mabonga, Thamsangqa Mabonga and Chipo Mabonga. You are the best family ever, let us keep united all the way. This achievement is because of your support. Thank you!

My heartfelt gratitude also goes to my BSB Team. I thank you for the pleasant moments we shared. Your support cannot go unmentioned. Thank you, a million times, for laying the foundation of my academic height and establishing it over the years with extraordinary encouragement and fun. Parting with you guys is a thorn in the flesh. I wish you all the best in your endeavours. Unmatchable special mention goes to my 'fellow researcher' Dr Priscilla Masamba. Your support, encouragement, resilience and the special refreshing moments we shared are much appreciated. "True friendship is never serene." I wish you all the best in your endeavors henceforth.

A very special gratitude goes to my late sister Memory Mabonga. I am greatly indebted for the support, encouragement, love, kindness you showed me until the last day I hugged you, kissed you and walked away from your deathbed with a sobbing heart; knowing I would never see you again. Dying of cancer left a deep scar in my heart. You always believed in my capabilities, you always had faith in my potential and loosing you at the end of my PhD left me heartbroken. However, I thank God for the good times we shared. I hope to meet someday, by and by. Finally, I thank my beloved God, Father and Friend for the great love and amazing grace. Jacob's Divine experience at Bethel (Genesis 28:10-22) is the reason why I look forward to that great day in 1 John 3vs 2. God, thank you for giving me life, wisdom, knowledge, intelligence, understanding, provisions, protection, family, and good health. May YAHWEH be glorified forever. Amen!

TABLE OF CONTENTS

LIST OF FIGURES

Chapter 2

Chapter 3

Chapter 4

Chapter 5

Chapter 6

APPENDIX I

LIST OF TABLES

Chapter 1

Chapter 2

Chapter 5

Chapter 6

LIST OF ABBREVIATIONS

3D	3-Dimensional
ADAM	A disintegrin and metalloproteinase
AKAP	A-kinase anchoring protein
AMPs	Antimicrobial Peptides
APC	Adenomatous polyposis coli
ATR	Ataxia telangiectasia and Rad3-related protein
Bcl-2	B cell lymphoma 2
BCL6	B cell lymphoma 6
B-CLL	B cell chronic lymphocytic leukaemia
Bcl-XL	B cell lymphoma-extra large
CADD	Computer-aided drug design
cAMP	Cyclic adenosine monophosphate
CBF	Core binding factor

CCR5	C-C chemokine receptor type 5
CDKN1A	Cyclin-dependent kinase inhibitor 1
CHK1	Checkpoint kinase 1
CRL	Cullin RING E3 ligase
CS	Computational solvent
CXCR4	C-X-C chemokine receptor type 4
CXCR7	C-X-C chemokine receptor type 7
DLBCL	Diffuse large B cell lymphoma
dnMAML1	Dominant-negative fragment of MAML1
DOS	Diversity-oriented synthesis
DPC	DNA-programmed chemistry
DWNN	Domain With No Name
FBDD	Fragment-based drug discovery
FBLD	Fragment-based lead discovery

FKBP12	FK-binding protein 12
FtsZ	Filamenting temperature-sensitive mutant Z
G-CSF	Granulocyte colony-stimulating factor
GPCR	G protein-coupled receptor
GSK-3	Glycogen synthase kinase 3
HDAC3	Histone deacetylase 3
HER2	Human epidermal growth factor receptor 2
HPV	Human papillomavirus
HTS	High-throughput screening
ICAM-1	Intercellular adhesion molecule-1
ICN1	Intracellular domain of NOTCH1
IL-2	Interleukin-2
IL-2Rα	Interleukin-2 receptor alpha chain
IPTG	Isopropyl β-D-1- thiogalactopyranoside

ITC	Isothermal Titration Calorimetry
K_D	Dissociation constant
LFA-1	Leukocyte integrin lymphocyte function-associated antigen 1
MAML	Mastermind-like
MDM2	Mouse double minute 2 homolog
MM	Multiple myeloma
MM/GBSA	Molecular Mechanics Generalized Born Surface Area
MST	MicroScale Thermophoresis
mTOR	Mammalian target of rapamycin
NMR	Nuclear magnetic resonance
PACs	Pancreatic acinar cells
PPI	Protein-protein interactions
Ptc1	Protein phosphatase 2C homolog 1
Phyre2	Protein Homology/analogy Recognition Engine 2

RBBP6	Retinoblastoma Binding Protein-6
RING	Really Interesting New Gene
SAHBs	Stabilised alpha-helix of Bcl-2 domains
SAR	Structure-activity relationship
SCLC	Small cell lung cancer
Shh	Sonic Hedgehog
SMRT	Silencing mediator for retinoid or thyroid-hormone receptors
SNRPG	Small Nuclear Ribonucleoprotein polypetide G
STATs	Signal transducers and activators of transcription
T-ALL	T cell acute lymphoblastic leukaemia
Tcf	T cell factor
TNF	Tumour necrosis factor
TNFR1	Tumour necrosis factor receptor 1
TNF-α	Tumour necrosis factor alpha

TPO	Thrombopoietin
TRAIL	TNF-related apoptosis-inducing ligand
Ub	Ubiquitin
UPP	Ubiquitin-proteasome pathway
UPS	Ubiquitin proteasome system
VEGF	Vascular endothelial growth factor
ZipA	Z interacting protein A

ABSTRACT

Cancer is the second leading cause of death globally after cardiovascular diseases, killing more people than HIV/AIDS, tuberculosis and malaria combined, which makes it a major public health concern. Increasing cancer morbidity and mortality rates has partly been due to the lack of specificity and side effects associated with most cancer drugs; hence over the years tremendous efforts have been linked to finding solutions to address these challenges. Novel techniques to design and develop cancer treatment methods using protein-protein interactions (PPIs) have become promising targets for therapeutic discovery. Suggestive evidence has proposed putative interactions between Small nuclear ribonucleoprotein polypeptide G (SNRPG) (also referred to as SmG in this study) and retinoblastoma binding protein 6 (RBBP6), which have been identified as potential diagnostic markers for cancer treatment. The broad focus of this study was to investigate the putative interactions between these two proteins. *In-silico* analysis and characterisation of the proteins using Autodock Vina revealed the binding and interaction patterns of SNRPG with the RBBP6 RING Finger domain with a docking score of -3.40kcal/mol. Using I-TASSER, a potential inhibitor known as (2R)-2-[(2-methyl-5-phenyl-pyrazol-3-yl) carbonyl amino]-3-napthalen-2-yl-propanoic acid) (4FI) was identified and MM/GBSA binding free energy analysis revealed a spontaneous reaction of SNRPG~RING Finger domain in complex with the inhibitor, due to a binding energy of -27.96kcal/mol. Some of the amino acid residues involved in the binding include Val222, Pro101 and Met194. Biophysical studies using MicroScale Thermophoresis (MST) confirmed the putative SNRPG~RING Finger domain interaction, and determined that the binding affinity was a K_d-value of 3.1596 nM under aqueous buffer conditions. The overall results from this study suggest the potential druggability of the SNRPG~RING Finger domain PPI. These findings will enhance our understanding in selective identification of small molecule inhibitors or peptides, which could be developed as novel therapeutic candidates in the diagnosis and treatment of cancer.

Keywords: Cancer, Peptidomimetics, Protein-protein interaction, retinoblastoma binding protein 6, Small ribonucleoprotein G

CHAPTER ONE

Introduction and Literature Review

1.1 General Introduction

Cancer remains one of the major public health challenges worldwide (Bhandari *et al*., 2014; Heneghan *et al.*, 2013). It kills more people globally than malaria, tuberculosis and HIV/AIDS put together (Jemal *et al*., 2010). Cancer is ranked second after cardiovascular diseases as the major cause of death globally (Ferlay *et al*., 2015). According to World Health Organization (2018) nearly 1 in 6 deaths worldwide is due to cancer (Du *et al.,* 2016). In 2008 an estimated 7.6 million people died of cancer globally and approximately 14.1 million new cancers cases were reported in 2012 resulting to an estimated 8.2 million deaths (Ferlay*et al*., 2015). In 2105 an estimated 8.8 million cancer deaths were reported globally with approximately 17.5 million new cancer cases reported (Du *et al.,* 2016). The number of new cancer cases is rapidly increasing and expected to rise by about 70% over the next two decades (WHO, 2018). Global economic costs of cancer are significantly rising above the projected US$1.16 trillion reported in 2010 (Jemal *et al*., 2011).

In Africa cancer is also a growing health challenge with increasing morbidity and mortality rates due to poor and limited health facilities and resources needed to prevent, diagnose, and treat cancer (Ferlay *et al*., 2015). In South Africa more than 100 000 people are diagnosed with cancer every year. And it is estimated that one in four people in South Africa is affected by cancer through diagnosis of family, friends or self. Cancer survival rate in South Africa is currently sitting on 2:3 (Jemal *et al*., 2010). In 2015, 114 091 new cancer cases were reported in South Africa leading to 58 237 cancer deaths in total (Randfontein herald, January 10, 2017). The worldwide increase of cancer burden is due to an aging and growing population, and the

adoption of economic development associated lifestyles and behaviors such as physical inactivity, unhealthy diets and smoking (Jemal *et al*., 2011). Lung, breast, prostate and colon cancer are the most commonly occurring cancers accounting for about 40 % of all cancers diagnosed. Breast cancer has been the most occurring having the highest mortality rate while lung cancer is the mostly diagnosed (Ferlay *et al*., 2015; Siegel *et al*., 2014, 2015).

Cancer often develops from uncontrolled cell division or proliferation of any kind of cells in the body (Hanahan & Weinberg, 2011). The cells divide when they are not expected to divide, continue dividing and do not undergo apoptosis. The uncontrolled cell cycle progression and inactivation of apoptotic mechanisms results from the switching off of genes that suppress tumors and/or switching on of oncogenes (Fearon & Vogelstein, 1999). In malignant cancers, cancer cells acquire enhanced cell motility (metastasis) due to the downregulation of cell adhesion receptors, activation of membrane metalloproteases and up-regulation of receptors (Fearon & Vogelstein, 1999; Hanahan & Weinberg, 2011). And this provides a physical pathway for the metastatic cancer cells to spread. However, there are various processes by which these genetic and cellular alterations occur (Hanahan & Weinberg, 2011).

The metastatic cancer cells move from their point of origin and invade other parts of the body, thereby forming malignant tumors and eventually causing death (Siegel *et al*., 2015). The canonical mechanisms are due to chromosomal translocation or deletion, mutation, genetic and chronic inflammation, dysregulated expression or activity of signaling pathways as well as exposure to risky lifestyle factors such as smoking, inadequate exercise, excessive alcohol intake, unhealthy diet, overweight and obesity (Hanahan & Weinberg, 2011; Jemal *et al*., 2011; Siegel *et al*., 2015). These events may activate oncogenes that promote dysregulated cell cycling and/or inactivate apoptotic pathways (Hanahan & Weinberg, 2011).

The challenge with cancer at the moment is the increasing morbidity and mortality rates, the lack of specificity of most cancer drugs and the side effects of cancer therapy drugs (Dardevet *et al.*, 2015). Over the years tremendous efforts have been linked to finding solutions to address the existing challenges. And protein-protein interaction (PPI) studies have emerged as promising targets for therapeutic discovery (Sable & Jois, 2015). PPI studies are not new in the field of smart drug development and drug design. They were known more than two decades ago however because of the wide interaction networks, small molecules that were targeted were often unsuccessful. This led to the view that targeting PPI had high risks and such targets were considered unfeasible (Sable & Jois, 2015).

However, this concept was challenged and now PPI have emerged as promising targets for therapeutic discovery (Sable & Jois, 2015). Currently there are drugs on the market (Fuller *et al*., 2009; Wishart *et al*., 2008; Xu *et al*., 2014). And potential drug-like candidates that target PPI are also in clinical trials (Domling, 2008). Inhibiting protein-protein interactions (PPI) using small molecules or peptides has both a controlling influence on biochemical pathways and therapeutic significance (Sable & Jois, 2015). The potential of PPI in therapeutics has prompted scientists to consider cancer-implicated protein-protein interaction networks as possible targets in alleviating the existing cancer challenges. Previous studies have provided encouraging proof on selective and efficient interruption of aberrant protein-protein interactions, opening up new avenues in the development of new cancer therapy drugs (Araujo *et al*., 2007; Fujii *et al*., 2007; Harris, 2006).

According to Johnson and coworkers (2006) most cancer-implicated proteins possess structural domains that have a higher ratio of infidelity as compared to non-cancer implicated proteins, making them more prone to interact with a wide diversity of proteins. Cancer-implicated proteins

have a large number of interacting proteins and occupy a central position in cancer-cell protein networks (Bhandari *et al*., 2014; Heneghan *et al*., 2013; Steinbrecher & Labahn, 2010). Protein interactions between cancer-implicated proteins have a higher probability of being related to the cancer processes than non-interacting proteins (Du *et al.*, 2016; Hanlon *et al*., 2010; Johnson *et al*., 2006).

Small nuclear ribonucleoprotein polypeptide G (SNRPG) and retinoblastoma binding protein 6 (RBBP6) are two good examples of proteins that have been implicated in cancer and whose functions are predominantly mediated by protein-protein interactions (Conte *et al*., 2002; Shi *et al*., 2009; Vo *et al*., 2001). SNRPG is a core-splicing protein essential in the biogenesis of small nuclear ribonucleoproteins (snRNPs) which are precursors of the spliceosome (Luhrmann *et al*., 1990; Palf *et al*., 2000; Stevens & Abelson, 1999; Will & Luhrmann, 2011). RBBP6 is a splicing-associated multi-domain and multi-functional nuclear protein (Pugh *et al*., 2006).

SNRPG and RBBP6 proteins are known to localize in the nucleolus and nuclear speckles and appear to have a closely linked role in pre-mRNA splicing (Hull *et al*., 2015). Varying expression levels of the two proteins have been reported in different types of cancer which include breast cancer, lung cancer, prostate cancer and colon cancer (Hull *et al*., 2015; Khan *et al*., 2014; Li *et al.*, 2007; Simons *et al.*, 1997). However, very little is known about the putative interactions between SNRPG and RBBP6 in different types of cancer.

Chibi and coworkers (2008) predicted possible interactions between SNRPG and the N terminal domains of RBBP6. The N terminal domains of RBBP6 consist of the Domain With No Name (DWNN), the Zinc knuckle and Really Interesting New Gene (RING) finger domain). Using the yeast 2-hybrid technique, Chibi and co-workers (2008) suggested that RBBP6 interacts with

SNRPG through its DWNN domain which is a crucial component of the RNA processing machinery in the cell. In another study, Kappo and co-workers (2012) also identified two copies of SNRPG (conformational isomers of the same protein) as part of the five substrates that bind to the N-terminal domain of RBBP6 (Kappo *et al.*, 2012). The findings substantiate the involvement of SNRPG with the N terminal domains of RBBP6 in tumorigenesis and tumour development. However, the precise mechanisms involved remain elusive.

Smith (Sm) proteins

The human spliceosomal Smith (Sm) proteins belong to a larger family of Sm and Sm-like (LSm) proteins. They are crucial elements in the biogenesis of the four major spliceosomaluridyl-rich small nuclear ribonucleoprotein particles (U snRNPs) (Schwer *et al.*, 2016; Ohtani 2018). U snRNPs are precursors of both the major and minor spliceosome which catalyze the essential RNA processing events, including the splicing of pre-mRNA introns and exon ligation. Thus, U snRNPs facilitate a crucial step in the formation of translatable mRNA (Schwer *et al.*, 2016; Weber *et al.*, 2010; Matera & Wang 2014). U snRNPs consist of snRNA (U1, U2, U4/U6 and U5) and numerous distinct proteins that are classified into two groups. The first group comprise of snRNP-specific proteins which facilitate snRNP-specific functions in pre-mRNA splicing. The second group consists of common snRNP proteins. These common proteins are collectively referred to as Sm proteins. They are originally identified from their antigenicity in the autoimmune disease systemic lupus erythematosus. The best characterized of these common Sm proteins are B' (29 kDa), B (28 kDa), D1 (16 kDa), D2 (16.5 kDa), D3 (18 kDa), E (12 kDa), F (11 kDa) and G (9 kDa). B and B′ are alternatively spliced products of the same gene (Kambach *et al.*, 1999; Stark *et al.*, 2001; Urlaub *et al.*, 2001).

Sm proteins are small proteins characterized by the presence of the highly conserved Sm fold containing two conserved motifs each of which consists largely of antiparallel β sheets β5↑•β1↓•β2↑•β3↓•β4↑ topology and an embedded RNA-binding site (shown in Figure 1.1A) (Hermann *et al*., 1995). The motifs are separated by a ‘‘linker’’ region that varies in length among different Sm proteins (Khusial *et al*., 2005; Hermann *et al*., 1995). When a single Sm protein is folded, both RNA-binding sites are found within loops located in proximity on one side of the protein as shown in Figure 1.1B for the human SNRPD1 protein. When an entire Sm ring is assembled, these loops are oriented toward the central pore of the ring (Kambach *et al*., 1999; Hermann *et al*., 1995). Sm subunits are decorated by additional unstructured C terminal extensions and secondary structure elements. The Sm motif encodes for a common folding domain (Sm domain) that is responsible for mediating the Sm-Sm protein interactome through the antiparallel β strands (Kambach *et al*., 1999). Sm proteins also possess two solvent-exposed hydrophobic interaction surfaces that are prone to nonspecific interactions under physiological conditions (Kambach *et al*., 1999; Pellizzoni *et al*., 2002; Kroiss *et al*., 2008; Zhang *et al*., 2011; Grimm *et al*., 2013; Neuenkirchen *et al*., 2015).

The morphogenesis of the spliceosomal U snRNPs depends, to a high degree, on the cytoplasmic Sm-Sm protein interactome and Sm protein-RNA interactome (Mura 2002). Binding of the Sm core proteins to the Sm site of the newly transcribed snRNAs in the cytoplasm is a prerequisite for the hypermethylation of the snRNA cap structure as well as for transport of the mature snRNP particle into the nucleus (Meister *et al*., 2001). In addition to the Sm-Sm protein interactome and Sm protein-RNA interactome, a significant fraction of Sm-proteic interactome has been reported (Kambach *et al*., 1999; Pellizzoni *et al*., 2002; Kroiss *et al*., 2008; Zhang *et al*., 2011; Grimm *et al*., 2013; Neuenkirchen *et al*., 2015; Hyjek *et al*., 2019).

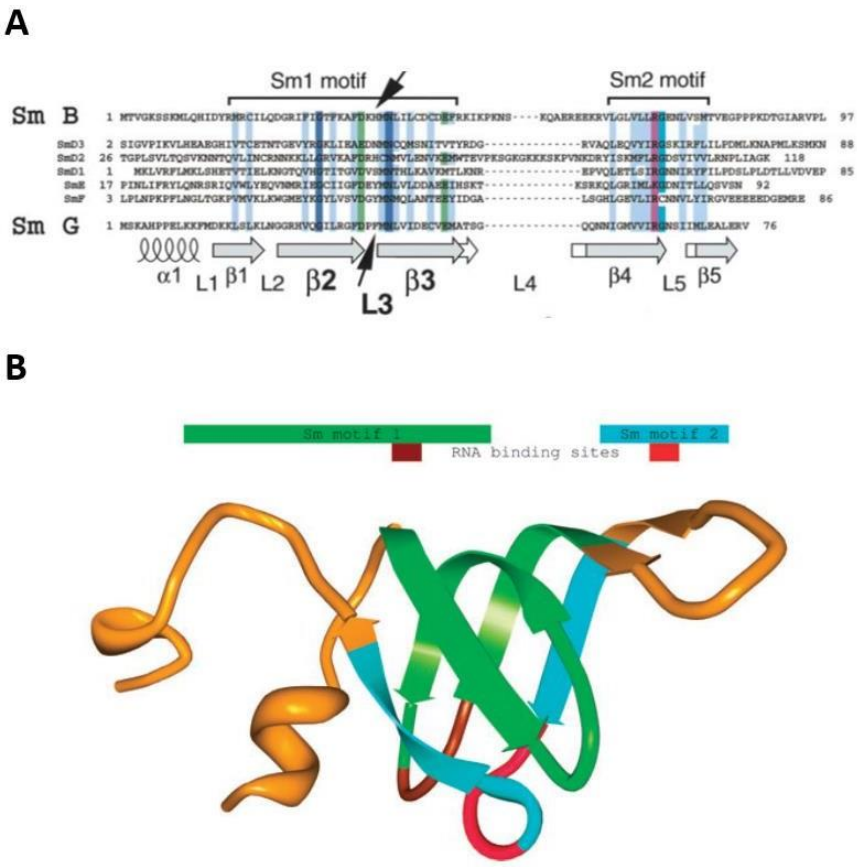


Figure 1.1: Human Sm protein primary and model structures (**A**) Sm protein primary structure alignment showing the Sm1 and Sm2 motifs. Conserved amino acids are highlighted as follows; Light blue (uncharged hydrophobic residues), Green (acidic amino acids), Purple (basic amino acids), Dark blue (100% conserved amino acids) and Turquoise (80% conserved glycine). Arrows mark the cross-linked amino acids in the protein sequences as identified by N-terminal sequencing. The cross-linking sites are located within loop L3 of the Sm1 motif. Figure extracted from Hermann *et al*., 1995. (**B**) Human SNRPD1 protein structural model (PDB ID 1B34). Green, light blue, brown, and red regions of the model correspond to the Sm motif and RNA-binding site sequences. Figure extracted from Kambach *et al*., 1999.

The interactome relates to ribosomes, translation, mitochondria, energy metabolism and, RNA- and DNA-binding proteins implicated in chromatin organization, regulation of transcription, translation and RNA metabolism. The Sm-proteic interactome is involved in the regulation of cell development where they perform essential roles in cell specification and gene expression (Hyjek *et al*., 2019).

Smith (Sm) proteins and the biogenesis of U snRNPs

Sm proteins play a presumably very critical role in the biogenesis of U snRNPs in normative cellular processes and disease states. They are the core elements of snRNPs (U1, U2, U4/U6, and U5 snRNPs), spliceosome (both minor and major) and subsequently play a critical role in the excision of introns and ligation of exons in pre-mRNA splicing (Montzka & Steitz, 1988). RNA processing events (on a cellular scale) can be traced back hierarchically to the Sm proteins (on a molecular scale) (see Figure 1.2) (Mura 2002). Placing the Sm protein family in a biochemical context underscores their central importance in RNA metabolism. To cope with the task of splicing virtually all intron-containing pre-mRNAs, the homeostatic control of Sm proteins is critical (Montzka & Steitz, 1988). The efficient production of the core components of the U snRNPs, maintenance of their steady-state levels and protraction of their proper functional interactome landscape appear to be indispensable in higher eukaryotes (Montzka & Steitz, 1988).

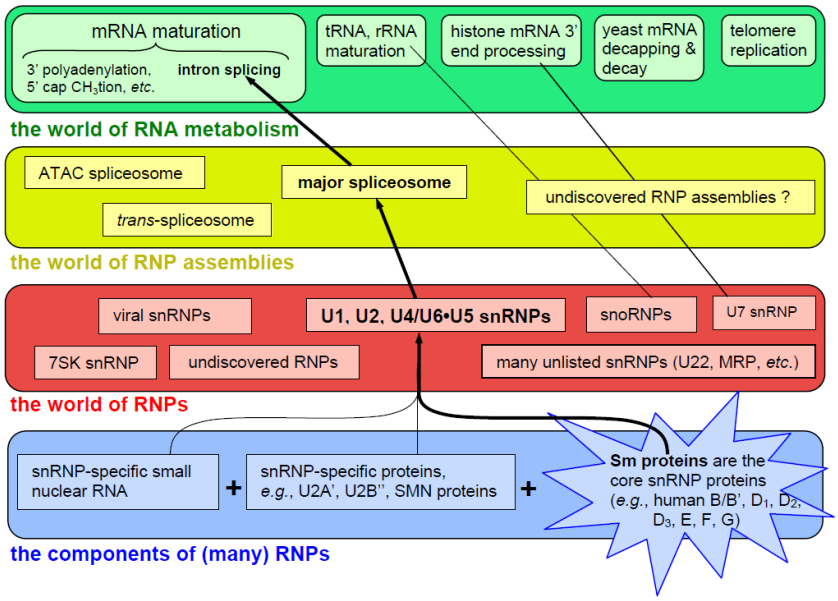


Figure 1.2: Sm proteins and the modular approach to RNA metabolism. Sm proteins are core elements of snRNPs (U1, U2, U4/U6, and U5 snRNPs), major spliceosome and subsequently play a critical role inthe excision of introns in pre-mRNA splicing. This figure is for illustrative purposes only. It is by no means comprehensive. Figure extracted from Mura (2002).

Sm proteins are synthesized in the cytoplasm by ribosomes translating Sm messenger RNA (Meister *et al*., 2001). They are stored in the cytoplasm in the form of three partially assembled ring complexes associated with the assembly chaperone pICln protein, namely, the 6S pentamer complex comprising of SNRPD1, SNRPD2, SNRPF, SNRPE and SNRPG with pICln; a 2-4S complex of SNRPB, possibly with SNRPD3 and pICln; and the 20S methylosome, which is a large complex of SNRPD3, SNRPB, SNRPD1, pICln and the arginine methyltransferase-5 (PRMT5) protein (shown in Figure 1.3A) (Meister *et al*., 2001). pICln acts as a specialized chaperone protein recruiting all newly synthesized Sm proteins to the PRMT5 complex and preventing the premature assembly of Sm proteins with snRNA (Meister *et al*., 2001). Furthermore, the formed and methylated 6S and pICln-D3/B interact with the SMN complex facilitating the transfer of the Sm proteins onto the SMN complex (1) (shown in Figure 1.3B). Gemin2 coordinates the binding of five Sm proteins (D1/D2/F/E/G) (Zhang *et al*., 2011). According to Chari and coworkers D3/B presumably binds to SMN, Gemin2 and/or Gemin8 (Neuenkirchen *et al*., 2015).

Concurrently, pre-U snRNA is transcribed by RNA polymerase II (pol II) and m7G-capped in the nucleus (Figure 1.4, step 1) (Neuenkirchen *et al*., 2015; Prusty *et al*., 2017). The transcribed pre-U snRNA is complexed with CBC, PHAX, CRM1 and RanGTP to form an export complex that is actively transported into the cytoplasm via the NPC (Figure 1.4, step 2). In the cytoplasm the export factors and pre-U snRNA disassemble (step 3) and Sm proteins provided by the SMN complex are assembled onto the ‘‘Sm-site’’ of pre-U snRNA (Figure 1.4(step 4) and Figure 1.4B (2)). The SMN complex prevents the association of Sm proteins with nontarget RNAs (Neuenkirchen *et al*., 2015; Prusty *et al*., 2017). After recruitment by the SMN-complex and Sm core domain the import factors SPN1 and importin β mediate translocation of the complex into

the nucleus (step 6). In the nucleus, both factors dissociate and are recycled into the cytoplasm (step 7) and the U snRNPs enrich in Cajal bodies (step 8). Following scaRNA guided pseudouridylation (ψ) and 2'-O-methylation (m; Figure 1.4, step 9) the mature U snRNP is channeled to the spliceosome (step 10), however the SMN-complex is exported into the cytoplasm (Figure 1.4, step 11) where it re-enters the biogenesis cycle (Figure 1.4, step 12) (Neuenkirchen *et al*., 2015; Prusty *et al*., 2017).

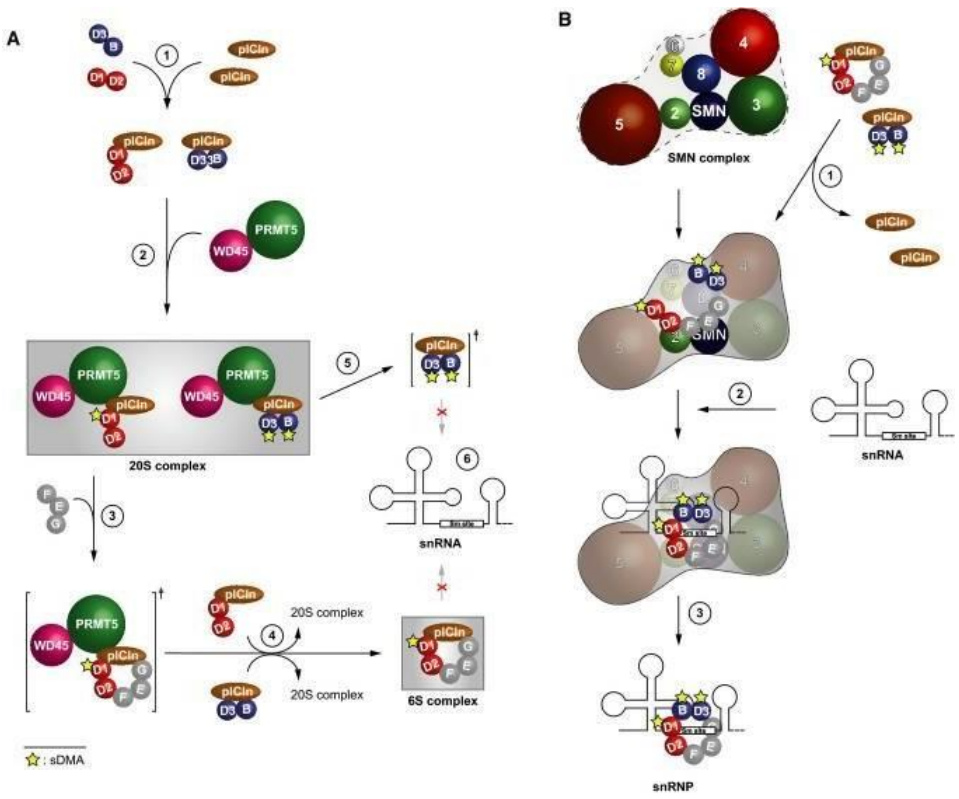


Figure 1.3: Model of Sm protein assembly. (**A**) Sm proteins are initially translated in the cytoplasm. Heterooligomers D1/D2 and D3/B interact with pICln to form pICln-Sm protein complexes (1) and recruited to PRMT5/WD45 forming a 20S complex (2). The addition of F/E/G on the 20S complex results in the formation of a ring-shaped 6S assembly intermediate (3). Further supplementation of pICln-Sm protein complexes releases 6S and again establishes a 20S complex (4). In either case, pICln serves as a molecular chaperone and prevents the interaction of the associated Sm proteins with snRNA (6). (**B**) The initially formed and methylated 6S and pICln-D3/B interact with the SMN complex and the Sm proteins are transferred onto the SMN complex (1). Gemin2 coordinates the binding of five Sm proteins (D1/D2/F/E/G), while D3/B is likely bound to SMN, Gemin2 and/or Gemin8. The snRNA associates with the loaded SMN complex and the Sm proteins are assembled around the Sm site (2). The complex is translocated to the nucleus where the mature snRNP dissociates from the SMN complex (3). Figure extracted from Neuenkirchen *et al*., 2015.

In accordance with this view the efficient production of Sm proteins, maintenance of their steady-state levels and perpetuation of their proper functional interactome landscape appear indispensable in mRNA metabolism. The accurate regulation of Sm proteins during the biogenesis of U snRNPs is a key aspect in the normative functioning of eukaryotic cells (Prusty *et al*., 2017). Kinetic defects in the U snRNP assembly pathway, which affect the rate of U snRNP formation, lead to a reduction in the levels of functional U snRNPs in the cell, which subsequently results to a disease state known as "U snRNP assembly disease" (Shukla & Parker 2014; Neuenkirchen *et al*., 2015; Prusty *et al*., 2017).

Cells use a plethora of PPI networks to help maintain the normative functions of Sm proteins in small nuclear ribonucleoprotein (snRNP) assembly, biogenesis of U snRNPs and the smooth running of the splicing machinery. Deregulation of spliceosomal Sm proteins has been significantly implicated in pathophysiological cues and disease states including cancer (Marabti & Younis 2018; Blijlevens *et al*., 2019).

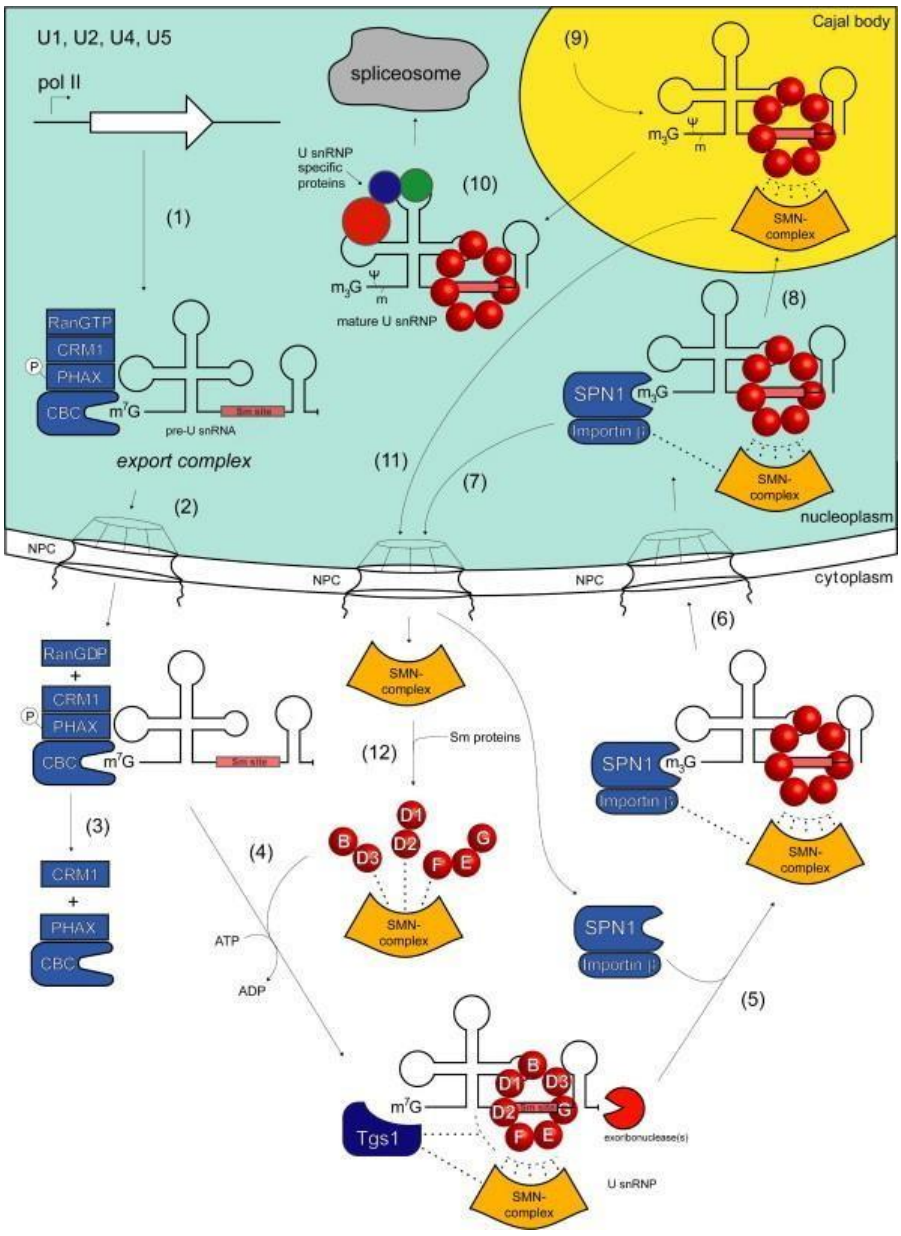


Figure 1.4: Sm proteins and the biogenesis pathway of spliceosomal U snRNPs. Sm proteins provided by the SMN-complex are assembled onto the "Sm-site" of pre-U snRNA (step 4). Following recruitment by the SMN-complex and Sm core domain, the hypermethylase Tgs1 modifies the m^7G-cap to m_3G (step 5), before the import factors SPN1 and importin β mediate translocation into the nucleus (step 6). There, both factors dissociate and are recycled into the cytoplasm (step 7), and U snRNPs associated with the SMN-complex enrich in Cajal bodies (step 8). After scaRNA guided pseudouridylation (Ψ) and 2′-*O*-methylation (m; step 9), the mature U snRNP is directed to the spliceosome, (step 10), whereas the SMN-complex is believed to be exported into the cytoplasm (step 11), where it can re-enter the biogenesis cycle (step 12) (Figure extracted from Neuenkirchen *et al.*, 2008).

Defects in Sm protein complex binding to snRNAs are known to reduce levels of snRNAs, suggesting an unknown quality control system for small nuclear ribonucleoprotein (snRNP) assembly (Shukla & Parker, 2014). In case of exhaustion or improper functioning of the assembly pathway, the cell activates fail-safe measures, including targeted autophagosome-mediated Sm protein degradation and exosome-processed Sm-encoding transcript degradation. These measures refine cellular quality control mechanisms to prevent proteotoxicity during imbalances in UsnRNP assembly. Dysregulation of Sm proteins during U snRNP assembly causes cellular proteotoxicity. Early phase plCln deficiency leads to degradation of Sm proteins via autophagy and mislocalisation of unassembled and/or misassembled Sm proteins. Late phase SMN deficiency leads to accumulation of Sm proteins over plCln (Prusty *et al*., 2017).

Sm proteins and Cancer

Sm proteins play an indispensable clinical and prognostic regulatory role in pathophysiological cues and disease progression including cancer (Marabti & Younis 2018; Blijlevens *et al*., 2019). They are master regulators of both co- and post-transcriptional assemblies, making them an attractive target in PPI-focused drug discovery technology (Valeur *et al*., 2019; Sorolla *et al*., 2019; Haymond *et al*., 2019). One example of this class of diseases is a neurodegenerative disease, spinal muscular atrophy (SMA), caused by ubiquitous deficiency in the survival motor neuron (SMN) protein due to mutations in the principal SMN-coding gene SMN1 leading to reduced snRNA and snRNP levels (Lorson *et al*., 2010; Lefebvre *et al*., 1995). One role of the SMN complex is to load the Sm protein complex onto the Sm site on snRNAs, which has a consensus sequence of PuAU4–6GPu (Battle *et al*., 2006; Fischer *et al*., 2011; Neuenkirchen *et al*., 2015).

Van Alstyne and coworkers added that the function of SMN in the assembly of spliceosomal Sm proteins regulates alternative splicing of Mdm2 and Mdm4, two non-redundant repressors of p53 (Alstyne *et al*., 2018). Decreased inclusion of critical Mdm2 and Mdm4 exons is most prominent in SMA motor neurons and correlates with both U snRNPs reduction and p53 activation *in vivo*. The findings reveal that loss of SMN-dependent regulation of Mdm2 and Mdm4 alternative splicing underlies p53-mediated death of motor neurons in SMA, establishing a causal link between U snRNPs dysfunction and neurodegeneration (Alstyne *et al*., 2018; Neuenkirchen *et al*., 2015; Fischer *et al*., 2011; Lorson *et al*., 2010; Battle *et al*., 2006; Lefebvre *et al*., 1995).

Sm proteins have attracted significant attention because of their implicated roles in tumorigenesis and tumor development (Marabti & Younis 2018; Blijlevens *et al*., 2019). Accumulating evidence suggests that Sm proteins contribute significantly to the initiation and progression of many disease states including cancer (Blijlevens *et al*., 2019). Varying expression levels have been reported in different types of cancers, which include breast cancer, lung cancer, prostate cancer, liver cancer and colon cancer (Hull *et al*., 2015; Ezkurdia *et al*., 2014; Khan *et al*., 2014; Li *et al*., 2007; Conte *et al*., 2002; Simons *et al*., 1997). The interference of Sm protein expression has been shown to induce apoptosis in non-small-cell lung cancer (NSCLC) cells (Blijlevens *et al*., 2019). For several cancers, increased expression levels of Sm proteins showed a positive correlation with disease severity establishing a causal link between Sm proteins and, tumorigenesis and tumor development. Immunohistochemical studies illustrating antibody staining of standard cancer tissues samples have highlighted the localization of Sm proteins in tumor cells. Tumor heterogeneity and inter-individual differences in variable immunohistochemical staining patterns reflected the diverse expression of Sm proteins (Ezkurdia

et al., 2014). Hence, Sm proteins present a particularly useful novel target for diagnostic and therapeutic intervention in cancer.

Using semi-quantitative RT-PCR, Anchi and co-workers reported the involvement of Sm proteins in cell proliferation and progression of high-grade prostate cancer through the regulation of androgen receptor expression (Anchi *et al*., 2012). The study highlighted that SmE over-expression promoted prostate cancer cell proliferation in high-grade prostate cancer cells compared with normal prostatic epithelial cells, indicating its oncogenic effects (Anchi *et al*., 2012). SmE knockdown expression by short interfering RNA (siRNA) resulted in the marked suppression of prostate cancer cell proliferation (Anchi *et al*., 2012).

The study further elucidated that the regulation of androgen receptor expression by SmE is essential for cell proliferation and progression of high-grade prostate cancer. In this context SmE presents a novel molecular target for cancer drugs (Anchi *et al*., 2012). However, targeting these splicing machinery components as drug development tools in cancer has been an overlooked and underdeveloped strategy over the years. Very little is known about their putative interactions in cancer-cell protein networks and their roles in different types of cancers remain elusive.

In addition to that Sm proteins have been shown to interact with the oncogenic multi-domain and multifunctional protein, RBBP6 (Kappo *et al*., 2012; Chibi *et al*., 2008; Simons *et al*., 1997). Using immunoblot analysis, Simons and co-workers concluded that Sm proteins interact particularly with the N-terminal domains of RBBP6 (Simons *et al*., 1997). The findings suggest possible connections between Sm proteins and tumor suppressor proteins. Sm proteins may directly influence tumor suppressor proteins in pre-mRNA splicing. Chibi and co-workers also postulated that SNRPG interacts with the N-terminal domain of RBBP6 which happens to be a

crucial component of the RNA processing machinery in the cell (Chibi *et al*., 2008). Kappo and co-workers further identified two copies of the SNRPG protein (conformational isomers of the same protein) as part of the five substrates that bind to the N-terminal domains of RBBP6 (Kappo *et al*., 2012). Conte and co-workers also identified SNRPG as one of the two potential binding partners for oncogenic TACC1 in breast cancer (Conte *et al*., 2002). The findings suggested a possible link between Sm proteins and TACC1 in the control of mRNA metabolism.

Conte and co-workers speculated that down-regulation of TACC1 may alter the control of mRNA homeostasis in polarized cells and subsequently participate in the oncogenic processes (Stark *et al*., 2006). Considering the oncogenic nature of the RBBP6 and TACC1 protein, the findings suggest that Sm proteins possibly play a role in pRb/p53 pathways, tumorigenesis and the progression of many cancers (Chibi *et al*., 2008; Simons *et al*., 1997). However, the precise mechanisms remain elusive and yet to be understood.

Inhibiting the interaction between Sm proteins and active oncogenic proteins using small molecules or peptides may modulate cancer-cell networks and have therapeutic significance. Thus, investigating Sm protein interactomes and their binding events in cancer-cell protein networks may help identify new avenues to design and develop Sm PPI-focused therapeutic drugs. Currently, the physiological relevance and functional basis of these different Sm proteins in cancer-cell networks still remain obscure up to this day. Very little is known about the nontarget Sm PPI in cancer cells. Sm proteins play a yet uncharacterized role in linking splicing machinery components to tumorigenesis and tumor development in various cancers. Perhaps understanding their interactome in cancer-cell networks may lead to new knowledge on how this family of proteins relate to tumorigenesis and tumour development. Perhaps finding small molecule inhibitors for Sm PPI in cancer cells may present potential drug targets in diagnostics

and therapeutic studies. Thus, new avenues to design and develop therapeutic drugs may be established.

Small nuclear ribonucleoprotein polypeptide G (SNRPG)

SNRPG is an 8.5 kDa splicing associated and cancer implicated Smith (Sm) protein whose functions are predominantly mediated by protein-protein interactions. Sm proteins are core-splicing proteins common to each uridyl-rich small nuclear ribonucleoprotein particles (U snRNP) particle (Weber *et al*., 2010). U snRNP complexes (U1, U2, U4 and U5) are precursors of the spliceosome. They catalyze the removal of intervening (non-coding) sequences from pre-mRNA (Green, 1986; Wahl *et al*., 2009). Sm proteins are named in order of their decreasing size are B' (29 kDa), B (28 kDa), D3 (18 kDa), D2 (16.5 kDa), D (16 kDa), E (12 kDa), F (11 kDa) and G (8.5 kDa) (Luhrmann *et al*., 1990; Will & Luhrmann, 2011).

SNRPG has been reported in more than 20 different cancers that include breast cancer, lung cancer, prostate cancer and colon cancer (Conte *et al*., 2002). According to Stark *et al*. (2006) SNRPG has a very wide interaction network comprising of more than 138 interactions with more than 115 identified interactors. Using the two-hybrid screen in yeast, GST pull-downs and co-immunoprecipitations Cont *et al*. (2002) identified SNRPG as one of the two potential binding partners for the oncogenic protein TACC1 in breast cancer. Previous research has identified SNRPG as a component of SMN-Sm protein complex, U1 snRNP, U2 snRNP, U12 type spliceosomal complex, U4 snRNP, U7 snRNP, U5 snRNP, spliceosomal tri-snRNP complex, catalytic step 2 spliceosome, Cytosol, methylosome, nucleoplasm, small nuclear ribonucleoprotein complex and spliceosomal complex (Stark *et al*., 2006).

The SNRPG encoding gene in humans is found on chromosome 2p13.3 and is made up of 8 exons. The gene comprises 455 nucleotides with an open reading frame encoding a predicted protein of 76 amino acid residues. SNRPG has a theoretical pl of 8.9 (Gasteiger *et al*., 2005). The percentage amino acid composition is shown in Table 1. According to Hermann *et al*. (1995) the SNRPG protein translated in vitro from a single SNRPG mRNA migrates as a doublet by high-TEMED SDS-PAGE. Hermann and co-workers (1995) suggested that the 2 bands represent conformational isomers of the same protein. Northern blot analysis revealed that the SNRPG gene is expressed as an approximately 0.5-kb mRNA in HeLa cells (Hermann *et al*., 1995). The SNRPGDNA coding sequence and amino acid sequence are shown in Figure 1.5 and 1.6.

```
  1 caacaatcca ttggctaacg aggtcgatga cgccagacgc aagacgccgg gcctacagcg
 61 ggagcgtgag gaaagccgtg cgttgcgttc caaggcatct gtgagcccgc ggagtataca
121 ccatgagcaa agctcaccct cccgagttga aaaaatttat ggacaagaag ttatcattga
181 aattaaatgg tggcagacat gtccaaggaa tattgcgggg atttgatccc tttatgaacc
241 ttgtgataga tgaatgtgtg gagatggcga ctagtggaca acagaacaat attggaatgg
301 tggtaatacg aggaaatagt atcatcatgt tagaagcctt ggaacgagta taaataatgg
361 ctgttcagca gagaaaccca tgtcctctct ccatagggcc tgttttacta tgatgtaaaa
421 attaggtcat gtacattttc atattagact ttttgttaaa taaacttttg taatagtcaa
481 aaatgctttc tcagatgttc tgaatataga atatcagctc tcattccagt tttttctaac
541 atgaattttc ctggttgaca ttgatttcaa agggttttat gcattaaagt gaaagaatct
601 tattaaatgt gaaacatggc aaggaaaaaa
```

Figure 1.5: SNRPG DNA coding sequence (CDS)

```
atgagcaaagctcaccctcccgagttgaaaaaatttatggacaagaagttatcattgaaa
 M  S  K  A  H  P  P  E  L  K  K  F  M  D  K  K  L  S  L  K
ttaaatggtggcagacatgtccaaggaatattgcggggatttgatccctttatgaaccttt
 L  N  G  G  R  H  V  Q  G  I  L  R  G  F  D  P  F  M  N  L
gtgatagatgaatgtgtggagatggcgactagtggacaacagaacaatattggaatggtg
 V  I  D  E  C  V  E  M  A  T  S  G  Q  Q  N  N  I  G  M  V
gtaatacgaggaaatagtatcatcatgttagaagccttggaacgagtataa
 V  I  R  G  N  S  I  I  M  L  E  A  L  E  R  V  -
```

Figure 1.6: SNRPG amino acid sequence

SNRPG consists of a total number of 8 negatively charged residues (Asp + Glu) and 10 positively charged residues (Arg + Lys) giving a positive net charge of 2 (Gasteiger *et al*., 2005). The protein does not contain any tryptophan residues and has a computed extinction coefficient of zero (0) assuming all pairs of Cysteine residues form cystines (oxidised dimers) and are reduced (Gasteiger *et al*., 2005). Considering the methionine on the N-terminal of the sequence, SNRPG has an estimated in vivo half-life of 30 hours, >20 hours and >10 hours in mammalian reticulocytes, yeast and *Escherichia coli* respectively (Gasteiger *et al*., 2005).

SNRPG protein is a stable protein with a computed instability index (II) of 35.06 (Gasteiger *et al*., 2005). The protein has an aliphatic index of 98.68 suggesting that the protein has a relative volume occupied by aliphatic side chains (alanine, valine, isoleucine and leucine) (Gasteiger *et al*., 2005). SNRPG has a grand average hydropathicity (GRAVY) of 0.066. The hydropathy index relates to the number of hydrophobic or hydrophilic properties of the side chains. The larger the number, the more hydrophobic is the protein (Gasteiger *et al*., 2005).

SNRPG is rich in hydrophobic amino acids (about 34%) most of which are located in the C-terminal half (Gasteiger *et al*., 2005; Hermann *et al*., 1995). The N-terminal portion of the sequence contains a larger number of positively charged residues while the rest of the sequence contains a larger number of negatively charged residues (Gasteiger *et al*., 2005; Hermann *et al*., 1995). The predictive physicochemical parameters of SNRPG make it feasible to carry out investigations on the putative interactions between SNRPG and the RBBP6 RING Finger domain.

Table 1.1: Percentage amino acid composition for SNRPG (Gasteiger *et al*., 2005)

Amino acid	Symbol	Number	% composition
Alanine	(A)	3	3.9%
Arginine	(R)	4	5.3%
Asparagine	(N)	5	6.6%
Aspartic acid	(D)	3	3.9%
Cysteine	(C)	1	1.3%
Glutamine	(Q)	3	3.9%
Glutamic acid	(E)	5	6.6%
Glycine	(G)	7	9.2%
Histidine	(H)	2	2.6%
Isoleucine	(I)	6	7.9%
Leucine	(L)	8	10.5%
Lysine	(K)	6	7.9%
Methionine	(M)	6	7.9%
Phenylalanine	(F)	3	3.9%

Proline	(P)	3	3.9%
Serine	(S)	4	5.3%
Threonine	(T)	1	1.3%
Tryptophan	(W)	0	0.0%
Tyrosine	(Y)	0	0.0%
Valanine	(V)	6	7.9%
Hydroxyproline	(O)	0	0.0%
Pyroglutamatic	(U)	0	0.0%

Retinoblastoma-binding protein-6 (RBBP6)

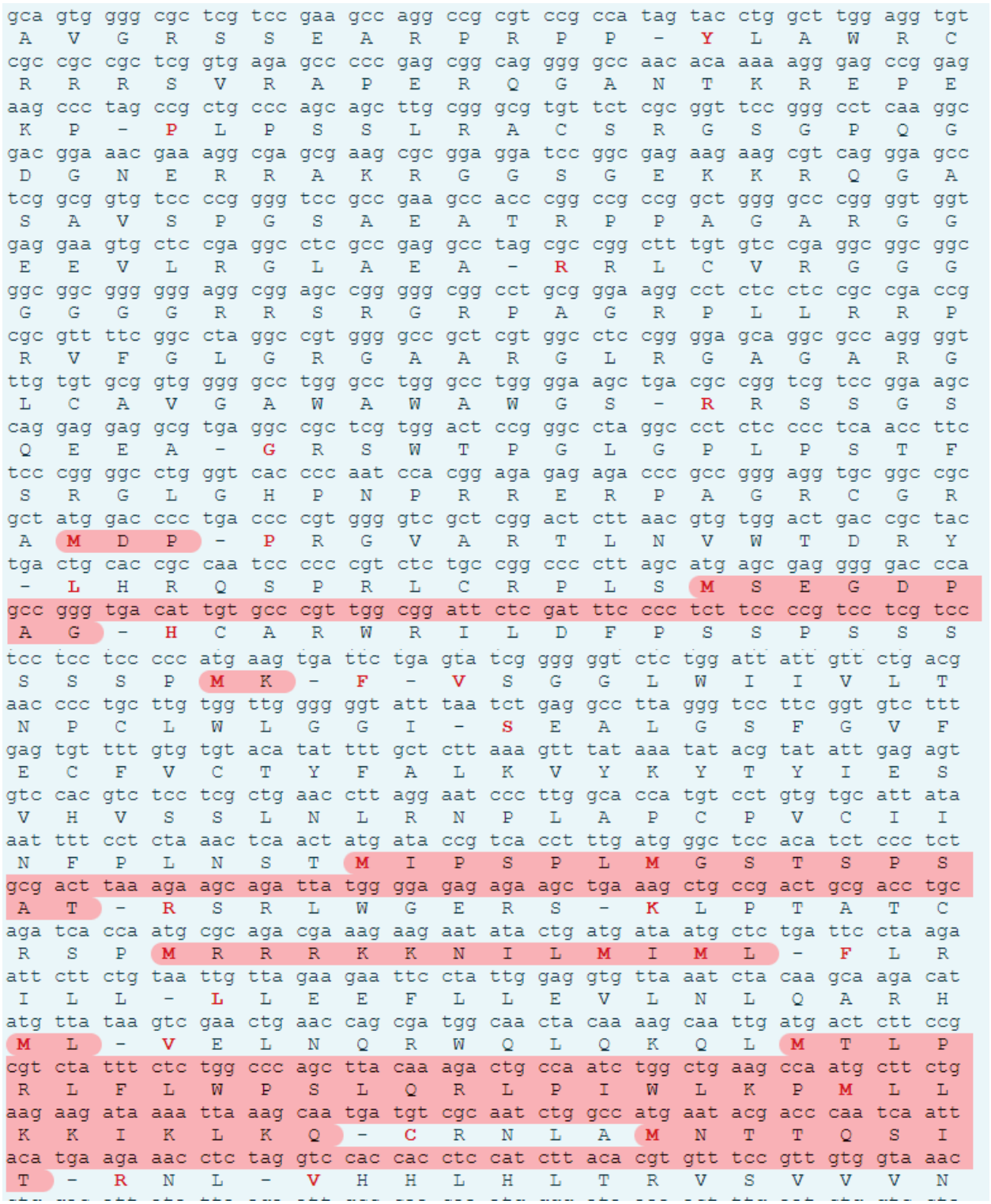
RBBP6 is a 250 kDa splicing-associated human protein initially known to bind to the retinoblastoma gene product, pRB (Sakai *et al.*, 1995). RBBP6 gene is known to possess six different domains (shown schematically in Figure 1.7) that have been characterized and linked with different types of cancer (Simons *et al*., 1997). RBBP6 has three well conserved N-terminal domains namely "Domain With No Name" (DWNN), Zinc knuckle domain and Really Interesting New Gene (RING) finger domain. It also contains three C-terminal domains namely proline-rich SR domain, Rb-binding domain and p53-binding domain (Pugh *et al*., 2006; Simons *et al*., 1997; Witte & Scott, 1997). The functions of the six identified domains are not yet fully understood however it is well understood that their functions are predominantly mediated by protein-protein interactions (Li *et al*., 2007).

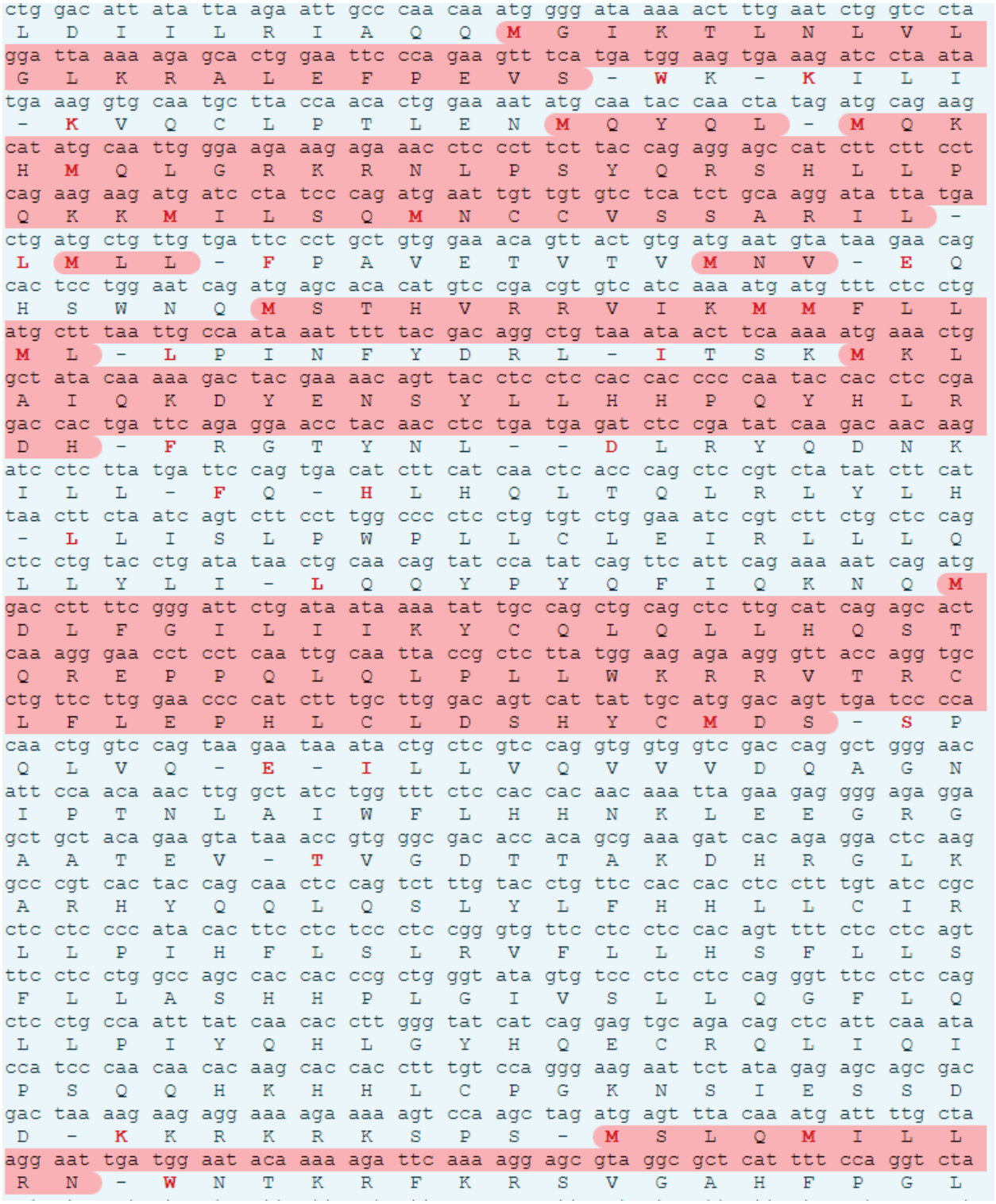
```
   1 gcagtggggc gctcgtccga agccaggccg cgtccgccat agtacctggc ttggaggtgt
  61 cgccgccgct cggtgagagc ccccgagcgg caggggggca acacaaaaag ggagccggag
 121 aagccctagc cgctgcccag cagcttgcgg gcgtgttctc gcggttccgg gcctcaaggc
 181 gacggaaacg aaaggcgagc gaagcgcgga ggatccggcg agaagaagcg tcagggagcc
 241 tcggcggtgt ccccggggtc cgccgaagcc acccggccgc cggctggggc ccggggtggt
 301 gaggaagtgc tccgaggcct cgccgaggcc tagcgccggc tttgtgtccg aggcggcggc
 361 ggcggcgggg ggaggcggag ccgggggcgg cctgcgggaa ggcctctcct ccgccgaccg
 421 cgcgttttcg gcctaggccg tggggccgct cgtggcctcc ggggagcagg cgccaggggt
 481 ttgtgtgcgg tgggggcctg ggcctgggcc tggggaagct gacgccggtc gtccggaagc
 541 caggaggagg cgtgaggccg ctcgtggact ccgggcctag gccctctccc ctcaaccttc
 601 tcccggggcc tgggtcaccc caatccacgg agagagagac ccgccgggag gtgcggccgc
 661 gctatggacc cctgaccccg tggggtcgct cggactctta acgtgtggac tgaccgctac
 721 tgactgcacc gccaatcccc ccgtctctgc cggcccctta gcatgagcga gggggaccca
 781 gccgggtgac attgtgcccg ttggcggatt ctcgatttcc cctcttcccc gtcctcgtcc
 841 tcctcctccc ccatgaagtg attctgagta tcgggggggtc tctggattat tgttctgacg
 901 aacccctgct tgtggttggg gggtatttaa tctgaggcct tagggtcctt cggtgtcttt
 961 gagtgttttg tgtgtacata ttttgctctt aaagtttata aatatacgta tattgagagt
1021 gtccacgtct cctcgctgaa ccttaggaat cccttggcac catgtcctgt gtgcattata
1081 aattttcctc taaactcaac tatgataccg tcacctttga tgggctccac atctccctct
1141 gcgacttaaa gaagcagatt atggggagag agaagctgaa agctgccgac tgcgacctgc
1201 agatcaccaa tgcgcagacg aaagaagaat atactgatga taatgctctg attcctaaga
1261 attcttctgt aattgttaga agaattccta ttggaggtgt taaatctaca agcaagacat
1321 atgttataag tcgaactgaa ccagcgatgg caactacaaa agcaattgat gactcttccg
1381 cgtctatttc tctggcccag cttacaaaga ctgccaatct ggctgaagcc aatgcttctg
1441 aagaagataa aattaaagca atgatgtcgc aatctggcca tgaatacgac ccaatcaatt
1501 acatgaagaa acctctaggt ccaccacctc catcttacac gtgtttccgt tgtggtaaac
1561 ctggacatta tattaagaat tgcccaacaa atggggataa aaactttgaa tctggtccta
1621 ggattaaaaa gagcactgga attcccagaa gtttcatgat ggaagtgaaa gatcctaata
1681 tgaaaggtgc aatgcttacc aacactggaa aatatgcaat accaactata gatgcagaag
1741 catatgcaat tgggaagaaa gagaaacctc ccttcttacc agaggagcca tcttcttcct
1801 cagaagaaga tgatcctatc ccagatgaat tgttgtgtct catctgcaag gatattatga
1861 ctgatgctgt tgtgattccc tgctgtggaa acagttactg tgatgaatgt ataagaacag
1921 cactcctgga atcagatgag cacacatgtc cgacgtgtca tcaaaatgat gtttctcctg
1981 atgctttaat tgccaataaa tttttacgac aggctgtaaa taacttcaaa aatgaaactg
2041 gctatacaaa aagactacga aaacagttac ctcctccacc acccccaata ccacctccga
2101 gaccactgat tcagaggaac ctacaacctc tgatgagatc tccgatatca agacaacaag
2161 atcctcttat gattccagtg acatcttcat caactcaccc agctccgtct atatcttcat
```

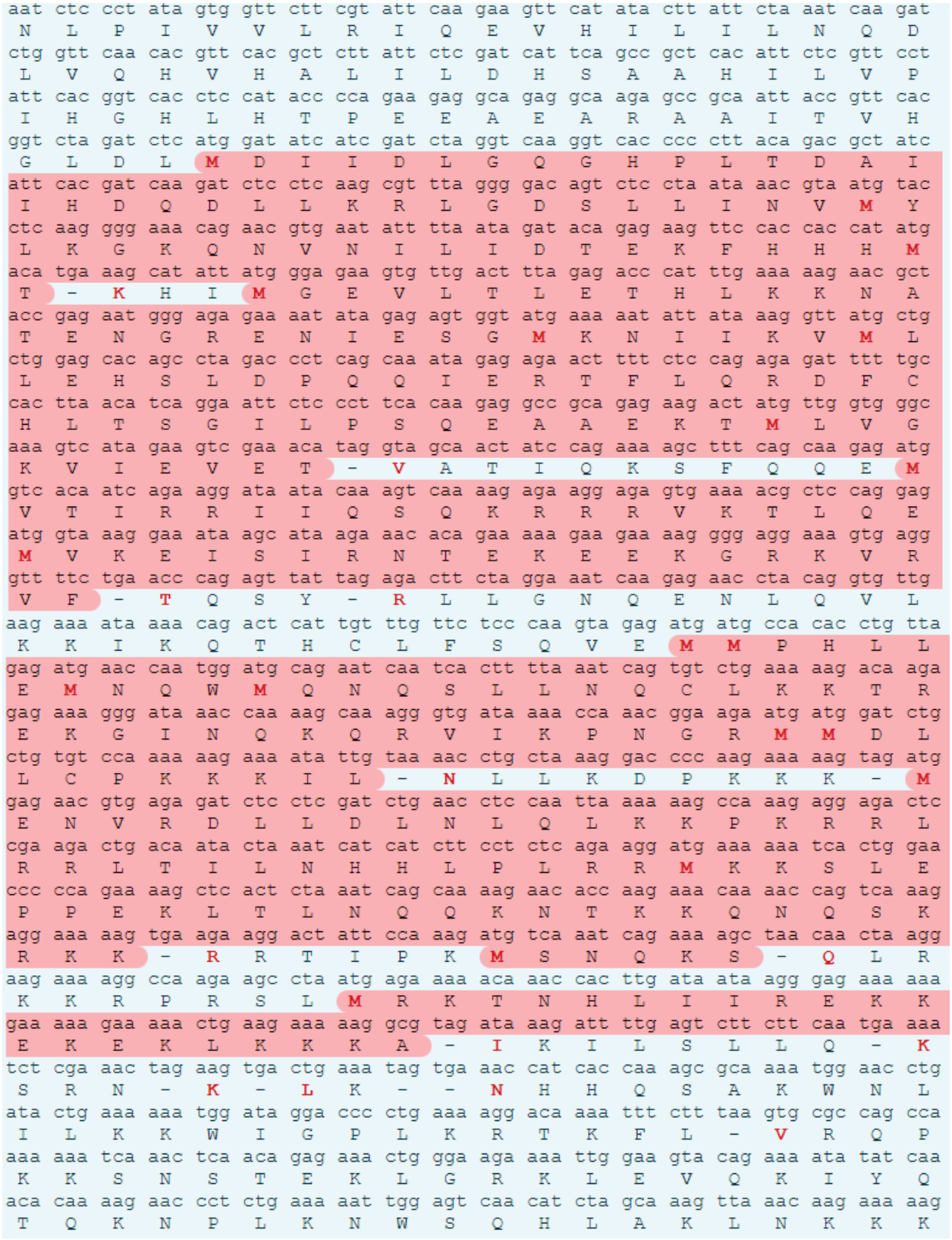
```
2221 taacttctaa tcagtcttcc ttggcccctc ctgtgtctgg aaatccgtct tctgctccag
2281 ctcctgtacc tgatataact gcaacagtat ccatatcagt tcattcagaa aaatcagatg
2341 gacctttcg ggattctgat aataaaatat tgccagctgc agctcttgca tcagagcact
2401 caaagggaac ctcctcaatt gcaattaccg ctcttatgga agagaagggt taccaggtgc
2461 ctgttcttgg aaccccatct ttgcttggac agtcattatt gcatggacag ttgatcccca
2521 caactggtcc agtaagaata aatactgctc gtccaggtgg tggtcgacca ggctgggaac
2581 attccaacaa acttggctat ctggtttctc caccacaaca aattagaaga ggggagagga
2641 gctgctacag aagtataaac cgtgggcgac accacagcga aagatcacag aggactcaag
2701 gcccgtcact accagcaact ccagtctttg tacctgttcc accacctcct ttgtatccgc
2761 ctcctcccca tacacttcct ctccctccgg gtgttcctcc tccacagttt tctcctcagt
2821 ttcctcctgg ccagccacca cccgctgggt atagtgtccc tcctccaggg tttcctccag
2881 ctcctgccaa tttatcaaca ccttgggtat catcaggagt gcagacagct cattcaaata
2941 ccatcccaac aacacaagca ccacctttgt ccagggaaga attctataga gagcagcgac
3001 gactaaaaga agaggaaaag aaaaagtcca agctagatga gtttacaaat gattttgcta
3061 aggaattgat ggaatacaaa aagattcaaa aggagcgtag gcgctcattt tccaggtcta
3121 aatctcccta tagtggttct tcgtattcaa gaagttcata tacttattct aaatcaagat
3181 ctggttcaac acgttcacgc tcttattctc gatcattcag ccgctcacat tctcgttcct
3241 attcacggtc acctccatac cccagaagag gcagaggcaa gagccgcaat taccgttcac
3301 ggtctagatc tcatggatat catcgatcta ggtcaaggtc acccccttac agacgctatc
3361 attcacgatc aagatctcct caagcgttta ggggacagtc tcctaataaa cgtaatgtac
3421 ctcaagggga aacagaacgt gaatatttta atagatacag agaagttcca ccaccatatg
3481 acatgaaagc atattatggg agaagtgttg actttagaga cccatttgaa aaagaacgct
3541 accgagaatg ggagagaaaa tatagagagt ggtatgaaaa atattataaa ggttatgctg
3601 ctggagcaca gcctagaccc tcagcaaata gagagaactt ttctccagag agatttttgc
3661 cacttaacat caggaattct cccttcacaa gaggccgcag agaagactat gttggtgggc
3721 aaagtcatag aagtcgaaac ataggtagca actatccaga aaagctttca gcaagagatg
3781 gtcacaatca gaaggataat acaaagtcaa aagagaagga gagtgaaaac gctccaggag
3841 atggtaaagg aaataagcat aagaaacaca gaaaaagaag aaaaggggag gaaagtgagg
3901 gttttctgaa cccagagtta ttagagactt ctaggaaatc aagagaacct acaggtgttg
3961 aagaaaataa aacagactca ttgtttgttc tcccaagtag agatgatgcc acacctgtta
4021 gagatgaacc aatggatgca gaatcaatca cttttaaatc agtgtctgaa aaagacaaga
4081 gagaaaggga taaaccaaaa gcaaagggtg ataaaaccaa acggaagaat gatggatctg
4141 ctgtgtccaa aaaagaaaat attgtaaaac ctgctaaagg accccaagaa aaagtagatg
4201 gagaacgtga gagatctcct cgatctgaac ctccaattaa aaaagccaaa gaggagactc
4261 cgaagactga caatactaaa tcatcatctt cctctcagaa ggatgaaaaa atcactggaa
4321 cccccagaaa agctcactct aaatcagcaa aagaacacca agaaacaaaa ccagtcaaag
4381 aggaaaaagt gaagaaggac tattccaaag atgtcaaatc agaaaagcta acaactaagg
```

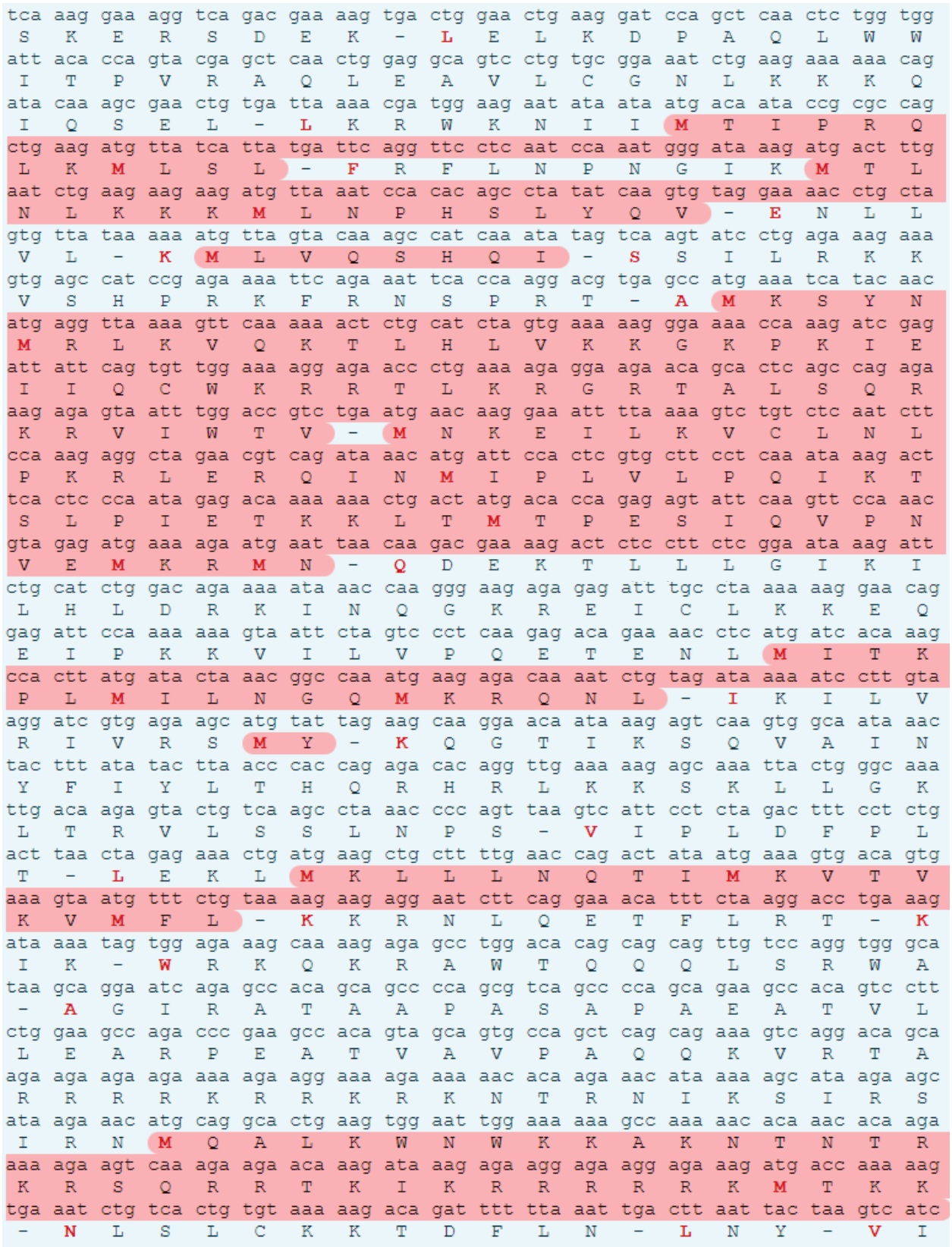
```
4441 aagaaaaggc caagaagcct aatgagaaaa acaaaccact tgataataag ggagaaaaaa
4501 gaaaaagaaa aactgaagaa aaaggcgtag ataaagattt tgagtcttct tcaatgaaaa
4561 tctcgaaact agaagtgact gaaatagtga aaccatcacc aaagcgcaaa atggaacctg
4621 atactgaaaa aatggatagg acccctgaaa aggacaaaat ttctttaagt gcgccagcca
4681 aaaaaatcaa actcaacaga gaaactggga agaaaattgg aagtacagaa aatatatcaa
4741 acacaaaaga accctctgaa aaattggagt caacatctag caaagttaaa caagaaaaag
4801 tcaaaggaaa ggtcagacga aaagtgactg gaactgaagg atccagctca actctggtgg
4861 attacaccag tacgagctca actggaggca gtcctgtgcg gaaatctgaa gaaaaaacag
4921 atacaaagcg aactgtgatt aaaacgatgg aagaatataa taatgacaat accgcgccag
4981 ctgaagatgt tatcattatg attcaggttc ctcaatccaa atgggataaa gatgactttg
5041 aatctgaaga agaagatgtt aaatccacac agcctatatc aagtgtagga aaacctgcta
5101 gtgttataaa aaatgttagt acaaagccat caaatatagt caagtatcct gagaaagaaa
5161 gtgagccatc cgagaaaatt cagaaattca ccaaggacgt gagccatgaa atcatacaac
5221 atgaggttaa aagttcaaaa aactctgcat ctagtgaaaa agggaaaacc aaagatcgag
5281 attattcagt gttggaaaag gagaaccctg aaaagaggaa gaacagcact cagccagaga
5341 aagagagtaa tttggaccgt ctgaatgaac aaggaaattt taaaagtctg tctcaatctt
5401 ccaaagaggc tagaacgtca gataaacatg attccactcg tgcttcctca aataaagact
5461 tcactcccaa tagagacaaa aaaactgact atgacaccag agagtattca agttccaaac
5521 gtagagatga aaagaatgaa ttaacaagac gaaaagactc tccttctcgg aataaagatt
5581 ctgcatctgg acagaaaaat aaaccaaggg aagagagaga tttgcctaaa aaaggaacag
5641 gagattccaa aaaaagtaat tctagtccct caagagacag aaaacctcat gatcacaaag
5701 ccacttatga tactaaacgg ccaaatgaag agacaaaatc tgtagataaa aatccttgta
5761 aggatcgtga gaagcatgta ttagaagcaa ggaacaataa agagtcaagt ggcaataaac
5821 tactttatat acttaaccca ccagagacac aggttgaaaa agagcaaatt actgggcaaa
5881 ttgacaagag tactgtcaag cctaaacccc agttaagtca ttcctctaga ctttcctctg
5941 acttaactag agaaactgat gaagctgctt ttgaaccaga ctataatgaa agtgacagtg
6001 aaagtaatgt ttctgtaaaa gaagaggaat cttcaggaaa catttctaag gacctgaaag
6061 ataaaatagt ggagaaagca aaagagagcc tggacacagc agcagttgtc caggtgggca
6121 taagcaggaa tcagagccac agcagcccca gcgtcagccc cagcagaagc cacagtcctt
6181 ctggaagcca gacccgaagc cacagtagca gtgccagctc agcagaaagt caggacagca
6241 agaagaagaa gaaaaagaag gaaaagaaaa aacacaagaa acataaaaag cataagaagc
6301 ataagaaaca tgcaggcact gaagtggaat tggaaaaaag ccaaaaacac aaacacaaga
6361 aaaagaagtc aaagaagaac aaagataaag agaaggagaa ggagaaagat gaccaaaaag
6421 tgaaatctgt cactgtgtaa aaagacagat tttttaaatt gacttaatta ctaagtcatc
6481 tgtattaaat tttgttataa tgtaaagaga ttcaagcctt gtaaataatg acatggaaga
6541 ccctgtgctg cacttaaaat attgctgctt gattatttga tttttacatc agagctttat
6601 aacacgaact tttgtacaga attgtgagtt gtgaccatgt aacatgagag gttttgctag
6661 ggcctattat ttttaaccac cattaattag ttggggtgga gtttactgta atgtgaaatt
6721 ttcacatttg aatttttaa ttgcctggca aaagctgata taagttctaa aatatcagca
6781 gaatgatttg ctgaattcat tacaaccctg ttatgtcact ttttgattac aataaaagtt
6841 ttcagtaaac ttttcaaa
```

Figure 1.7: RBBP6 DNA coding sequence (CDS)


gca gtg ggg cgc tcg tcc gaa gcc agg ccg cgt ccg cca tag tac ctg gct tgg agg tgt
A V G R S S E A R P R P P - Y L A W R C
cgc cgc cgc tcg gtg aga gcc ccc gag cgg cag ggg gcc aac aca aaa agg gag ccg gag
R R R S V R A P E R Q G A N T K R E P E
aag ccc tag ccg ctg ccc agc agc ttg cgg gcg tgt tct cgc ggt tcc ggg cct caa ggc
K P - P L P S S L R A C S R G S G P Q G
gac gga aac gaa agg cga gcg aag cgc gga gga tcc ggc gag aag aag cgt cag gga gcc
D G N E R R A K R G G S G E K K R Q G A
tcg gcg gtg tcc ccg ggg tcc gcc gaa gcc acc cgg ccg ccg gct ggg gcc cgg ggt ggt
S A V S P G S A E A T R P P A G A R G G
gag gaa gtg ctc cga ggc ctc gcc gag gcc tag cgc cgg ctt tgt gtc cga ggc ggc ggc
E E V L R G L A E A - R R L C V R G G G
ggc ggc ggg ggg agg cgg agc cgg ggg cgg cct gcg gga agg cct ctc ctc cgc cga ccg
G G G G R R S R G R P A G R P L L R R P
cgc gtt ttc ggc cta ggc cgt ggg gcc gct cgt ggc ctc cgg gga gca ggc gcc agg ggt
R V F G L G R G A A R G L R G A G A R G
ttg tgt gcg gtg ggg gcc tgg gcc tgg gcc tgg gga agc tga cgc cgg tcg tcc gga agc
L C A V G A W A W A W G S - R R S S G S
cag gag gag gcg tga ggc cgc tcg tgg act ccg ggc cta ggc cct ctc ccc tca acc ttc
Q E E A - G R S W T P G L G P L P S T F
tcc cgg ggc ctg ggt cac ccc aat cca cgg aga gag aga ccc gcc ggg agg tgc ggc cgc
S R G L G H P N P R R E R P A G R C G R
gct atg gac ccc tga ccc cgt ggg gtc gct cgg act ctt aac gtg tgg act gac cgc tac
A M D P - P R G V A R T L N V W T D R Y
tga ctg cac cgc caa tcc ccc cgt ctc tgc cgg ccc ctt agc atg agc gag ggg gac cca
- L H R Q S P R L C R P L S M S E G D P
gcc ggg tga cat tgt gcc cgt tgg cgg att ctc gat ttc ccc tct tcc ccg tcc tcg tcc
A G - H C A R W R I L D F P S S P S S S
tcc tcc tcc ccc atg aag tga ttc tga gta tcg ggg ggt ctc tgg att att gtt ctg acg
S S S P M K - F - V S G G L W I I V L T
aac ccc tgc ttg tgg ttg ggg ggt att taa tct gag gcc tta ggg tcc ttc ggt gtc ttt
N P C L W L G G I - S E A L G S F G V F
gag tgt ttt gtg tgt aca tat ttt gct ctt aaa gtt tat aaa tat acg tat att gag agt
E C F V C T Y F A L K V Y K Y T Y I E S
gtc cac gtc tcc tcg ctg aac ctt agg aat ccc ttg gca cca tgt cct gtg tgc att ata
V H V S S L N L R N P L A P C P V C I I
aat ttt cct cta aac tca act atg ata ccg tca cct ttg atg ggc tcc aca tct ccc tct
N F P L N S T M I P S P L M G S T S P S
gcg act taa aga agc aga tta tgg gga gag aga agc tga aag ctg ccg act gcg acc tgc
A T - R S R L W G E R S - K L P T A T C
aga tca cca atg cgc aga cga aag aag aat ata ctg atg ata atg ctc tga ttc cta aga
R S P M R R R K K N I L M I M L - F L R
att ctt ctg taa ttg tta gaa gaa ttc cta ttg gag gtg tta aat cta caa gca aga cat
I L L - L L E E F L L E V L N L Q A R H
atg tta taa gtc gaa ctg aac cag cga tgg caa cta caa aag caa ttg atg act ctt ccg
M L - V E L N Q R W Q L Q K Q L M T L P
cgt cta ttt ctc tgg ccc agc tta caa aga ctg cca atc tgg ctg aag cca atg ctt ctg
R L F L W P S L Q R L P I W L K P M L L
aag aag ata aaa tta aag caa tga tgt cgc aat ctg gcc atg aat acg acc caa tca att
K K I K L K Q - C R N L A M N T T Q S I
aca tga aga aac ctc tag gtc cac cac ctc cat ctt aca cgt gtt tcc gtt gtg gta aac
T - R N L - V H H L H L T R V S V V V N


ctg gac att ata tta aga att gcc caa caa atg ggg ata aaa act ttg aat ctg gtc cta
L D I I L R I A Q Q M G I K T L N L V L
gga tta aaa aga gca ctg gaa ttc cca gaa gtt tca tga tgg aag tga aag atc cta ata
G L K R A L E F P E V S - W K - K I L I
tga aag gtg caa tgc tta cca aca ctg gaa aat atg caa tac caa cta tag atg cag aag
- K V Q C L P T L E N M Q Y Q L - M Q K
cat atg caa ttg gga aga aag aga aac ctc cct tct tac cag agg agc cat ctt ctt cct
H M Q L G R K R N L P S Y Q R S H L L P
cag aag aag atg atc cta tcc cag atg aat tgt tgt gtc tca tct gca agg ata tta tga
Q K K M I L S Q M N C C V S S A R I L -
ctg atg ctg ttg tga ttc cct gct gtg gaa aca gtt act gtg atg aat gta taa gaa cag
L M L L - F P A V E T V T V M N V - E Q
cac tcc tgg aat cag atg agc aca cat gtc cga cgt gtc atc aaa atg atg ttt ctc ctg
H S W N Q M S T H V R R V I K M M F L L
atg ctt taa ttg cca ata aat ttt tac gac agg ctg taa ata act tca aaa atg aaa ctg
M L - L P I N F Y D R L - I T S K M K L
gct ata caa aaa gac tac gaa aac agt tac ctc ctc cac cac ccc caa tac cac ctc cga
A I Q K D Y E N S Y L L H H P Q Y H L R
gac cac tga ttc aga gga acc tac aac ctc tga tga gat ctc cga tat caa gac aac aag
D H - F R G T Y N L - - D L R Y Q D N K
atc ctc tta tga ttc cag tga cat ctt cat caa ctc acc cag ctc cgt cta tat ctt cat
I L L - F Q - H L H Q L T Q L R L Y L H
taa ctt cta atc agt ctt cct tgg ccc ctc ctg tgt ctg gaa atc cgt ctt ctg ctc cag
- L L I S L P W P L L C L E I R L L L Q
ctc ctg tac ctg ata taa ctg caa cag tat cca tat cag ttc att cag aaa aat cag atg
L L Y L I - L Q Q Y P Y Q F I Q K N Q M
gac ctt ttc ggg att ctg ata ata aaa tat tgc cag ctg cag ctc ttg cat cag agc act
D L F G I L I I K Y C Q L Q L L H Q S T
caa agg gaa cct cct caa ttg caa tta ccg ctc tta tgg aag aga agg gtt acc agg tgc
Q R E P P Q L Q L P L L W K R R V T R C
ctg ttc ttg gaa ccc cat ctt tgc ttg gac agt cat tat tgc atg gac agt tga tcc cca
L F L E P H L C L D S H Y C M D S - S P
caa ctg gtc cag taa gaa taa ata ctg ctc gtc cag gtg gtg gtc gac cag gct ggg aac
Q L V Q - E - I L L V Q V V V D Q A G N
att cca aca aac ttg gct atc tgg ttt ctc cac cac aac aaa tta gaa gag ggg aga gga
I P T N L A I W F L H H N K L E E G R G
gct gct aca gaa gta taa acc gtg ggc gac acc aca gcg aaa gat cac aga gga ctc aag
A A T E V - T V G D T T A K D H R G L K
gcc cgt cac tac cag caa ctc cag tct ttg tac ctg ttc cac cac ctc ctt tgt atc cgc
A R H Y Q Q L Q S L Y L F H H L L C I R
ctc ctc ccc ata cac ttc ctc tcc ctc cgg gtg ttc ctc ctc cac agt ttt ctc ctc agt
L L P I H F L S L R V F L L H S F L L S
ttc ctc ctg gcc agc cac cac ccg ctg ggt ata gtg tcc ctc ctc cag ggt ttc ctc cag
F L L A S H H P L G I V S L L Q G F L Q
ctc ctg cca att tat caa cac ctt ggg tat cat cag gag tgc aga cag ctc att caa ata
L L P I Y Q H L G Y H Q E C R Q L I Q I
cca tcc caa caa cac aag cac cac ctt tgt cca ggg aag aat tct ata gag agc agc gac
P S Q Q H K H H L C P G K N S I E S S D
gac taa aag aag agg aaa aga aaa agt cca agc tag atg agt tta caa atg att ttg cta
D - K K R K R K S P S - M S L Q M I L L
agg aat tga tgg aat aca aaa aga ttc aaa agg agc gta ggc gct cat ttt cca ggt cta
R N - W N T K R F K R S V G A H F P G L




tca aag gaa agg tca gac gaa aag tga ctg gaa ctg aag gat cca gct caa ctc tgg tgg
S K E R S D E K - L E L K D P A Q L W W
att aca cca gta cga gct caa ctg gag gca gtc ctg tgc gga aat ctg aag aaa aaa cag
I T P V R A Q L E A V L C G N L K K K Q
ata caa agc gaa ctg tga tta aaa cga tgg aag aat ata ata atg aca ata ccg cgc cag
I Q S E L - L K R W K N I I M T I P R Q
ctg aag atg tta tca tta tga ttc agg ttc ctc aat cca aat ggg ata aag atg act ttg
L K M L S L - F R F L N P N G I K M T L
aat ctg aag aag aag atg tta aat cca cac agc cta tat caa gtg tag gaa aac ctg cta
N L K K K M L N P H S L Y Q V - E N L L
gtg tta taa aaa atg tta gta caa agc cat caa ata tag tca agt atc ctg aga aag aaa
V L - K M L V Q S H Q I - S S I L R K K
gtg agc cat ccg aga aaa ttc aga aat tca cca agg acg tga gcc atg aaa tca tac aac
V S H P R K F R N S P R T - A M K S Y N
atg agg tta aaa gtt caa aaa act ctg cat cta gtg aaa aag gga aaa cca aag atc gag
M R L K V Q K T L H L V K K G K P K I E
att att cag tgt tgg aaa agg aga acc ctg aaa aga gga aga aca gca ctc agc cag aga
I I Q C W K R R T L K R G R T A L S Q R
aag aga gta att tgg acc gtc tga atg aac aag gaa att tta aaa gtc tgt ctc aat ctt
K R V I W T V - M N K E I L K V C L N L
cca aag agg cta gaa cgt cag ata aac atg att cca ctc gtg ctt cct caa ata aag act
P K R L E R Q I N M I P L V L P Q I K T
tca ctc cca ata gag aca aaa aaa ctg act atg aca cca gag agt att caa gtt cca aac
S L P I E T K K L T M T P E S I Q V P N
gta gag atg aaa aga atg aat taa caa gac gaa aag act ctc ctt ctc gga ata aag att
V E M K R M N - Q D E K T L L L G I K I
ctg cat ctg gac aga aaa ata aac caa ggg aag aga gag att tgc cta aaa aag gaa cag
L H L D R K I N Q G K R E I C L K K E Q
gag att cca aaa aaa gta att cta gtc cct caa gag aca gaa aac ctc atg atc aca aag
E I P K K V I L V P Q E T E N L M I T K
cca ctt atg ata cta aac ggc caa atg aag aga caa aat ctg tag ata aaa atc ctt gta
P L M I L N G Q M K R Q N L - I K I L V
agg atc gtg aga agc atg tat tag aag caa gga aca ata aag agt caa gtg gca ata aac
R I V R S M Y - K Q G T I K S Q V A I N
tac ttt ata tac tta acc cac cag aga cac agg ttg aaa aag agc aaa tta ctg ggc aaa
Y F I Y L T H Q R H R L K K S K L L G K
ttg aca aga gta ctg tca agc cta aac ccc agt taa gtc att cct cta gac ttt cct ctg
L T R V L S S L N P S - V I P L D F P L
act taa cta gag aaa ctg atg aag ctg ctt ttg aac cag act ata atg aaa gtg aca gtg
T - L E K L M K L L L N Q T I M K V T V
aaa gta atg ttt ctg taa aag aag agg aat ctt cag gaa aca ttt cta agg acc tga aag
K V M F L - K K R N L Q E T F L R T - K
ata aaa tag tgg aga aag caa aag aga gcc tgg aca cag cag cag ttg tcc agg tgg gca
I K - W R K Q K R A W T Q Q Q L S R W A
taa gca gga atc aga gcc aca gca gcc cca gcg tca gcc cca gca gaa gcc aca gtc ctt
- A G I R A T A A P A S A P A E A T V L
ctg gaa gcc aga ccc gaa gcc aca gta gca gtg cca gct cag cag aaa gtc agg aca gca
L E A R P E A T V A V P A Q Q K V R T A
aga aga aga aga aaa aga agg aaa aga aaa aac aca aga aac ata aaa agc ata aga agc
R R R R K R R K R K N T R N I K S I R S
ata aga aac atg cag gca ctg aag tgg aat tgg aaa aaa gcc aaa aac aca aac aca aga
I R N M Q A L K W N W K K A K N T N T R
aaa aga agt caa aga aga aca aag ata aag aga agg aga agg aga aag atg acc aaa aag
K R S Q R R T K I K R R R R R K M T K K
tga aat ctg tca ctg tgt aaa aag aca gat ttt tta aat tga ctt aat tac taa gtc atc
- N L S L C K K T D F L N - L N Y - V I

```
tgt att aaa ttt tgt tat aat gta aag aga ttc aag cct tgt aaa taa tga cat gga aga
 C   I   K   F   C   Y   N   V   K   R   F   K   P   C   K   -   -   H   G   R
ccc tgt gct gca ctt aaa ata ttg ctg ctt gat tat ttg att ttt aca tca gag ctt tat
 P   C   A   A   L   K   I   L   L   L   D   Y   L   I   F   T   S   E   L   Y
aac acg aac ttt tgt aca gaa ttg tga gtt gtg acc atg taa cat gag agg ttt tgc tag
 N   T   N   F   C   T   E   L   -   V   V   T   M   -   H   E   R   F   C   -
ggc cta tta ttt tta acc acc att aat tag ttg ggg tgg agt tta ctg taa tgt gaa att
 G   L   L   F   L   T   T   I   N   -   L   G   W   S   L   L   -   C   E   I
ttc aca ttt gaa ttt ttt aat tgc ctg gca aaa gct gat ata agt tct aaa ata tca gca
 F   T   F   E   F   F   N   C   L   A   K   A   D   I   S   S   K   I   S   A
gaa tga ttt gct gaa ttc att aca acc ctg tta tgt cac ttt ttg att aca ata aaa gtt
 E   -   F   A   E   F   I   T   T   L   L   C   H   F   L   I   T   I   K   V
ttc agt aaa ctt ttc aaa
 F   S   K   L   F   K
```

Figure 1.8: SNRPG amino acid sequence

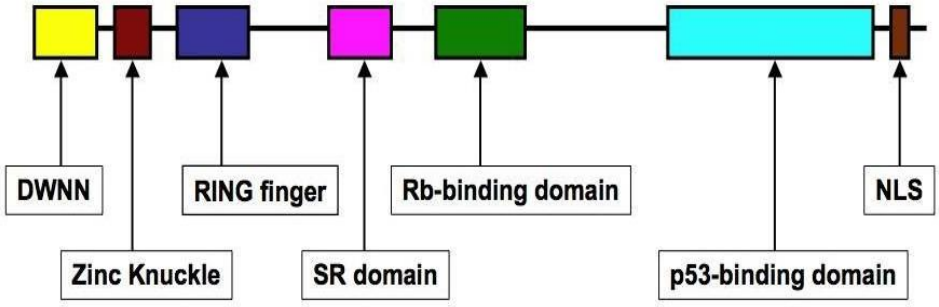


Figure 1.9: The domain organization in human RBBP6. The RBBP6 protein is a modular protein that consists of N-terminal domains (DWNN, Zinc knuckle and RING finger) and a C-terminal region (SR domain, Rb-binding domain, p53-binding domain and a nuclear localization signal) (Figure adapted from Muleya *et al*. (2010).

RBBP6 promotes cell proliferation in several cancers (Moela *et al*., 2014) and is found at varying expression levels in cancer cells (Khan *et al*., 2014). RBBP6 is highlighted significantly in more than fourteen different types of cancers that include but not limited to breast cancer, lung cancer, prostate cancer, colon cancer, pancreatic cancer, bladder cancer, cervical cancer, oesophagal cancer, liver cancer, gastric cancer, ovarian cancer and pituitary cancer (Khan *et al*., 2014). RBBP6 interacts with the two prototypical tumour suppressor proteins p53 and pRB (Khan *et al*., 2014; Li *et al*., 2007; Simons *et al*., 1997). However, the physiological relevance of the interactions and the functional basis for the association remains elusive and less clear.

RBBP6 enhances p53 ubiquitination and degradation by facilitating the interaction between p53 and its negative regulator Mdm2. RBBP6 also interferes with the binding of p53 to DNA (Li *et al.*, 2007). According to Pugh *et al*. (2006) RBBP6 directly interacts with the pro-proliferative transcription factor Y-box-binding protein-1 (YB-1). The studies have shown that over-expression of RBBP6 causes suppression of the anti-apoptotic YB-1 in a proteasome-dependent manner in cultured mammalian cells (Pugh *et al*. 2006). Considering that RBBP6 down-regulates both the pro-apoptotic p53 and the anti-apoptotic YB-1, the effect of RBBP6 in tumorigenesis remains elusive and yet to be clearly understood (Kappo *et al*., 2012).

Immunohistochemical studies by Yoshitake and coworkers (2004) reported that RBBP6 is the significantly expressed antigen in esophageal cancer. RBPP6 is strongly upregulated in esophageal cancer cells. High expression levels of RBBP6 correlate with higher rates of proliferation in cultured esophageal cancer cell and subsequent low survival rates in cancer patients (Yoshitake *et al*., 2004). Yoshitake and co-workers demonstrated that RBBP6 promotes the proliferation of oesophageal cancer cells and that antibodies against RBBP6 can be successfully used in immunotherapy against oesophageal cancer, leading to decreases in tumour size.

In another study using quantitative RT-PCR and fluorescence in situ hybridization (FISH) Motadi and coworkers (2011) found that RBBP6 has an anti-apoptotic function in lung cancer and degradative effects on the tumor suppressor proteins, p53 and pRb (Motadi *et al*., 2011). RBBP6 has also been previously identified as a strong candidate diagnostic marker for colon cancer prognosis for overall and disease-free survival in colon cancer patients (Chen *et al*., 2013). On the other hand, studies done on human MCF breast cancer cells suggest that RBBP6 promotes apoptosis (Gao *et al*., 2002a, b).

Although more commonly associated with cancer RBBP6 also plays an essential role in mRNA splicing (Chibi *et al*., 2008; Di Giammartino *et al*., 2014; Miotto *et al*., 2014). Its mice homologue PACT (p53 Associated Cellular protein-Testes) forms part of the cellular pre-mRNA splicing machinery while P2P-R (Proliferation Potential Related protein) localizes primarily to nucleoli, the site of mRNA processing where it associates with heterogeneous nuclear ribonucleoprotein particles (Witte & Scott, 1997). Mpe1, another homologue of RBBP6, is a component of the Cleavage and Polyadenylation Factor (CPF) in yeast (Shi *et al.*, 2009).

Despite RBBP6's implications in cancer and its apparent close association with mRNA splicing; very little is known about its putative interactions with mRNA splicing-associated proteins. RBBP6 plays an as-yet uncharacterized role in linking core pre-mRNA splicing proteins to various cancers. In this study we explore the structural and biophysical studies of the putative interactions between N-terminal domains of RBBP6 and a core splicing-associated protein, SNRPG.

The N terminal domains of RBBP6

The N terminus of RBBP6 is a 335-residue fragment which comprises of the DWNN, Zinc knuckle domain and RING finger domain. The three domains are well conserved and exist as single copy genes in eukaryotic genomes analyzed to date, including the single celled parasite *E. cuniculi*, in which it is very much reduced in size. In vertebrates and insects, the protein includes a long C-terminal extension containing p53 and Rb-interaction domains in human and mouse. A short form consisting of the DWNN domain and a poorly conserved C-terminal tail is also found in vertebrates (Pugh *et al*., 2006).

The "Domain With No Name" (DWNN)

Domain With No Name (DWNN) is a 13kDa ubiquitin-like domain. DWNN is found at the N-terminus of RBBP6 family of splicing-associated proteins (Altschul *et al*., 1997, Pugh *et al*., 2006). In vertebrates DWNN is expressed independently as a single domain splicing-associated protein and as the N-terminal domain of the full length RBBP6 isoforms 1 and 2 (Pugh *et al*., 2006). In addition to the 81-residues, DWNN domain contains a poorly conserved C-terminal tail. However, the domain is stable without the tail (Pugh *et al*. 2006). DWNN is alternatively expressed both as a ubiquitin-like protein (UBL) and ubiquitin domain protein (UDP) (Altschul *et al*., 1997, Pugh *et al*., 2006).

The DWNN domain contains a di-glycine peptide at position 83 and 84 (shown in Figure 1.8) similar to the conserved di-glycine motif found at position 75 and 76 in ubiquitin. The di-glycine peptide is essential for conjugation of the domain to other proteins. The presence of the di-glycine motif in the DWNN domain has led to the suggestion that the DWNN domain may be covalently linked to proteins in the same manner as ubiquitin (Pugh *et al.*, 2006). The similarity of DWNN domain to ubiquitin suggests that the domain has ubiquitin-like modifier functions. It may directly or indirectly alter protein-protein interactions and possibly plays a role in the regulation of the spliceosome (Cano *et al*., 2010). However, it remains unclear whether this involves ubiquitin-associated signaling pathways or the proteasome (Pugh *et al*., 2006).

The ubiquitin-like structure of the DWNN domain explains the reasons why RBBP6 is implicated in ubiquitin-like modification processes. And the fact that the DWNN domain is independently expressed as a single copy gene in higher vertebrates leads to the proposition that the domain independently functions as a novel ubiquitin-like modifier to other proteins (Pugh *et al.*, 2006). The three-dimensional structure of the DWNN is shown in Figure 1.8 highlighting the

superposition of the 25 lowest energy conformers (Figure 1.9A) and the cartoon representation of the overall fold and secondary structure (Figure 1.9B). As illustrated in Figure 1.9B the secondary structure of the DWNN domain consists of eight elements namely β1, β2, β3, α, β4, β5, β6 and β7. The backbone structure resembles a ubiquitin-like β-grasp fold. The central α-helix is packed against a five-stranded β-sheet consisting of β1, β2, β4, β5 and β7 strands. The N-terminal end of the DWNN central α-helix contains an additional double-stranded β-sheet comprising the β3 and β6 strands (Pugh *et al*., 2006).

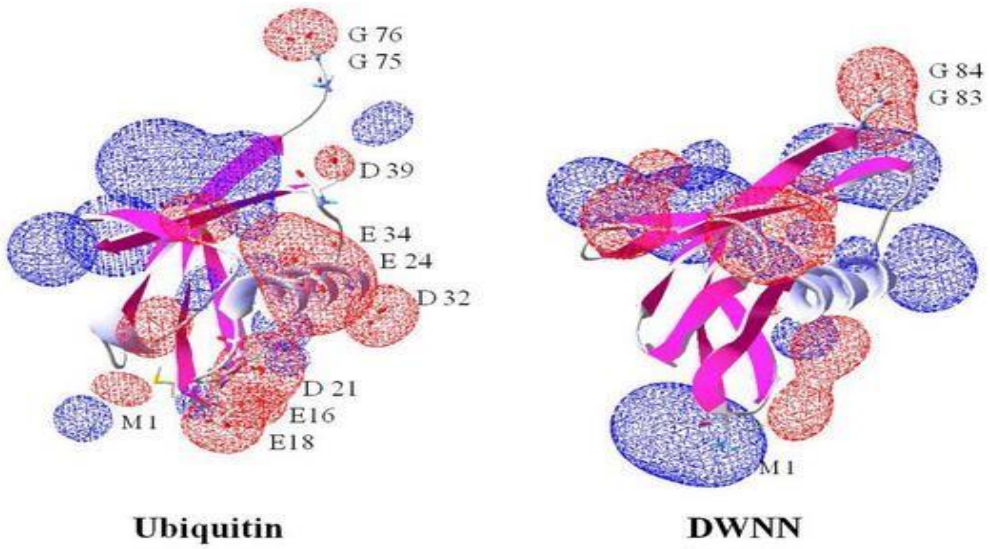


Figure 1.10: Comparison of the DWNN and Ubiquitin showing the GG-motif (Figure extracted from Pugh *et al*., 2006).

According to Li *et al*. (2007) the DWNN domain may be the medium through which RBBP6 modulates the ubiquitination of p53 by Hdm2 in an E4-like manner. E4 ligases have been described in the ubiquitination mechanisms of a few yeast and human proteins. It is thought that these conjugation factors may be required for polyubiqutination of protein substrates (Grossman *et al.*, 2003; Haenzelmann *et al.*, 2010; Pant & Lozano, 2014).

As is common in ubiquitin-like proteins, the presence of the C-terminal GG motif in human and mouse DWNN at structurally identical positions suggests that the DWNN domain may be involved in a process of conjugation called "DWNNylation". This gives an interesting hypothesis that one of the functions of RBBP6 may be to DWNNylate other proteins (Pugh *et al*., 2006). However, the mechanism remains elusive and is yet to be fully investigated.

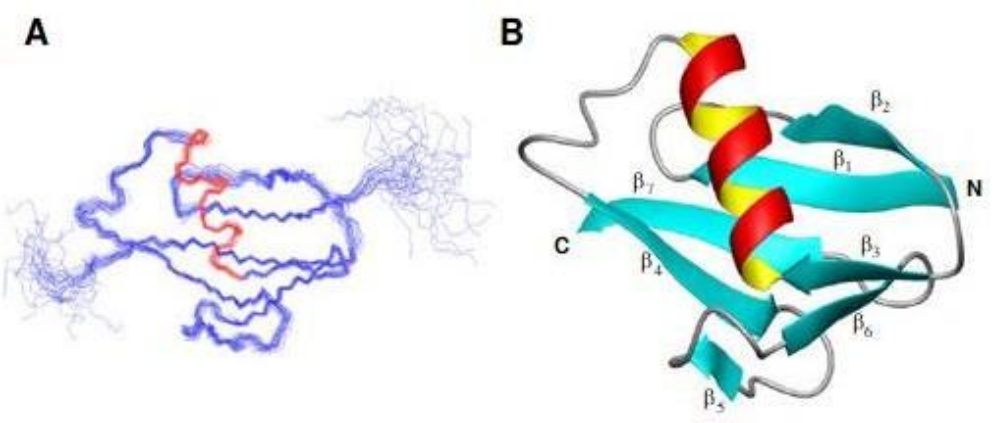


Figure 1.11: Three-dimensional structure of the DWNN domain. (**A**) Superposition of the 25 lowest energy conformers. (**B**) Cartoon representation of the overall fold and secondary structure. Figure adapted from Pugh *et al*. 2006.

DWNN mediates the interaction of the PACT (RBBP6 homologue in mice) with the tumour suppressor protein p53 (Li *et al.,* 2007). The interaction negatively regulates p53 by increasing the ubiquitination and resultant proteasomal degradation of p53 by MDM-2. DWNN promotes the ubiquitination of p53 by HDM2 in an E4-like manner. The down regulation of p53 as a result of the DWNN mediated PACT/p53 interaction has considerable implications in tumorigenesis, since p53 plays a pivotal role in driving the process of programmed cell death (Li *et al.,* 2007). Li and co-workers have established that disruption of the DWNN function in PACT leads to early embryonic death accompanied by widespread apoptosis (Li *et al.,* 2007). This suggests that DWNN is anti-apoptotic in mice and is essential for development. These findings are coherent with Mather and coworkers' suggestion that RBBP6 is essential for embryonic development

(Mather *et al.,* 2005). In another study Mbita and coworkers (2012) reported that the RBBP6 isoform 3, which comprises the DWNN only, is down regulated in human cancers while RBBP6 isoforms 1 and 2 tend to be up regulated (Mbita *et al*., 2012). According to Di Giammartino *et al*. (2014) the possible explanation for this is that DWNN as an independent module that antagonizes the larger isoforms by competition. Mbita and coworkers (2012) further expounded that DWNN controls the cell cycle at G2/M transition since the reduced expression of the DWNN leads to fewer cells at G2/M. Previous studies have also shown that there are connections between the G2/M transition and pre-mRNA splicing (Burns & Gould, 1999; Burns *et al*., 2002).

DWNN is involved in mRNA 3′-end processing and splicing (Pugh *et al*., 2006; Shi *et al.*, 2009; Vo *et al.,* 2001) and is a component of the polyadenylation machinery (Di Giammartino *et al*., 2014; Shi *et al.*, 2009). DWNN regulates the 3'-end processing of RNAs with AU-rich 39 un-translated regions (UTRs) such as c-Fos and c-Jun (Di Giammartino *et al*., 2014). Di Giammartino and coworkers showed that a derivative of RBBP6 lacking the DWNN domain was unable to reconstitute cleavage activity, indicating that the DWNN domain is essential in 3'-end processing. According to Giammartino *et al*. (2014) DWNN inhibits the binding to RNA, because a derivative of RBBP6 lacking the DWNN domain has enhanced RNA-binding affinity.

The overexpression of DWNN results in the inhibition of the 3' end pre-mRNA cleavage (Di Giammartino *et al*., 2014). DWNN out-competes RBBP6 for binding to CstF, thus inhibiting cleavage of SVL pre-mRNA (Di Giammartino *et al*. 2014) indicating that the DWNN domain is essential in 3'-end processing. However, the co-immunoprecipitation of RBBP6 with 3′ processing factors and the strong binding to the cleavage stimulatory factor-64 (CstF64) is mediated by the DWNN (Di Giammartino *et al*., 2014).

Using the yeast 2-hybrid technique, it has been postulated that RBBP6 may perhaps interact with SNRPG through its DWNN domain (Chibi *et al.,* 2008) which is a crucial component of the RNA processing machinery in the cell. These suggestions may substantiate the involvement of RBBP6 through its DWNN domain in pathways linked to the pre-mRNA splicing machinery. However, the precise mechanisms involved remain to be discovered.

The DWNN domain of RBBP6 has been shown to interact with two of the seven core splicing proteins, SmB (Simons *et al.*, 1997) and SNRPG (Chibi *et al*., 2008; Kappo *et al*., 2012). However, the mechanisms and functional basis are yet to be fully understood. Considering that DWNN mediates the ubiquitination of p53 and pRb; interacts with SNRPB and SNRPG; and localises in the nucleolus and nuclear speckles like SNRPG, it has been speculated that DWNN may link mRNA 3′-end processing to pRb/p53 pathways and tumorigenesis (Shi *et al.*, 2009). Perhaps understanding the binding events in the putative interactions between DWNN and SNRPG may lead to new knowledge on how the two proteins relate in regulating splicing, tumorigenesis and tumor development. Thus, new avenues to design and develop new therapeutic drugs may be established.

Zinc finger domain

The Zinc finger (Znf) domain is a relatively small protein motif which contains multiple finger-like protrusions that make tandem contacts with their target molecule. The binding properties of the Znf domain depend on the amino acid sequence, the finger-linker, the higher-order structures and the number of fingers (Krishna *et al*., 2003; Lopato *et al*., 1999). Generally, Znf domains are often clustered with the fingers having different binding specificities (Lopato *et al*., 1999). There are quite a number of superfamilies of Znf motifs and many of them vary in both their sequence

and structure. The RBBP6 Znf domain is also known as a "zinc knuckle" and is of the CysCysHisCys (CCHC) type comprising of 18 residues (Krishna *et al*., 2003).

The CCHC type Znf domains have the sequence: C-X2-C-X4-H-X4-C where X can be any amino acid and the number indicates the number of residues. One example of Znf domain structure is shown in Figure 1.10 below. The CCHC type Znf domains occur in a number of proteins associated with the mRNA including splicing factors (Krishna *et al*., 2003; Lopato *et al*., 1999). Often times the appearance of Znf domains and splicing factors occurs concurrently with the development of pre-mRNA splicing (Anantharaman *et al*., 2002). Znf domains are stable structures. They rarely undergo conformational changes upon target binding. The Zinc component is thought to play a major structural role in contributing to the stability of the domain (Krishna *et al*., 2003).

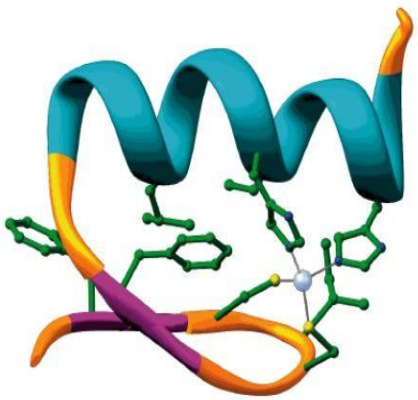


Figure 1.12: Schematic secondary structure of the Zinc finger domain. The Zinc finger motif consist of an α helix and an antiparallel β sheet. The zinc ion (white) is coordinated by cysteine and histidine residues. Adapted from Brown, 2005.

Zinc Finger Proteins in Cancer Progression

Over the years increasing evidence on the potential roles of Znf domains in cancer has been reported. However, the underlying mechanisms of Znf domains in cancer vary n different cancer types under different types of stress (Jen & Wang, 2016). Studies have revealed that aberrant expression of Znf proteins contributes to tumorigenesis and tumour development in different ways. Yang and co-workers (2008a, b) reported a Znf protein, ZKSCAN3, as a novel "driver" in colorectal cancer progression. It has been reported that ZKSCAN3 amplified and over-expressed in prostate cancer and multiple myeloma enhancing cell proliferation (Yang *et al*., 2008a, b; Yang *et al*., 2011; Zhang *et al*., 2012).

In lung cancer another Znf protein, ZNF322A, promotes cell proliferation, metastasis and invasion through transcriptionally activating cyclin D1 and alpha-adducin, and suppressing p53 (Jen *et al*., 2015). In another study ZNF304 has been shown to play a critical role in silencing tumour suppressors (Serra *et al*., 2014). Experimental validations have revealed the involvement of ZNF304 in ovarian cancer metastasis (Aslan *et al*., 2015). ZNF139 (another Znf protein) is significantly over-expressed in gastric cancer patients and is thought to promote cancer migration and invasion (Li *et al*., 2014a, b). Fan and co-workers (2015) also added that ZNF139 is implicated in cell proliferation and the inhibition of apoptosis.

Over-expression of Znf protein, X-linked (ZFX), has been shown to promote cell growth and metastasis in breast cancer, gallbladder cancer, laryngeal squamous cell carcinoma, glioma, non-small cell lung cancer, gastric cancer and oral squamous cell carcinoma (Jiang *et al*., 2012; Lai *et al*., 2014; Li *et al*., 2015; Ma *et al*., 2015; Weng *et al*., 2015; Wu *et al*., 2013; Yang *et al*., 2014; Zhu *et al*., 2013).

Zinc Finger Proteins as Tumour Suppressors

In addition to cancer promotion a number of Znf proteins have been found to function as tumor suppressors. In nasopharyngeal, esophageal, lung, gastric, colon and breast cancer ZNF545 acts as a tumor suppressor by inducing cell apoptosis (Cheng *et al*., 2012). ZNF331 inhibits cell growth and cell migration/invasion by down-regulating genes (Yu *et al*., 2013). ZNF24 negatively regulates tumor growth by inhibiting angiogenesis in breast cancer tumorigenesis (Harper *et al*., 2007; Jia *et al*., 2013). In breast cancer ZNF668 facilitates p53 stabilization and activity. It disrupts MDM2-mediated ubiquitination and degradation (Hu *et al*., 2011). In hepatocellular carcinoma and gastric cancer Zinc-fingers and homeoboxes-1 (ZHX1) are down-regulated and are known to enhance apoptosis (Ma *et al*., 2016; Wang *et al*., 2013; Wang *et al*., 2014).

Taken together these findings indicate the potential of Znf proteins as therapeutic targets. Considering that Znf proteins vary in their roles in different types of cancer drugs targeting specific Znf protein-protein interactions can be targeted for therapeutic strategy against tumors in a specific stage of cancer progression. Hence in this study we aim to investigate the putative interactions between the RBBP6 Zinc finger domain and SNRPG.

Really Interesting New Gene (RING) finger domain

Another functionally relevant domain in human RBBP6 is an N-terminal cysteine-rich Really Interesting New Gene (RING) finger domain. The RING finger domain promotes the ubiquitination of p53 by Hdm2 (Li *et al*., 2007). The cysteine-rich domain of RBBP6 has been classified both as a RING finger and U-box due to the presence of eight conserved cysteine residues and the conserved pattern of hydrophobic residues (Aravind & Koonin, 2000).

RING fingers are small domains approximately 70 residues in length (Kappo *et al*., 2012). The structure of the RING finger like-domain adopts a homodimeric configuration very similar to that present in U-boxes (shown in Figure 1.11). The interface of the structure consists of three strands of β-sheets (green) and the C-terminal helices (pink). Zinc ions are shown as yellow spheres. The interface contains a significant number of hydrophobic residues (brown) which overlaps inside the opposing monomer. At the centre of the interface is a pair of reciprocal hydrogen bonds between Asn312 and its counterpart. All RING fingers have two large loops which fold separately with the help of two zinc (Zn^{2+}) ions. The zinc ions are stabilised in a "cross-braced" fashion by four pairs of cysteine or histidine residues (Kappo *et al*., 2012).

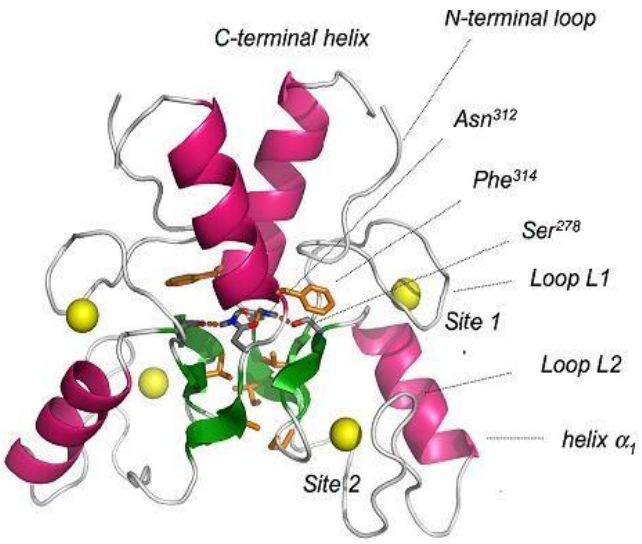


Figure 1.13: Schematic secondary structure of the RING Finger homodimer in solution. (Figure adapted from Kappo *et al*. 2012).

One ion is coordinated by the first and third Cys/His pairs while the other ion is coordinated by the second and fourth Cys/His pairs (Kappo *et al*., 2012). RING fingers are categorized

according to the pattern of zinc coordinating residues (Kostic *et al*., 2006). The C3HC4 is the most common however C4C4, C2H2C4 and C3HHC3 RING fingers do also exist. RBBP6 consists of eight conserved cysteine residues suggesting that it contains a C4C4 RING finger domain binding two zinc ions (Kappo *et al*., 2012; Kostic *et al*., 2006).

The two large loops lie parallel to an α-helix packing against a tripple-stranded β-sheet identical to a U-box fold. However, the link of zinc ions is substituted by salt bridges and hydrogen bonds (Ohi *et al*., 2003). RING fingers have been shown to form homodimers in solution (Kostic *et al*., 2006) and in some cases dimerization is essential for ubiquitination activity (Hashizume *et al*., 2001; Nikolay *et al*., 2004). The dimerization is stabilized by interactions between the monomers at the N and C termini and it occurs along the β-sheet (Nikolay *et al*., 2004).

Ubiquitination is a concerted action of three different enzymes: ubiquitin-activating enzyme (E1), ubiquitin-conjugating enzyme (E2), and substrate-specific ubiquitin-protein ligase (E3) (Kerscher *et al.*, 2006). Ubiquitination is a highly specific, energy-dependant degradation process that is achieved through the attachment of one or more ubiquitin moieties to a target protein (Figure 1.12). Once tagged with ubiquitin, the target molecules are then degraded in the proteasome, a barrel-shaped protein complex, where proteins are disassembled by proteases (Kerscher *et al.*, 2006).

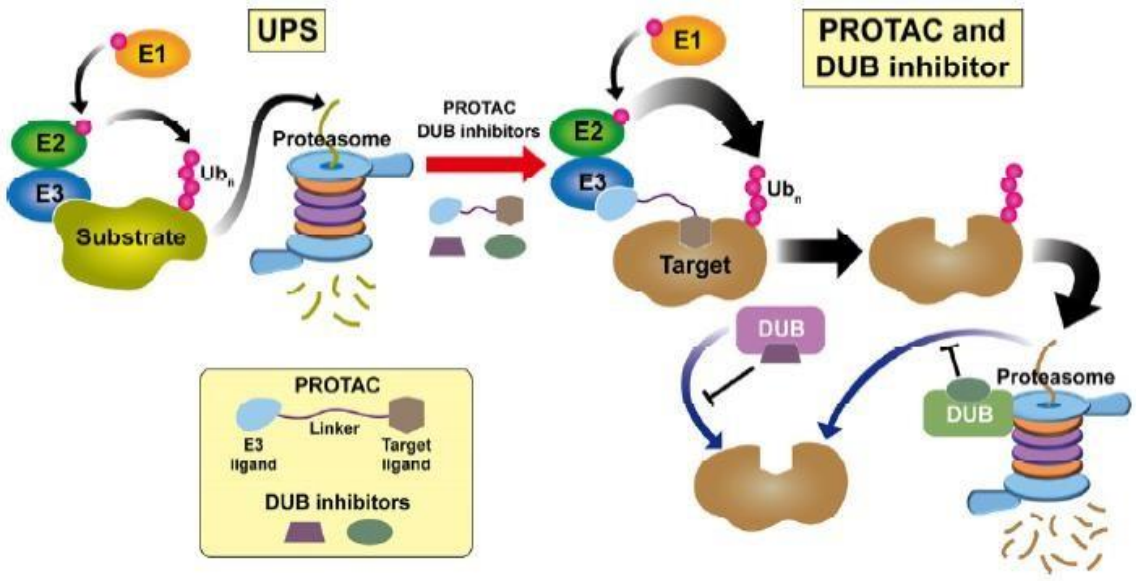


Figure 1.14: General scheme of ubiquitin-proteasome system. An enzymatic cascade of E1-E2-E3 transfers ubiquitin to Lys residues on the substrate. Therefore, the polyubiquitinated substrate is recognized by the 26S proteasome and undergoes degradation. (Right) Induced proteolysis by PROTAC and/or DUB inhibitor. PROTAC links E3 and the target protein, enhances E3-mediated ubiquitination, and promotes degradation of the target molecule. By contrast, DUBs acting upstream of the proteasome (violet) or on the proteasome (olive drab) can inhibit degradation through substrate deubiquitination. Therefore, DUB inhibitors may facilitate the proteasomal degradation pathway by antagonizing deubiquitination. Adapted from Moon and Lee, 2018.

The primary function of RING fingers is to recruit the ubiquitin conjugated E2. This allows ubiquitin to be transferred from the E2 to the substrate and is further recruited by the substrate-binding domain of the E3 (Cyr *et al.*, 2002). The presence of the RING finger domain confers E3-ubiqutin ligase activity to RBBP6 thereby implicating the nuclear protein in the ubiquitination and subsequent degradation of substrate proteins in a proteasome-dependent manner (Chibi *et al.*, 2008).

Ubiquitin-protein ligases (E3) are classified into four types namely the RING (really interesting new gene), U-box (a modified RING motif without the full complement of Zn^{2+}-binding ligands), HECT (homologous to E6- associated protein C-terminus) and N-end rule E3 ligases. The RING finger motif-containing E3s harbors a RING whose function is to bind the ubiquitin-

bound E2, bringing it into contact with the substrate which is bound by a separate substrate binding domain of the E3 (Deshaies & Joazeiro, 2009).

RING E3s interact simultaneously with E2 and with substrate suggesting that they facilitate the recruitment of E2s to the vicinity of proteins to be ubiquitinated (Joazeiro *et al.*, 1999). Human RBBP6 has intrinsic E3 ubiquitin ligase activity due to the presence of the RING finger domain in its structure (Chibi *et al.*, 2008). RBBP6 has been shown to ubiquitinate YB-1 leading to the suppression of YB-1 levels in a proteasome-dependant manner (Chibi *et al.*, 2008). RBBP6 has also been shown to ubiquitinate the methyl-binding transcriptional repressor Zinc finger and BTB domain containing 38 (zBTB38) (Miotto *et al.*, 2014). Taking all this into account it is possible that other substrates ubiquitinated by RBBP6 are still yet to be identified.

A report from Li and co-workers suggests that RBBP6 also has E4 ligase activity (Li *et al.*, 2007). E4 enzymes help to elongate short ubiquitin chains (Hoppe 2005). According to Li *et al.* (2007) RBBP6 is not in itself responsible for ubiquitination of p53. Research has shown that RBBP6 enhances the polyubiquitination of p53 by the Mdm2 E3 ligase. How? It acts directly as a ubiquitin ligase on the oncogenic transcription factor YB-1 (Chibi *et al.*, 2008; Li *et al.*, 2007). RBBP6 may play the role of a molecular scaffold, recruiting p53 and MDM-2 and thereby facilitating their interaction (Li *et al.*, 2007). Li and co-workers also demonstrated that RBBP6 interacts directly with HDM2 and promotes the ubiquitination and degradation of p53. While over-expression of RBBP6 in HEK293 cells had no effect on p53 ubiquitination, co-transfection of RBBP6 and HDM2 led to enhanced ubiquitination and degradation of p53. Replacement of RBBP6 by a mutant lacking the RING finger abolished the enhancement of p53 degradation (Li *et al.*, 2007).

A study conducted by co-transfection of a p53-responsive reporter plasmid and increasing amounts of RBBP6 showed that RBBP6 inhibited the transcriptional activity of p53 in a dose-dependent manner while inactivation of RBBP6 stabilized p53 and induced p53-dependent apoptosis (Li *et al.*, 2007). The results from this study establish RBBP6 as a negative regulator of p53 and are coherent with other studies that RBBP6 is essential for early embryonic development. RBBP6 may also function as a scaffold for the assembly of the p53-HMD2 complex (Li *et al.*, 2007). The fact that RBBP6 down-regulates both the pro-apoptotic p53 and the anti-apoptotic YB-1 its roles on tumorigenesis are tremendously complex and yet to be fully understood.

The reversible addition of ubiquitin has been shown to alter protein-protein interactions in processes of mRNA biology such as transcription, splicing and nuclear export (Cano *et al.*, 2010). However, there is no direct role for ubiquitinin in mRNA 3'-end processing that has been demonstrated as yet. According to Cano *et al.* (2012) RING fingers are essential for efficient interaction with RNA. Several RING-type E3 ubiquitin ligases possess canonical RNA-binding domains (RBDs) and are known to bind RNA and regulate the translation or stability of mRNAs. However, the role of ubiquitination in this process remains elusive and yet to be elucidated (Cano *et al.*, 2012).

PROBLEM STATEMENT

The increasing cancer morbidity and mortality rates, the lack of specificity of most cancer drugs and side effects of cancer therapy drugs remain a global public health concern. Novel techniques to design and develop cancer therapy drugs and treatment methods are of great interest. Protein-protein interaction studies have become promising targets for therapeutic discovery. Recent

studies have provided encouraging proof of selective and efficient inhibition of protein-protein interactions to open new avenues in the development of new cancer therapy drugs (Araujo *et al*., 2007; Fujii *et al*., 2007; Harris, 2006).

Previous studies proposed putative interactions between the two cancer implicated proteins, SNRPG and the N terminal domains of RBBP6. The N terminal domain of RBBP6 has been shown to interact with two of the seven core splicing proteins SNRPB (Simons *et al.*, 1997) and SNRPG (Chibi *et al*., 2008). In another study, Kappo and co-workers (2012) also identified two copies of SNRPG (conformational isomers of the same protein) as part of the five substrates that bind to the N-terminal domain of RBBP6 (Kappo *et al*., 2012). The findings substantiate the involvement of SNRPG with the N terminal domains of RBBP6 in tumorigenesis and tumour development. However, the precise mechanisms involved remain elusive and the mechanisms behind the binding events are yet to be fully investigated. The purpose of the study is to investigate the putative interactions between SNRPG and RBBP6 RING Finger domain. The study is essential to facilitate the development of new cancer drug designs that aim at breaking and or inhibiting protein-protein interactions between SNRPG and RBBP6 RING Finger domain.

AIM AND OBJECTIVES

The aim of the study is to conduct structural and biophysical studies of the putative interactions between SNRPG and the RING finger domain of RBBP6. The aim will be accomplished through the following objectives:

1) Examine *in silico* the domain structures and interactions between human SNRPG and the human RING finger domain of RBBP6

- Homology Modeling of SNRPG and the RING finger domain of RBBP6

- Molecular Docking using MGL tools and AutoDock Vina and ligand prediction using I-TASSER
- Molecular Dynamics (MD) Simulation and Post-Dynamic Analyses using AMBER18 software suite
- Determination of binding free energies by the Molecular Mechanics/Generalized Born Surface Area (MM/GBSA) method

2) Determine *in vitro* the binding affinity and strength of the determined interactions between human SNRPG and the human RING finger domain of RBBP6

- Recombinant expression of sufficient amounts of SNRPG and RING finger domain of RBBP6
- Purification of the SNRPG protein to homogeneity using Ni-NTA affinity chromatography and RING finger domain of RBBP6 using glutathione-agarose (GST) affinity chromatography
- Determination of binding affinity and strength between RING finger domain of RBBP6 and SNRPG using MicroScale Themophoresis (MST) assay

REFERENCES

Alstyne MV, Simon CM, Sardi SP, Shihabuddin LS, Mentis GZ and Pellizzoni L (2018). Dysregulation of Mdm2 and Mdm4 alternative splicing underlies motor neuron death in spinal muscular atrophy. *Genes and Development* **32**(15-16): 1045–1059.

Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W and Lipman DJ (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research* **25**(17): 3389-3402.

Anantharaman V, Koonin EV and Aravind L (2002). Comparative genomics and evolution of proteins involved in RNA metabolism. *Nucleic Acids Research* **30**(7): 1427-1464.

Anchi T, Tamura K,Furihata M, Satake H, Sakoda H, Kawada C, Kamei M, Shimamoto T, Fukuhara H, Fukata S, Ashida S, Karashima T, Yamasaki I, Yasuda M, Kamada M, Inoue K and Shuin T (2012). SNRPE is involved in cell proliferation and progression of high-grade prostate cancer through the regulation of androgen receptor expression. *Oncology Letters* **3**(2): 264-268.

Araujo RP, Liotta LA and Petricoin EF (2007). Proteins, drug targets and the mechanisms they control: the simple truth about complex networks. *Nature Reviews Drug Discovery* **6**(11): 871-880.

Aravind L and Koonin EV (2000). The Ubox is a modified RING finger. A common domain in ubiquitination. *Current Biology* **10**(4): R132–R134.

Aslan B, Monroig P, Hsu MC, Pena GA, Rodriguez-Aguayo C, Gonzalez-Villasana V, Rupaimoole R, Nagaraja AS, Mangala S, Han HD and Yuca E(2015). The ZNF304-integrin axis protects against anoikis in cancer. *Nature Communications* **6**(1): 1-12.

Battle DJ, Kasim M, Yong J, Lotti F, Lau CK, Mouaikel J, Zhang Z, Han K, Wan L and Dreyfuss G (2006). The SMN complex: An assembly machine for RNPs. Cold Spring Harbor

Laboratory Press **71**: 313-320.

Bhandari GP, Angdembe MR, Dhimal M, Neupane S and Bhusal C (2014). State of non-communicable diseases in Nepal. *BMC Public Health* **14**(1): 1-9.

Blijlevens M, van der Meulen-Muileman IH, de Menezes RX, Smit EF and van Beusechem VW (2019). High-throughput RNAi screening reveals cancer-selective lethal targets in the RNA spliceosome. *Oncogene* **38**(21): 4142-4153.

Bothwell J and Griffin J (2010). An introduction to biological nuclear magnetic resonance spectroscopy. *Biological Reviews* **86**(2): 493-510.

Boxer DH, Zhang H, Gourley DG, Hunter WN, Kelly SM and Price NC (2004). Sensing of remote oxyanion binding at the DNA binding domain of the molybdate-dependent transcriptional regulator, ModE. *Organic and Biomolecular Chemistry* **2**(19): 2829-2837.

Bronowska AK (2011). Thermodynamics of ligand-protein interactions: Implications for molecular design. *In Thermodynamics—Interaction Studies—Solids, Liquids and Gases*. IntechOpen.

Brown RS (2005). Zinc finger proteins: getting a grip on RNA. *Current Opinion in Structural Biology* **15**(1): 94-98.

Burns CG and Gould KL (1999). Connections between pre-mRNA processing and regulation of the eukaryotic cell cycle. *Frontiers of hormone research* **25**: 59–82.

Burns CG, Ohi R, Mehta S, O'Toole ET, Winey M, Clark TA, Sugnet CW, Ares Jr M and Gould KL (2002). Removal of a single alpha-tubulin gene intron suppresses cell cycle arrest phenotypes of splicing factor mutations in Saccharomyces cerevisiae. *Molecular and Cellular Biology* **22**(3): 801–815.

Cano F, Miranda-Saavedra D and Lehner PJ (2010). RNA-binding E3 ubiquitin ligases: novel players in nucleic acid regulation. *Biochemical Society Transactions* **38**(6): 1621–1626.

Chen J, Tang H, Wu Z, Zhou C, Jiang T, Xue Y, Huang G, Yan D and Peng Z (2013). Overexpression of RBBP6, Alone or Combined with Mutant TP53, Is Predictive of Poor Prognosis in Colon Cancer. *PLoS One* **8**(6): e66524.

Cheng Y, Liang P, Geng H, Wang Z, Li L, Cheng SH, Ying J, Su X, Ng KM, Ng MH and Mok TS (2012). A novel 19q13 nucleolar zinc finger protein suppresses tumor cell growth through inhibiting ribosome biogenesis and inducing apoptosis but is frequently silenced in multiple carcinomas. *Molecular Cancer Research* **10**(7): 925- 36.

Chibi M, Meyer M, Skepu A, Rees DJ, Moolman-Smook JC and Pugh DJ (2008). RBBP6 interacts with multifunctional protein YB-1 through its RING finger domain, leading to ubiquitination and proteosomal degradation of YB-1. *Journal of Molecular Biology* **384**(4): 908–916.

Conte N, Charafe-Jauffret E, Delaval B, Adelaide J, Ginestier C, Geneix J, Isnardon D, Jacquemier J and Birnbaum D (2002). Carcinogenesis and translational controls: TACC1 is down-regulated in human cancers and associates with mRNA regulators. *Oncogene* **21**(36): 5619- 5630.

Cyr DM, Hohfeld J and Patterson C (2002). Protein quality control: U-box-containing E3 ubiquitin ligases join the fold. *Trends in Biochemical Sciences* **27**(7): 368-375.

Dardevet L, Rani D, Aziz TAE, Bazin I, Sabatier JM, Fadl M and De Waard M (2015). Chlorotoxin: A Helpful Natural Scorpion Peptide to Diagnose Glioma and Fight Tumor Invasion. *Toxins* **7**(4): 1079-1101.

Deshaies RJ and Joazeiro CAP (2009). RING Domain E3 Ubiquitin Ligases. *Annual Review of*

Biochemistry 78: 399-434.

Dessau M, Chamovitz DA and Hirsch JA (2006). Expression, purification and crystallization of a PCI domain from the COP9 signalosome subunit 7 (CSN7). *Acta Crystallographica Section F: Structural Biology and Crystallization Communications* **62**(11):1138-1140.

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.

Dipankar C and Murugan S (2012). The green synthesis, characterization and evaluation of the biological activities of silver nanoparticles synthesized from Iresineherbstii leaf aqueous extracts. *Colloids and Surfaces. B, Biointerfaces* **98**: 112-119.

Domling A (2008). Small molecular weight protein-protein interaction antagonists: An insurmountable challenge? *Current Opinion in Chemical Biology* **12**(3): 281-291.

Du X, Li Y, Xia Y, Ai S, Liang J, Sang P, Ji X Liu S (2016). Insights into Protein– Ligand Interactions: Mechanisms, Models and Methods. *International journal of molecular sciences* **17**(2): 144.

Ezkurdia I, Juan D and Rodriguez JM (2014). Multiple evidence strands suggest that there may be as few as 19 000 human protein-coding genes. *Human molecular genetics* 23(22): 5866-5878.

Fearon ER and Vogelstein B (1999). A genetic model for colorectal tumorigenesis. *Cell* **61**(5): 759-767.

Ferlay J, Soerjomataram I, Dikshit R, Eser S, Mathers C, Rebelo M, Parkin DM, Forman D and Bray F (2015). Cancer incidence and mortality worldwide: Sources, methods and major patterns in GLOBOCAN 2012. *International Journal of Cancer* **136**(5): E359-E386.

Fischer U, Englbrecht C and Chari A (2011). Biogenesis of spliceosomal small nuclear ribonucleoproteins. *Wiley Interdisciplinary Reviews: RNA* **2**(5): 718–731.

Fujii N, You L, Xu Z, Uematsu K, Shan J, He B, Mikami I, Edmondson LR, Neale G, Zheng J, Guy RK and Jablons DM (2007) An antagonist of dishevelled protein-protein interaction suppresses beta-catenin-dependent tumor cell growth. *Cancer Research* **67**(2): 573-579.

Fuller JC, Burgoyne NJ and Jackson RM (2009). Predicting druggable binding sites at the protein-protein interface. *Drug Discovery Today* **14**(3-4):155-161.

Gao S and Scott RE (2002a). P2P-R protein overexpression restricts mitotic progression at prometaphase and promotes mitotic apoptosis. *Journal of Cellular Physiology* **193**(2): 199-207.

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

Gasteiger E, Hoogland C, Gattiker A, Duvaud S, Wilkins MR, Appel RD, Bairoch A (2005). Protein identification and Analysis Tools on the ExPasy Server; (In) John M. Walker (ed): The Proteomics Protocols Handbook, Human Press, pp. 571-607.

Green MR (1986). Pre-mRNA splicing. *Annual Review of Genetics* **20**(1): 671-708.

Grimm C, Chari A, Pelz JP, Kuper J, Kisker C, Diederichs K, Stark H, Schindelin H and Fischer U. (2013). Structural basis of assembly chaperone- mediated snRNP formation. *Molecular Cell* **49**(4): 692-703.

Grossman SR, Deato ME, Brignone C, Chan HM, Kung AL, Tagami H, Nakatani Y and Livingston DM (2003). Polyubiquitination of p53 by a ubiquitin ligase activity of p300. *Science* **300**(5617): 342–344.

Hanahan D and Weinberg RA. Hallmarks of cancer: The next generation. *Cell* **144**(5): 646-674.

Hanlon L, Avila JL, Demarest RM, Troutman S, Allen M, Ratti F, Rustgi AK, Stanger BZ, Radtke F, Adsay V, Long F, Capobianco AJ and Kissil JL (2010). Notch1 functions as a tumor suppressor in a model of K-ras-induced pancreatic ductal adenocarcinoma. *Cancer Research* **70**(11):4280-4286.

Hanzelmann P, Stingele J, Hofmann K, Schindelin H and Raasi S (2010). The yeast E4 ubiquitin ligase Ufd2 interacts with the ubiquitin-like domains of Rad23 and Dsk2 via a novel and distinct ubiquitin-like binding domain. *Journal of Biological Chemistry* **285**(26): 20390-20398.

Harris CC (2006). Protein-protein interactions for cancer therapy. *Proceedings of the National Academy of Sciences* **103**(6): 1659-1660.

Harper J, Yan L, Loureiro RM, Wu I, Fang J, D'Amore PA and Moses MA (2007). Repression of vascular endothelial growth factor expression by the zinc finger transcription factor ZNF24. *Cancer Research* **67**(18): 8736-8741.

Hashizume R, Fukuda M, Maeda I, Nishikawa H, Oyake D, Yabuki Y, Ogata H and Ohta T (2001). The RING heterodimer BRCA1-BARD1 is a ubiquitin ligase inactivated by a breast cancer-derived mutation. *Journal of Biological Chemistry* **276**(18): 14537-14540.

Haymond A, Dey D, Carter R, Dailing A, Nara V, Nara P, Venkatayogi S, Paige MA, Liotta LA and Luchini A (2019). Protein painting, an optimized MS-based technique, reveals functionally relevant interfaces of the PD-1/PD-L1 complex and the YAP2/ZO-1 complex. *Journal of Biological Chemistry* **294**(29): 11180-11198.

Heneghan C, Blacklock C, Perera R, Davis R, Banerjee A, Gill P and Ward A (2013). Evidence for non-communicable diseases: analysis of Cochrane reviews and randomised trials by World Bank classification. *British Medical Journa*l **3**(7): e003298.

Hermann H, Fabrizio P, Raker VA, Foulaki K, Hornig H, Brahms H and Lührmann R (1995). snRNP Sm proteins share two evolutionarily conserved sequence motifs which are involved in Sm protein–protein interactions. *The EMBO Journal* **14**(9): 2076-2088.

Hoppe T (2005). Multiubiquitylation by E4 enzymes: 'one size' doesn't fit all. *Trends in Biochemical Science* **30**(4): 183-187.

Hu R, Peng G, Dai H, Breuer EK, Stemke-Hale K, Li K, Gonzalez-Angulo AM, Mills GB and Lin SY (2011). ZNF668 functions asa tumor suppressor by regulating p53 stability and function in breast cancer. *Cancer Research* **71**(20): 6524-6534.

Hull R, Oosthuysen B, Cajee U, Mokgohloa L, Nweke E, Antunes RJ, Coetzer THT 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.

Hyjek-Składanowska M, Bajczyk M, Gołębiewski M, Nuc P, Kołowerzo-Lubnau A, Jarmołowski A and Smoliński DJ (2020). Core spliceosomal Sm proteins as constituents of cytoplasmic mRNPs in plants. *The Plant Journal* **103**(3): 1155-1173.

Jecklina MC, Schauerb S, Dumelinc CE and Zenobi R (2009). Label-free determination of protein–ligand binding constants using mass spectrometry and validation using surface plasmon resonance and isothermal titration calorimetry. *Journal of Molecular Recognition* **22**(4): 319-329.

Jemal A, Siegel R, Xu J and Ward E (2010). Global cancer statistics. *A Cancer Journal for Clinicians* **60**(5): 277-330.

Jemal A, Bray F, Center MM, Ferlay J, Ward E and Forman D (2011). Global cancer statistics. *CA: A Cancer Journal for Clinicians* **61**(2): 69–90.

Jen J, Lin LL, Chen HT, Liao SY, Lo FY, Tang YA, Su WC, Salgia R, Hsu CL, Huang HC and Juan HF (2015). Oncoprotein ZNF322A transcriptionally deregulates alpha-adducin, cyclin D1 and p53 to promote tumor growth and metastasis in lung cancer. *Oncogene* **35**(18): 2357-2369.

Jen J and Wang Y (2016). Zinc finger proteins in cancer progression. *Journal of Biomedical Science* 23(1): 1-9.

Jia D, Hasso SM, Chan J, Filingeri D, D'Amore PA, Rice L, Pampo C, Siemann DW, Zurakowski D, Rodig SJ and Moses MA (2013). Transcriptional repression of VEGF by ZNF24: mechanistic studies and vascular consequences in vivo. *Blood* **121**(4): 707-715.

Jiang M, Xu S, Yue W, Zhao X, Zhang L, Zhang C and Wang Y (2012). The role of ZFX in non-small cell lung cancer development. *Oncology Research Featuring Preclinical and Clinical Cancer Therapeutics* **20**(4): 171-178.

Joazeiro CA, Wing SS, Huang H, Leverson JD, Hunter T and Liu YC (1999). The tyrosinekinase negative regulator c-Cbl as a RING-type, E2-dependent ubiquitin–protein ligase. *Science* **286**(5438): 309–312.

Johnson PF, Cavanna T, Zicha D and Bates PA (2006). Cluster analysis of networks generated through homology: automatic identification of important protein communities involved in cancer metastasis. *BMC Bioinformatics* **7**(1): 1-13.

Jones DT, Tress M, Bryson K and Hadley C (1999). Successful recognition of protein folds using threading methods biased by sequence similarity and predicted secondary structure. *Proteins: Structure, Function, and Bioinformatics* 37(S3): 104-111.

Kappo MA, Eiso AB, Hassem F, Atkinson RA, Faro A, Muleya V, Mulaudzi T, Poole JO, McKenzie JM, Chibi M, Moolman-Smook JC, Rees DJG and Pugh DJR (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**(10): 7146-7158.

Kambach C, Walke S, Young R, Avis JM, de la Fortelle E, Raker VA, Lührmann R, Li J and Nagai K. (1999). Crystal structures of two Sm protein complexes and their implications for the assembly of the spliceosomal snRNPs. *Cell* **96**(3): 375-387.

Kerscher O, Felberbaum R and Hochstrasser M. (2006). Modification of proteins by ubiquitin and ubiquitin-like proteins. *Annual Review of Cell and Developmental Biology* **22**: 159-180.

Khan F, Allam M, Tincho MB and Pretorius A (2014). Implications of RBBP6 in various types of Cancer. *Proceedings IWBBIO* 2014. Granada 7-9 April 2014.

Khusial P, Plaag R and Zieve GW (2005). LSm proteins form heptameric rings that bind toRNA via repeating motifs. *Trends in Biochemical Sciences* **30**(9): 522-528.

Kostic M, Matt T, Martinez-Yamout MA, Dyson HJ and Wright PE (2006). Solution Structure of the Hdm2 C2H2C4 RING, a Domain Critical for Ubiquitination of p53. *Journal of Molecular Biology* **363**(2): 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.

Kroiss M, Schultz J, Wiesner J, Chari A, Sickmann A and Fischer U (2008). Evolution of an RNP assembly system: a minimal SMN complex facilitates formation of UsnRNPs in Drosophila melanogaster. *Proceedings of the National Academy of Sciences* **105**(29): 10045-10050.

Kwan AH, Mobli M, Gooley PR, King GF and Mackay JP (2011). Macromolecular NMR spectroscopy for the non-spectroscopist. *The FEBS Journal* **278**(5): 687-703.

Laemmli UK (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**(5259): 680-685.

Lai KP, Chen J, He M, Ching AK, Lau C, Lai PB, To KF and Wong N (2014). Overexpression of ZFX confers self-renewal and chemoresistance properties in hepatocellular carcinoma. *International journal of cancer* **135**(8): 1790-1799.

Lefebvre S, Bürglen L, Reboullet S, Clermont O, Burlet P, Viollet L, Benichou B, Cruaud C, Millasseau P, Zeviani M, Paslier D, Frézal J, Cohen D, Weissenbach J, Munnich A and Melki J (1995). Identification and characterization of a spinal muscular atrophy-determining gene. *Cell* **80**(1): 155–165.

Lopato S, Gattoni R, Fabini G, Stevenin J and Barta A (1999). A novel family of plant splicing factors with a Zn knuckle motif: examination of RNA binding and splicing activities. *Plant Molecular Biology* **39**(4): 761-773.

Lorber B, Sauter C, Theobald-Dietrich A, Moreno A, Schellenberger P, Robert MC, Capelle B, Sanglier S, Potier N and Giege R (2009). Crystal growth of proteins, nucleic acids, and viruses in gels. *Progress in Biophysics and Molecular Biology* **101**(1-3): 13-25.

Lorson CL, Rindt H and Shababi M (2010). Spinal muscular atrophy: Mechanisms and therapeutic strategies. *Human Molecular Genetics* **19**(R1): R111–R118.

Lovell SC, Davis IW, Arendall WB, De Bakker PI, Word JM, Prisant MG, Richardson JS and Richardson DC (2003). Structure validation by Cα geometry: ϕ, ψ and Cβ deviation. *Proteins: Structure, Function, and Bioinformatics* **50**(3): 437-450.

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

Li Y, Tan BB, Zhao Q, Fan LQ, Wang D and Liu Y (2014a). ZNF139 promotes tumor metastasis by increasing migration and invasion in human gastric cancer cells. *Neoplasma* **61**(3): 291-298.

Li Y, Zhao Q, Fan LQ, Wang LL, Tan BB, Leng YL, Liu Y and Wang D (2014b). Zinc finger protein 139 expression in gastric cancer and its clinical significance. *World Journal of Gastroenterology* **20**(48): 18346- 18353.

Li Y, Yan X, Yan L, Shan Z, Liu S, Chen X, Zou J, Zhang W and Jin Z (2015). High expression of Zinc finger protein X-linked is associated with reduced E-cadherin expression and unfavorable prognosis in nasopharyngeal carcinoma. *International journal of clinical and experimental pathology* **8**(4): 3919-3927.

Luhrmann R, Kastner B and Bach M (1990). Structure of spliceosomal snRNPs and their role in pre-mRNA splicing. *Biochimica et Biophysica Acta (BBA)-Gene Structure and Expression* **1087**(3): 265-292.

Ma H, Yang F, Lian M, Wang R, Wang H, Feng L, Shi Q, Fang J (2015). Dysregulation of zinc finger protein, X-linked (ZFX) impairs cell proliferation and induces apoptosis in human oral squamous cell carcinorma. *Tumor Biology* **36**(8): 6103-6112.

Ma X, Huang M, Wang Z, Liu B, Zhu Z and Li C (2016). ZHX1 inhibits gastric cancer cell growth through inducing cell-cycle arrest and apoptosis. *Journal of Cancer* **7**(1): 60-68.

Marabti E and Younis I (2018). The Cancer Spliceome: Reprograming of Alternative Splicing in Cancer. *Frontiers in Molecular Biosciences* **5**: 80.

Matera A and Wang Z (2014). Erratum: A day in the life of the spliceosome. *Nature Reviews Molecular Cell Biology* **15**(2): 108-121.

Mather A, RakgothoM and Ntwasa M (2005). SNAMA, a novel protein with a DWNN domain and a RING finger-like motif: A possible role in apoptosis. *Biochimica et Biophysica Acta (BBA)-Gene Structure and Expression***1727**(3): 169-176.

Mbita Z, Meyer M, Skepu A, Hosie M, Rees J and Dlamini Z (2012). De-regulation of the RBBP6 isoform 3/DWNN in human cancers. *Molecular and Cellular Biochemistry* **362**(1): 249-262.

McGuffin LJ, Bryson K and Jones DT (2000). The PSIPRED protein structure prediction server. *Bioinformatics* **16**(4): 404-405.

Meister G, Eggert C, Bühler D, Brahms H, Kambach C and Fischer U (2001). "Methylation of Sm proteins by a complex containing PRMT5 and the putative U snRNP assembly factor pICln". *Current Biology* **11**(24): 1990-1994.

Miotto B, Chibi M, Xie P, Koundrioukoff S, Moolman-Smook H, Pugh D, Debatisse M, He F, Zhang L and Defossez P (2014). The RBBP6/ZBTB38/MCM10 Axis Regulates DNA Replication and Common Fragile Site Stability. *Cell Reports* **7**(2): 575–587.

Moela P, Choene MMS 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.

Montzka KA and Steitz JA (1988). Additional low-abundance human small nuclear ribonucleoproteins: U11, U12, etc. *Proceedings of the National Academy of Sciences* **85**(23): 8885–8889.

Moon S and Lee B (2018). Chemically Induced Cellular Proteolysis: An Emerging Therapeutic Strategy for Undruggable Targets. *Molecules and Cells* **41**(11): 933-942

Morgera FL, Vaccari L, Antcheva N and Tossi A (2010). "Structure and activities on antimicrobial peptides at the bacterial membrane" *In*: M.A.R.B. Castanho," Membrane-Active Peptides: Methods and Results on Structure and Function", IUL Biotechnology Series, La Jolla USA.

Motadi LR, Bhoola KD and Dlamini Z (2011) Expression and function of retinoblastoma binding protein 6 (RBBP6) in human lung cancer. *Immunobiology* **216**(10): 1065-1073

Muleya V. (2010). Structural characterisation of the interaction between RBBP6 and the multifunctional protein YB-1. Msc Thesis. Department of Biotechnology, University of Western Cape. South Africa.URI: http://hdl.handle.net/11394/2324 [Accessed March 27, 2017].

Mura C. (2002). The Structures, Functions, and Evolution of Sm-like Archaeal Proteins (SmAPs). PhD Thesis. University of California. Available at: https://arxiv.org/pdf/1606.03630.pdf/ [Accessed: 05 December 2019].

Myszka DG (1997). Kinetic analysis of macromolecular interactions using surface plasmon resonance biosensors. *Current Opinion in Biotechnology* **8**(1): 50–57.

Myszka DG (1999). Survey of the 1998 optical biosensor literature. *Journal of Molecular Recognition* **12**(6): 390- 408.

Navratilova I, Papalia GA, Rich RL, Bedinger D, Brophy S, Condon B, Deng T, Emerick AW, Guan HW, Hayden T, Heutmekers T, Hoorelbeke B, McCroskey MC, Murphy MM, Nakagawa T, Parmeggiani F, Qin X, Rebe S, Tomasevic N, Tsang T, Waddell MB, Zhang FF, Leavitt S, Myszka DG (2007). Thermodynamic benchmark study using Biacore technology. *Analytical Biochemistry* **364**(1): 67-77.

Neuenkirchen N, Chari A and Fischer U (2008). Deciphering the assembly pathway of Sm-class U snRNPs. *FEBS Letters* **582**(14): 1997–2003.

Neuenkirchen N, Englbrecht C, Ohmer J, Ziegenhals T, Chari A and Fischer U (2015). Reconstitution of the human U snRNP assembly machinery reveals stepwise Sm protein organization. *The EMBO Journal* **34**(14): 1925-1941.

Nikolay R, Wiederkehr T, Rist W, Kramer G, Mayer MP and Bukau B (2004) Dimerization of the human E3 ligase CHIP via a coiled-coil domain is essential for its activity. *Journal of Biological Chemistry* **279**(4): 2673-2678.

O'Connell MR, Gamsjaeger R and Mackay JP (2009). The structural analysis of protein- protein interactions by NMR spectroscopy. *Proteomics* **9**(23): 5224-5232.

Ohi MD, Vander Kooi CW, Rosenberg JA, Chazin WJ and Gould KL (2003). Structural insights into the U-box, a domain associated with multi-ubiquitination. *Nature Structural & Molecular Biology* **10**(4),250-255.

Ohtani M (2018). Plant snRNP biogenesis: a perspective from the nucleolus and cajal bodies. *Frontiers in Plant Science* **8**: 2184.

Onyemata J, Meyer M, McKenzie JM, Rees DJG and Pugh DJR (2005). Investigation of the binding of ceramide and palmitoyl-CoA to murine t-ACBP using heteronuclear NMR spectroscopy. *South African Journal of Science* **101**(9): 430-434.

Palfi Z, Lucke S, Lahm HW, Lane WS, Kruft V, Bragado-Nilsson E, Seraphin B and Bindereif A (2000). The spliceosomal snRNP core complex of Trypanosoma brucei: cloning and functional analysis reveals seven Sm protein constituents. *Proceedings of the National Academy of Sciences* **97**(16): 8967-8972.

Pant V and Lozano G (2016). Limiting the power of p53 through the ubiquitin proteasome pathway. *Genes & Development* **28**(16): 1739-1751

Pellizzoni L, Yong J and Dreyfuss G (2002). Essential role for the SMN complex in the specificity of snRNP assembly. *Science* **298**(5599):1775-1779.

Perkins DN, Pappin DJ, Creasy DM and Cottrell JS (1999). Probability-based protein identification by searching sequence databases using mass spectrometry data. *ELECTROPHORESIS: An International Journal* **20**(18): 3551-3567.

Perozzo R, Folkers G and Scapozza L (2004). Thermodynamics of protein-ligand interactions: History, presence, and future aspects. *Journal of Receptors and Signal Transduction* 24(1-2): 1-52.

Petre BA, Youhnovski N, Lukkari J, Weber R and Przybylski M (2005). Structural characterisation of tyrosine-nitrated peptides by ultraviolet and infrared matrix-assisted laser desorption/ ionisation Fourier transform ion cyclotron resonance mass spectrometry. *European Journal of Mass Spectrometry* **11**(5): 513-518.

Pierce MM, Raman CS and Nall BT (1999). Isothermal titration calorimetry of protein- protein interactions. *Methods* **19**(2): 213-221.

Pugh DJR (2005). Biomolecular NMR spectroscopy in South Africa: the first five years. *South African Journal of Science* **101**(9): 421-429.

Pugh DJR, Eiso AB, Faro A, Lutya PT, Hoffmann E and Rees DJG (2006). DWNN, a novel ubiquitin-like domain, implicates RBBP6 in mRNA processing and ubiquintin-like pathway. *BMC Structural Biology* **6**(1): 1-12.

Prusty AB, Meduri R, Prusty BK, Vanselow J, Schlosser A and Fischer U (2017). Impaired spliceosomalUsnRNP assembly leads to Sm mRNA down-regulation and Sm protein degradation. *Journal of Cell Biology* **216**(8): 2391-2407.

Radloff R, Bauer W and Vinograd J (1967). A dye-buoyant-density method for the detection and isolation of closed circular duplex DNA: The closed circular DNA in Hela Cells. *Proceedings of the National Academy of Sciences* **57**(5): 1514-1521.

Randfontein herald, January 10 2017. Shocking cancer facts and statistics. https://randfonteinherald.co.za/237824/shocking-cancer-facts-and-statistics/ [Accessed on 23-04-2018].

Sable R and Jois S (2015). Surfing the Protein-Protein Interaction Surface Using Docking Methods: Application to the Design of PPI Inhibitors. *Molecules* **20**(6): 11569-11603.

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**(1): 98-101.

Sali A and Blundell TL (1993). Comparative protein modeling by satisfaction of spatial restraints. *Journal of Molecular Biology* **234**(3): 779-815.

Schasfoort RBM and Tudos AJ (2008). Handbook of Surface Plasmon Resonance. Cambridge: RCS Publishing.UK, p401.

Schwer B, Kruchten J and Shuman S (2016). Structure–function analysis and genetic interactions of the SmG, SmE, and SmF subunits of the yeast Sm protein ring. *RNA* **22**(9): 1320-1328.

Serra RW, Fang M, Park SM, Hutchinson L and Green MR (2014). A KRAS-directed transcriptional silencing pathway that mediates the CpG island methylator phenotype. *Elife* **3**: e02313.

Sharon MK, Thomas JJ and Nicholas CP (2005). How to study proteins by circular dichroism.

Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics **1751**(2): 119-139.

Shi Y, Di Giammartino DC, Taylor D, Sarkeshik A, Rice WJ, Yates JR, Frank J and Manley JL (2009). Molecular architecture of the human pre-mRNA 3′-processing complex. *Molecular Cell* **33**(3): 365-376.

Shukla S and Parker R (2014). Quality control of assembly-defective U1 snRNAs by decapping and 5′-to-3′ exonucleolytic digestion. *Proceedings of the National Academy of Sciences* **111**(32): E3277-E3286.

Siegel R, Ma J, Zou Z and Jemal A (2014). Cancer statistics, 2014. *CA: A Cancer Journal for Clinicians* **64**(1): 9-29.

Siegel R.L., Miller K.D. and Jemal A. (2015). Cancer Statistics, 2015. *A Cancer Journal for Clinicians* **65**(1): 5-29.

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**(2): 145-155.

Smith AL (1997). "Applied infrared spectroscopy: fundamentals, techniques, and analytical problem-solving", A Wiley-Interscience Publication, John Wiley and Sons, New York, Chichester, Brisbane, and Toronto

Sorolla A, Wang E, Golden E, Duffy C, Henriques ST, Redfern AD and Blancafort P (2019). Precision medicine by designer interference peptides: applications in oncology and molecular therapeutics. *Oncogene* **39**(6): 1167-1184.

Spengler B and Cotter RJ (1990). Ultraviolet laser desorption/ionization mass spectrometry of proteins above 100,000 daltons by pulsed ion extraction time-of-flight analysis. *Analytical*

Chemistry **62**(8): 793-796.

Stark C, Breitkreutz BJ, Reguly T, Breitkreutz A and Tyers M (2006). BioGRID: a general repository for interaction datasets. *Nucleic Acids Research* **34**(suppl_1): D535-D539.

Stark H, Dube P, Lührmann R and Kastner B (2001). Arrangement of RNA and proteins in the spliceosomal U1 small nuclear ribonucleoprotein particle. *Nature* **409**(6819): 539-42.

Stevens SW and Abelson J (1999). Purification of the yeast U4/U6·U5 small nuclear ribonucleoprotein particle and identification of its proteins. *Proceedings of the National Academy of Sciences* **96**(13): 7226-7231.

Steinbrecher T and Labahn A (2010). Towards accurate free energy calculations in ligand protein-binding studies. *Current medicinal chemistry* **17**(8): 767-785.

Takagi T, Naito Y, Okada H, Okayama T, Mizushima K, Yamada S, Fukumoto K, Inoue K, Takaoka M, Oya-Ito T, Uchiyama K, Ishikawa T, Handa O, Kokura S, Yagi N, Ichikawa H,Kato Y, Osawa T and Yoshikawa T (2011). Identification of dihalogenated proteins in rat intestinal mucosa injured by indomethacin. *Journal of Clinical Biochemistry and Nutrition* **48**(2): 178-182.

Urlaub H, Raker VA, Kostka S and Lührmann R (2001). Sm protein–Sm site RNA interactions within the inner ring of the spliceosomal snRNP core structure. *The EMBO Journal* **20**(1-2): 187–196.

Valeur E, Narjes F, Ottmann C and Plowright AT (2019). Emerging modes-of-action in drug discovery. *MedChemComm* **10**(9): 1550-1568.

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

Wang J, Liu D, Liang X, Gao L, Yue X, Yang Y, Ma C and Liu J (2013). Construction of a recombinant eukaryotic human ZHX1 gene expression plasmid and the role of ZHX1 in hepatocellular carcinoma. *Molecular Medicine Reports* 8: 1531-1536.

Wang Z, Ma X, Cai Q, Wang X, Yu B, Cai Q, Zhu Z and Li C (2014). MiR-199a-3p promotes gastric cancer progression by targeting ZHX1. *FEBS Letters* **588**(23): 4504-4512.

Wahl MC, Will CL and Lührmann R (2009). The spliceosome: Design principles of a dynamic RNP machine. *Cell* **136**(4): 701-718.

Weber G, Trowitzsch S, Kastner B, Lührmann R and Wahl MC (2010). Functional organization of the Sm core in the crystal structure of human U1 snRNP. *EMBO Journal* **29**(24): 4172- 4184.

Weng H, Wang X, Li M, Wu X, Wang Z, Wu W, Zhang Z, Zhang Y, Zhao S, Liu S and Mu J (2015). Zinc finger X-chromosomal protein (ZFX) is a significant prognostic indicator and promotes cellular malignant potential in gallbladder cancer. *Cancer Biology & Therapy* **16**(10): 1462-70.

Will CL and Luhrmann R (2011). Spliceosome structure and function. *Cold Spring Harbor Perspectives in Biology* **3**(7): a003707.

Wishart DS, Knox C, Guo AC, Cheng D, Shrivastava S, Tzur D, Gautam B and Hassanali MD (2008). A knowledgebase for drugs, drug actions and drug targets. *Nucleic Acids Research* **36**(suppl_1): D901-D906.

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.

World Health Organisation. (2018). Riviere du Rempart Youth Center, 08 February 2018.

Wu S, Lao XY, Sun TT, Ren LL, Kong X, Wang JL, Du W, Yu YN, Weng YR and Hong J (2013). Knockdown of ZFX inhibits gastric cancer cell growth in vitro and in vivo via downregulating the ERK-MAPK pathway. *Cancer Letters* **337**(2): 293-300.

Xu GG, Guo J and Wu Y (2014). Chemokine receptor ccr5 antagonist maraviroc: Medicinal chemistry and clinical applications. *Current Topics in Medicinal Chemistry* **14**(13): 1504-1514.

Yang L, Hamilton SR, Sood A, Kuwai T, Ellis L, Sanguino A, Lopez-Berestein G and Boyd DD (2008a). The previously undescribed ZKSCAN3 (ZNF306) is a novel "driver" of colorectal cancer progression. *Cancer Research* **68**(11): 4321-4330.

Yang L, Zhang L, Wu Q and Boyd DD (2008b). Unbiased screening for transcriptional targets of ZKSCAN3 identifies integrin beta 4 and vascular endothelial growth factor as downstream targets. *Journal of Biological Chemistry* **283**(50): 35295-35304.

Yang L, Wang H, Kornblau SM, Graber DA, Zhang N, Matthews JA, Wang M, Weber DM, Thomas SK, Shah JJ and Zhang L (2011). Evidence of a role for the novel zinc-finger transcription factor ZKSCAN3 in modulating Cyclin D2 expression in multiple myeloma. *Oncogene* 30(11): 1329-1340.

Yang H, Lu Y, Zheng Y, Yu X, Xia X, He X, Feng W, Xing L and Ling Z (2014). shRNA-mediated silencing of ZFX attenuated the proliferation of breast cancer cells. *Cancer Chemotherapy and Pharmacology* **73**(3): 569- 576.

Yang J, Yan R, Roy A, Xu D, Poisson J and Zhang Y (2015). The I-Tasser Suite: Protein structure and function prediction. *Nature Methods* **12**(1): 7-8.

Yoshitake Y, Nakatsura T, Monji M, Senju S, Matsuyoshi H, Tsukamoto H, Hosaka S, Komori

H, Fukuma D, Ikuta Y, Katagiri T, Furukawa Y, Ito H, Shinohara M, Nakamura Y and Nishimura Y (2004). Proliferation Potential-Related Protein, an Ideal Esophageal Cancer Antigen for Immunotherapy, Identified Using Complementary DNA Microarray Analysis. *Clinical Cancer Research* **10**(19): 6437-6448,

Yu J, Liang QY, Wang J, Cheng Y, Wang S, Poon TC, Go MY, Tao Q, Chang Z and Sung JJ (2013). Zinc-fingerprotein 331, a novel putative tumor suppressor, suppresses growth and invasiveness of gastric cancer. *Oncogene* **32**(3): 307-317.

Zhang R, So BR, Li P, Yong J, Glisovic T, Wan L and Dreyfuss G (2011). Structure of a key intermediate of the SMN complex reveals Gemin2's crucial function in snRNP assembly. *Cell* **146**(3): 384-395.

Zhang X, Jing Y, Qin Y, Hunsucker S, Meng H, Sui J, Jiang Y, Gao L, An G, Yang N and Orlowski RZ (2012) The zinc finger transcription factor ZKSCAN3 promotes prostate cancer cell migration. *The International Journal of Biochemistry & Cell Biology* **44**(7): 1166-1173.

Zhu Z, Li K, Xu D, Liu Y, Tang H, Xie Q, Liu J, Wang H, Gong Y and Hu Z. (2013). ZFX regulates glioma cell proliferation and survival in vitro and in vivo. *Journal of Neuro-oncology* **112**(1): 17-25.

CHAPTER TWO

A paper already published in *Biophysical Reviews* 2019; 11: 559-581.
https://doi.org/10.1007/s12551-019-00570-x

Protein-Protein Interaction Modulators: Advances, Successes and Remaining Challenges

Lloyd Mabonga and Abidemi Paul Kappo

Preface − About the Manuscript

The manuscript is about the protein–protein interactions (PPIs) modulators as potential tools in the field of smart-drug discovery. The modulation of PPIs is considered a promising strategy towards next-generation therapeutics. We therefore expound on the advances, successes and the remaining challenges in this fascinating field of science. Over the years the proteome space has been prototypically branded as the 'undruggable and intractable' world and for a very long time this concept has obliterated the potential of PPIs to yield new drugs. However, this topic is slowly gathering attention and it is likely that very soon, based on improvements in technology, this space may become not so much "undruggable" in the near future. Drugging this area of macromolecular science seeks to expand and explore the intractable space as well as provide novel solutions to many diseases that have been deemed incurable. In this context, the manuscript highlights the potential of finding small-molecule inhibitors for the two oncogenic proteins, SNRPG and RBBP6,whose functions are predominantly mediated by PPIs. As the quest for the cure rages on, drugging the undruggable proteome space remains the drug discovery anecdote in cancer. Therefore, the manuscript presents the milestones and challenges of this underexplored field of research.

Protein-protein interaction modulators: advances, successes and remaining challenges

Lloyd Mabonga & Abidemi Paul Kappo

Biophysical Reviews

ISSN 1867-2450

Biophys Rev
DOI 10.1007/s12551-019-00570-x

Biophysical Reviews
https://doi.org/10.1007/s12551-019-00570-x

REVIEW

Protein-protein interaction modulators: advances, successes and remaining challenges

Lloyd Mabonga[1] · Abidemi Paul Kappo[1]

Received: 29 May 2019 / Accepted: 24 June 2019

Abstract

Modulating disease-relevant protein-protein interactions (PPIs) using small-molecule inhibitors is a quite indispensable diagnostic and therapeutic strategy in averting pathophysiological cues and disease progression. Over the years, targeting intracellular PPIs as drug design targets has been a challenging task owing to their highly dynamic and expansive interfacial areas (flat, featureless and relatively large). However, advances in PPI-focused drug discovery technology have been reported and a few drugs are already on the market, with some potential drug-like candidates already in clinical trials. In this article, we review the advances, successes and remaining challenges in the application of small molecules as valuable PPI modulators in disease diagnosis and therapeutics.

Keywords Protein-protein interactions · Small molecules · Modulators · Drug-like · Macrocycles · Small-molecule inhibitors

Abbreviations

PPI	Protein-protein interactions
IL-2	Interleukin-2
IL-2Rα	Interleukin-2 receptor alpha chain
FBDD	Fragment-based drug discovery
FBLD	Fragment-based lead discovery
SAR	Structure-activity relationship
NMR	Nuclear magnetic resonance
Bcl-2	B cell lymphoma 2
Bcl-XL	B cell lymphoma-extra large
CS	Computational solvent
MDM2	Mouse double minute 2 homolog
HPV	Human papillomavirus
TNF	Tumour necrosis factor
TNF-α	Tumour necrosis factor alpha
TNFR1	Tumour necrosis factor receptor 1
FtsZ	Filamenting temperature-sensitive mutant Z
ZipA	Z interacting protein A
HTS	High-throughput screening
CBF	Core binding factor
DOS	Diversity-oriented synthesis
Shh	Sonic Hedgehog
Ptc1	Protein phosphatase 2C homolog 1
mTOR	Mammalian target of rapamycin
FKBP12	FK-binding protein 12
DPC	DNA-programmed chemistry
STATs	Signal transducers and activators of transcription
mAb	Monoclonal antibody
HER2	Human epidermal growth factor receptor 2
ICAM-1	Intercellular adhesion molecule-1
VEGF	Vascular endothelial growth factor
G-CSF	Granulocyte colony-stimulating factor
TPO	Thrombopoietin
LFA-1	Leukocyte integrin lymphocyte function-associated antigen 1
GPCR	G protein-coupled receptor
CCR5	C-C chemokine receptor type 5
CXCR4	C-X-C chemokine receptor type 4
CXCR7	C-X-C chemokine receptor type 7
MM	Multiple myeloma
B-CLL	B cell chronic lymphocytic leukaemia
APC	Adenomatous polyposis coli
GSK-3	Glycogen synthase kinase 3
Tcf	T cell factor

This article is part of a Special Issue on 'Biophysics & Structural Biology at Synchrotrons' edited by Trevor Sewell.

* Abidemi Paul Kappo
KappoA@unizulu.ac.za

[1] Biotechnology and Structural Biology (BSB) Group, Department of Biochemistry and Microbiology, University of Zululand, KwaDlangezwa 3886, South Africa

Published online: 12 July 2019

DLBCL	Diffuse large B cell lymphoma
BCL6	B cell lymphoma 6
SMRT	Silencing mediator for retinoid or thyroid-hormone receptors
HDAC3	Histone deacetylase 3
CHK1	Checkpoint kinase 1
CDKN1A	Cyclin-dependent kinase inhibitor 1
ATR	Ataxia telangiectasia and Rad3-related protein
CADD	Computer-aided drug design
AKAP	A-kinase anchoring protein
cAMP	Cyclic adenosine monophosphate
Ub	Ubiquitin
UPS	Ubiquitin proteasome system
UPP	Ubiquitin-proteasome pathway
CRL	Cullin RING E3 ligase
K_i	Inhibition constant
TRAIL	TNF-related apoptosis-inducing ligand
SCLC	Small cell lung cancer
PACs	Pancreatic acinar cells
SAHBs	Stabilised alpha-helix of Bcl-2 domains
T-ALL	T cell acute lymphoblastic leukaemia
ADAM	A disintegrin and metalloproteinase
ICN1	Intracellular domain of NOTCH1
MAML	Mastermind-like
dnMAML1	Dominant-negative fragment of MAML1

Introduction

Modern drug discovery is driven by molecular targets with the aim of identifying new therapeutic agents that can selectively target disease-specific molecular mechanisms or pathways (Díaz-Eufracio et al. 2018). In this context, protein-protein interactions (PPIs) are an attractive emerging class of molecular targets and are critically important in the progression of many disease states (Robertson and Spring 2018; Zhang et al. 2018). PPIs are engineered to provide a therapeutically tractable way of tweaking and manipulating the interplay in order to address the progression of many disease states (Du et al. 2018). They are involved in hubs of reversible and irreversible cellular processes, assembling and disassembling rapidly, reassembling and rearranging in order to restore normative cellular functions (Robertson and Spring 2018). There are more than 645,000 reported disease-relevant PPIs in the human interactome. However, only 2% of these had been targeted with drugs by 2011. Most of the remaining disease-relevant PPIs in protein complexes, such as transcription factors and many other signalling proteins, have been widely considered 'undruggable' and remain elusive, under-explored and yet to be fully understood (Gonzalez and Kann 2012; Díaz-Eufracio et al. 2018; Robertson and Spring 2018; Zhang et al. 2018).

Inhibiting PPIs using small molecules is a tremendously important diagnostic and therapeutic strategy that may lead to greatly protracted remissions and even curative therapies for a number of diseases (Stevers et al. 2018). The emergence of new technologies has unveiled the potential of PPIs in drug discovery and has enabled regular discovery of small-molecule PPI modulators as significant smart-drug targets (Grossmann et al. 2015; Jana et al. 2017). Over the years, PPI-focused drug technology has been regarded as prototypically intractable because of the highly dynamic and expansive PPIs interfacial areas (Taylor et al. 2018). However, recent advances have resulted in a few drugs being placed on the market, with some potential drug-like candidates already in clinical trials. In this study, we review the advances, successes and remaining challenges in the application of small molecules as valuable PPI modulators in disease diagnosis and therapeutics.

Strategies for targeting protein-protein interactions

Over the years, technological progress has played an imperative role in the identification of small-molecule modulators of PPIs that have to date reached clinical production (Stevers et al. 2018). The use of structural biology to determine 'hotspots' in PPIs' binding interfaces has been an important strategy in discovering small-molecule modulators (Robertson and Spring 2018; Zhang et al. 2018). Despite the large sizes of PPIs' interfaces, only a small subset of amino acid residues that comprise the hotspot contributes most of the binding free energy. These 'hotspot' regions are potential targets for drug discovery (Zhang et al. 2018). A classic way of identifying and defining hotspots in PPIs has been the combination of alanine-scanning mutagenesis and X-ray crystallography (Moreira et al. 2007; Wells and McClendon 2007). The initial application of this strategy was used to identify a hotspot in the binding interface between the extracellular domain of human growth hormone and its receptor (Clackson and Wells 1995).

Using alanine-scanning mutagenesis, other classic PPI hotspots of high-fidelity protein regions, such as the Fc fragment hinge region-binding domain, have been identified (Wells and McClendon 2007). Mutagenesis and structural studies of the binding events of interleukin-2 (IL-2) and the IL-2 receptor alpha chain (IL-2Rα) provide more classical insight (Wilson and Arkin 2011). The first small molecule (Ro26-4550) capable of inhibiting the IL-2/IL-2Rα interaction was discovered in 1997 (Wilson and Arkin 2011). Despite the compound not qualifying as a drug, it provided proof-of-principle that small-molecule PPI inhibitor drug discovery or design might be feasible. Moreover, structural studies of the Ro26-4550/IL-2 complex helped to characterise the

IL-2 binding site, and served as the starting point for the design of higher affinity small-molecule IL-2/IL-2Rα PPI inhibitors (Wilson and Arkin 2011).

Fragment-based drug discovery (FBDD) (often referred to as fragment-based lead discovery, FBLD) is a key strategy for the discovery and design of small-molecule modulators of PPIs (Everts 2008; Erlanson et al. 2016; Robson-Tull 2018). It involves first identifying small chemical fragments (~ 200 Da), which may only bind at millimolar affinity to their targets (shown in Fig. 1). The fragments are then expanded or linked to other fragments that bind to nearby regions on the target in order to design a 'lead' with stronger affinity (Everts 2008; Erlanson et al. 2016). These 'leads' are then optimised via medicinal chemistry and may then be entered into preclinical and eventually clinical studies. Tethering and structure-activity relationship by nuclear magnetic resonance (SAR by NMR) are the two methods for FBDD/FBLD of potential modulators of PPIs (Erlanson et al. 2016; Robson-Tull 2018). Tethering involves constructing mutant forms of the target protein that contain the amino acid cysteine near a domain involved in PPIs (shown in Fig. 2), and then exposing it to a fragment library of disulphide moiety-linked organic compounds of less than 200 Da in molecular weight (Wilson and Arkin 2011). The goal is to select for compounds that bind weakly to the PPIs' binding site near a native or engineered cysteine residue. The tethering technology was used to explore the IL-2 binding site for Ro26-4550 and to discover the IL-2Rα receptor antagonist (SP4206), which is an inhibitor of the IL-2/IL-2Rα PPIs (Arkin et al. 2003; Wilson and Arkin 2011).

SAR by NMR involves the use of a high-throughput NMR technique to screen chemical libraries for fragment-sized compounds that bind to a protein sub-site with micromolar binding constants (Ma et al. 2016; Rüdisser et al. 2016). Using structural information from NMR to locate the binding sites for the compounds within a sub-site, ligands that bind to distinct but nearby sites within the sub-site can be identified. The two ligands can then be linked together to produce a new compound that binds to the domain with high affinity. This compound can thereafter be further optimised via medicinal chemistry to yield 'lead' compounds and ultimately, drug candidates (Ma et al. 2016; Robson-Tull 2018). SAR by NMR has been used to discover inhibitors of Bcl-2 family members that inhibit apoptosis (Oltersdorf et al. 2005). A very good example is ABT-263 (also known as Navitoclax, shown in Fig. 9a), which works by stimulating apoptosis in tumours (Duan et al. 2018).

Computational identification of hotspots for fragment-based drug design is another useful strategy in the discovery and design of lead compounds that address PPIs (Kozakov et al. 2011; Pan et al. 2018). It predicts the ability to bind fragment-sized small molecules and the side-chain flexibility necessary for the expansion of 'pockets'. This strategy is known as computational solvent (CS) mapping (Landon et al. 2007; Hall et al. 2012). The emergence of CS mapping has enabled the virtual identification of druggable binding sites within PPIs and the subsequent discovery or design of small-molecule PPI modulators (Cencic et al. 2011; Kozakov et al. 2011). Druggable hotspots for well-studied PPI targets that were identified via CS mapping include IL-2/IL-2Ra, Bcl-XL/BAK, MDM2/p53, HPV-11 E2/HPV-11 E1, ZipA/FtsZ and TNF-α/TNFR1. CS mapping constitutes a powerful enabling technology that can help move the field of small-molecule PPI inhibitors up the technology curve and enable the discovery and development of numerous drugs that target PPIs (Kozakov et al. 2011; Pan et al. 2018).

Another indirect strategy for discovering small-molecule PPI modulators involves searching for allosteric modulators. An allosteric site is a region of a protein that lies outside the binding site for the protein natural ligand, but when modulated, changes the conformation of the protein in such a way that it affects the activity of the binding site (Hansen et al. 2018; Trinh et al. 2018). For example, the anti-HIV/AIDS drug maraviroc, an allosteric PPI inhibitor of the CCR5 chemokine receptor, was discovered via high-throughput screening (HTS) (Dorr et al. 2005). Another serendipitous discovery of

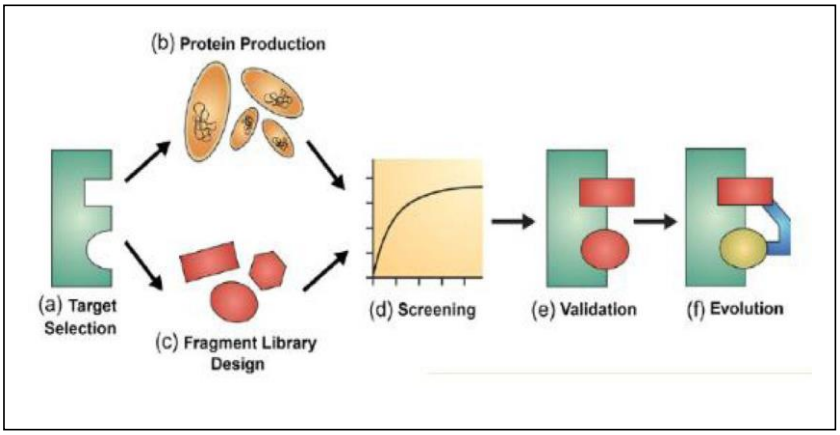


Fig. 1 Fragment-based drug discovery strategy. a Selection of a target compatible with the biophysical screening technique. b Production and purification of the target protein. c The fragment library design. d Biophysical screening of the fragment library. e Validation of hits to identify the fragment binding mode. f Development of the fragment(s) into a lead molecule (figure taken from Robson-Tull 2018)

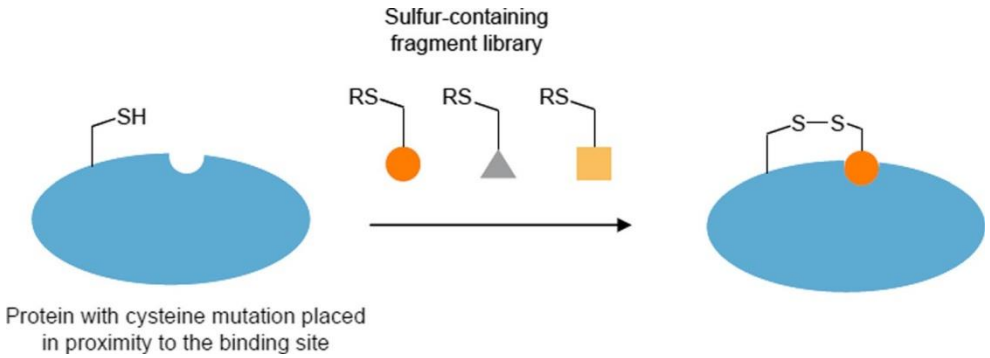


Fig. 2 The application of tethering in identifying leads in fragment-based drug design. Protein mutant forms with the cysteine mutation near a domain involved in PPIs are constructed. The mutant is exposed to a fragment library of disulphide moiety-linked organic compounds of less than 200 Da and compounds that bind weakly to the PPIs' binding site near a native or engineered cysteine residue are selected (figure taken from Haberman 2012)

small-molecule allosteric modulators of PPIs is the discovery of compounds that act at an allosteric site to inhibit the sub-units of the heterodimeric transcription factor core binding factor (CBF) CBFβ and Runx1 (also known as CBFα) (Gorczynski et al. 2007). Developing methods for identification of allosteric sites in proteins and discovering drugs that modulate the activity of these proteins by binding to these allosteric sites is now the focus of pharmaceutical studies (Hansen et al. 2005; Hansen et al. 2018; Trinh et al. 2018).

The design of improved chemical libraries for targeting PPIs using methods such as diversity-oriented synthesis (DOS) has also been of benefit to PPI-focused drug technology (Hajduk et al. 2011; Basso et al. 2019; Luise and Wyatt 2019). In DOS, chemical libraries are developed to cover larger portions of chemical space than the standard libraries derived from combinatorial chemistry. Synthetic schemes are developed to maximise the number of structures and scaffolds produced in as few steps as possible to fill the largest amount of chemical space (Nielsen and Schreiber 2008; Schreiber 2009; Luise and Wyatt 2019). DOS is the best way to create drug-size and structurally diverse molecules efficiently. It is applicable in cases where a drug for a specific disease has to be developed without knowledge of the specific targets involved in the pathophysiological cues (Basso et al. 2019; Luise and Wyatt 2019). Library screening of structurally diverse compounds using the chemical genetics approach augments the ideal strategy to identify the targets. DOS compounds that inhibit key biological targets such as PPIs involved in signal transduction pathways are ideal probes and such compounds are very useful in exploring biological pathways in vitro and in small, easily permeated organisms, such as zebrafish embryos (Nielsen and Schreiber 2008; Schreiber 2009).

In recent years, the synthesis of diverse libraries in as small a number of steps as possible has been developed through a modular strategy known as 'Build/Couple/Pair' (B/C/P) (Galloway et al. 2010). Compounds that have been discovered via screening of DOS libraries using such a modular strategy have been found to modulate PPIs, transcription factor/DNA interactions and multidrug resistance in pathogens (Galloway et al. 2010). One example is robotnikinin. It inhibits the interaction between the 12-pass transmembrane receptor Patched 1 (Ptc1) and the extracellular protein Sonic Hedgehog (Shh) (Hitzenberger et al. 2017; Carballo et al. 2018). The Shh/Ptc1 interaction activates the cancer-implicated Hedgehog pathway, which is necessary for embryonic development (Stanton et al. 2009; Carballo et al. 2018). Robotnikinin is a useful probe, not a drug, and like many macrocycles (robotnikinin is itself a macrocycle) has drug-like physicochemical properties (Stanton et al. 2009; Hitzenberger et al. 2017).

With the rise in DOS and the B/P/C strategy, macrocycles have also become an emerging and promising strategy in targeting PPIs (Driggers et al. 2008; Song et al. 2017). Macrocyclic natural products have provided many drugs, including macrolide antibiotics (e.g. erythromycin and azithromycin), other antibiotics (rifampin and vancomycin), immunosuppressors (e.g. cyclosporine, rapamycin and sirolimus) and cancer chemotherapy drugs (temsirolimus, everolimus and epothilone) (Song et al. 2017). Some of these compounds are modulators of PPIs, e.g. paclitaxel, epothilone B, dictyostatin and halichondrin B (Driggers et al. 2008; Miller et al. 2018). These compounds modulate (stabilise or disrupt) the interaction between α and β sub-units of the tubulin heterodimer, thus disturbing microtubule dynamics and acting as antimitotics. Other PPI modulators include the mammalian target of rapamycin (mTOR) inhibitor. Rapamycin (also known as sirolimus) forms a complex with FK-binding protein 12 (FKBP12). This complex forms a PPI with the mTOR complex 1, thus inhibiting its activity (Driggers et al. 2008). Rapamycin/sirolimus' anticancer derivatives temsirolimus and everolimus work via the same mechanism as well. Similarly, the macrocyclic natural product cyclosporine A, an immunosuppressant, forms a complex with cyclophilin A, which forms a PPI with calcineurin.

This complex inhibits the action of calcineurin, which when not inhibited activates the expression of IL-2 (Driggers et al. 2008).

Macrocycles demonstrate drug-like physicochemical properties with respect to factors such as solubility and lipophilicity. They display oral bioavailability, metabolic stability and good pharmacokinetic and pharmacodynamic properties (Driggers et al. 2008; Alihodzić et al. 2018). However, macrocyclic natural products have not received much attention in addressing challenging PPI targets and in solving such pressing problems as the need for new drugs to combat microbial antibiotic resistance (Alihodzić et al. 2018). The emergence of strategies for the synthesis of macrocycle-rich chemical libraries such as B/C/P and the recent breakthrough in respect of the synthesis of organic compounds, such as olefin metathesis, have put macrocycles within easier reach of medicinal chemists (Lee and Grubbs 2001; Yu et al. 2011). The synthesis of macrocyclic compounds via DNA-programmed chemistry (DPC) technology has also beneficiated the PPI-focused drug technology (Franzini and Randolph 2016; Zhou et al. 2018). DPC allows the control of high-fidelity chemical reactions needed for synthesis of the desired libraries through tagging specific reactants with hybridised DNA molecules. Once libraries are synthesised, screening for biological activity via affinity selection follows. The DPC strategy has led to the discovery of macrocycles, such as E-32712, that disrupt the interaction of tumour necrosis factor (TNF) with its receptor, TNFR (Drahl 2009; Franzini and Randolph 2016; Zhou et al. 2018).

These new technologies, combined with cellular screening assays (including high-content screening), help to discover and design PPI modulators. Cellular assays help to assess the activity of whole intracellular pathways or portions of pathways to identify compounds that inhibit these pathways via disruption of key PPIs (Mella et al. 2018; Booij et al. 2019). Two examples of cellular screening assays are ligand signal transducers and activators of transcription (STATs) technology and BioImage Redistribution technology. Ligand STATs technology takes advantage of the activation of the intracellular signal transduction pathway through STATs and high-throughput fluorescent reporter assay (Pándy-Szekeres et al. 2018), whereas BioImage redistribution technology focuses on pathways that involve intracellular translocation of a signalling protein from one intracellular compartment to another, such as from the cytoplasm to the nucleus or vice versa (Lage et al. 2018). Taken together, these strategies help to discover and design PPI modulators. Potent PPI modulators have already been developed and some potential drug-like candidates have already reached clinical production.

The challenges of targeting protein-protein interactions

Undruggability of PPIs

PPIs represent challenging targets for small-molecule drugs that have the potential to become available orally (Zhang et al. 2018). They have highly dynamic and expansive interfacial areas of approximately 1500-3000 Å, compared with those involved in interactions between proteins and small molecules, which are approximately 300-1000 Å (Gonzalez and Kann 2012; Robertson and Spring 2018). PPIs' contact surfaces are usually flat, featureless and relatively large. They are deficient in the kind of cavities present in the surfaces of proteins that bind to small-molecule ligands. Previously, PPIs were regarded as prototypically 'intractable' and 'undruggable' (Robertson and Spring 2018; Zhang et al. 2018). PPIs do not have natural small-molecule ligands; hence, these ligands cannot be used to initiate design of drug molecules. The contact surface area in PPIs often involves ramified amino acid residues whose sequences in the polypeptide chain are not juxtaposed. The amino acid residues are only augmented through the three-dimensional folding state of the native protein (Jochim and Arora 2010). This makes it impossible to use short peptide chains derived from the protein structure as starting points for the design of peptidomimetic drugs. Furthermore, high-throughput screening using combinatorial libraries rarely identifies compounds that address PPIs (Haberman 2012).

Target validation and druggability

In PPI-focused drug technology, target validation and druggability have been the focal points in selecting targets for drug discovery (Feng et al. 2017). Target validation refers to a process of determining that a target is critically involved in a disease pathway and that modulating the target with a drug is likely to have a positive therapeutic effect (Modell et al. 2016; Feng et al. 2017). However, druggable targets refer to biomolecules that can be modulated with drugs, usually using well-proven drug discovery science and technology aimed at developing both large-molecule and small-molecule drugs (Modell et al. 2016). Over the years, large-molecule drugs or biologics have been the fastest-growing and most successful class of biologics. Examples include monoclonal antibody (mAb) drugs and recombinant proteins (Sinha et al. 2012). Most of these drugs are involved in PPIs and are indicated for oncology and inflammatory diseases. They address appropriate targets, such as cell-surface receptors (e.g. HER2 in breast cancer and CD20 in non-Hodgkin's lymphoma) and cytokines (e.g. TNF-alpha (TNF-α) in inflammatory diseases such as rheumatoid arthritis) (Allison 2009; Feng et al. 2017).

Large-molecule biologics circumvent the druggability problems encountered when developing small-molecule drugs in PPI-focused drug technology (Wan 2016). However, they have limitations and setbacks. Large-molecule biologics do not address intracellular targets, are usually not orally available and are administered parenterally (Debouck and Metcalf 2000). Thus, small-molecule drugs have been synthesised to curb these challenges. Small-molecule drugs are orally available, less expensive, easy to administer and convenient for patients. The discovery of small-molecule drugs typically entails the use of medicinal and combinatorial chemistry, in some cases augmented by structure-based drug design (Debouck and Metcalf 2000; Wan 2016). The central goal of medicinal chemistry was to design and select small-molecule compounds that have 'drug-like properties' with good absorption, distribution, metabolism and elimination properties needed for orally deliverable drugs. Hence, the 'Rule of Five', a set of parameters that predicts a compound's solubility and permeability, was often used (Lipinski et al. 2012). However, natural products and some small-molecule drugs did not fit the Rule of Five criteria. This led to the ascendancy of two waves of technology-driven drug discovery techniques: combinatorial chemistry combined with HTS and genomics-driven drug discovery (Debouck and Metcalf 2000; Wan 2016).

Combinatorial chemistry produces libraries of small organic compounds, which are subjected to HTS to discover compounds that address targets derived from genomics (Taylor et al. 2018). Once an active skeleton has been identified, combinatorial chemistry is a superb technology for optimising the structures of lead compounds. The use of combinatorial chemistry and HTS permits the examination of large numbers of compounds in a short time, implying that a few drugs could reach the market and achieve blockbuster status (Li and Vederas 2009; Taylor et al. 2018). Moreover, emphasis have been on organic compounds that can be synthesised using combinatorial chemistry because they are more amenable to use in HTS and have relatively simpler structures than natural products. However, in the majority of cases, combinatorial libraries of these synthetic organic compounds are based on structural modification of existing drugs, thus rendering other classes of drug targets such as PPIs 'hard targets,' 'challenging targets,' or simply 'undruggable' (Bauer et al. 2010). Hence, approximately 10-14% of human proteins are druggable according to the current libraries of drug-like molecules. The use of current combinatorial libraries of synthetic organic molecules for de novo drug discovery has been sorely lacking (Newman and Cragg 2007).

In contrast to combinatorial libraries of synthetic compounds, natural products play a crucial role in the discovery of drug leads. Current synthetic libraries are indirectly based on natural products (Bauer et al. 2010; Wan 2016). Thus, the envisioned governing paradigm of medicinal and combinatorial chemistry is to develop new libraries based on natural product scaffolds that are underrepresented in current libraries. Approximately 83% of small natural product scaffolds and 20% of small metabolite scaffolds are not represented in commercially available libraries. These underrepresented scaffolds may enable to address what are now considered 'undruggable,' 'hard' or 'challenging' targets (Haberman 2012). Hence, drugs or drug candidates that modulate PPIs (for example natural products or natural product-like compounds), screening natural product libraries and synthesising natural product-like compounds remain one approach to discovering drugs that address PPIs (Bauer et al. 2010; Haberman 2012; Wan 2016).

Targeted protein-protein interactions

Cell-surface receptors

Many cell-surface receptors modulate their physiological functions through PPIs and these include receptors for cytokines, chemokines, growth factors and integrins. Integrins bind to extracellular matrix proteins or to members of the immunoglobulin superfamily, such as intercellular adhesion molecule-1 (ICAM-1) and Fc receptors (which bind to the Fc regions of antibodies) (Guidolin et al. 2019; Husain et al. 2019). Large-molecule drugs such as recombinant proteins and mAbs target many of the cell-surface receptors. These include biologic inhibitors of VEGF and TNF, and recombinant versions of proteins that are ligands for cell-surface receptors such as erythropoietin and granulocyte colony-stimulating factor (G-CSF) (Parveen et al. 2019). Erythropoietin is used for cancer chemotherapy and the treatment of anaemia associated with dialysis undertaken in the treatment of chronic renal failure (Fecková et al. 2019). G-CSF is used for the treatment of neutropenia associated with bone marrow transplantation, cancer chemotherapy and increasing the number of blood haematopoietic stem cells in haematopoietic stem cell transplantation (Zhao et al. 2019). In cases where developing biologic drugs to modulate such cell-surface receptors has not been possible, small-molecule receptor modulators are orally available, less expensive and safer than the corresponding biologics (Parveen et al. 2019). The development of small-molecule drugs that modulate these receptors is challenging; however, some potential drugs have already reached clinical production (Fecková et al. 2019; Zhao et al. 2019). The modulators of cell-surface receptors that interact with proteins or peptides other than chemokine receptors are listed in Table 1.

Table 1 PPI modulators in development targeting cell-surface receptors (Haberman 2012)

Receptor	Compound	Comment
Thrombopoietin receptor	Eltrombopag	Agonist approved for idiopathic thrombocytopenic purpura; in Phase III trials in hepatitis
Thrombopoietin receptor	GSK2285921	Agonist, phase II, Oncology
Erythropoietin receptor	LG5640; lead agonist compound; ligand	Agonist, advanced discovery stage
Granulocyte colony-stimulating factor	Lead optimisation; ligand	Agonist
Lymphocyte function associated antigen/intercellular adhesion molecule-1 PPI	SAR code: SAR1118	Antagonist; phase III, dry eye syndrome. It is a target in inflammatory conditions including dry eye.
Tumour necrosis factor/tumour necrosis factor receptor PPI	E-32712 and other lead macrocyclic compounds	Antagonist

Cytokine receptors

Most small-molecule PPI modulators of cytokine receptors are PPI inhibitors or antagonists. Using the proprietary cellular screening assay, STATs technology, small-molecule cytokine mimetics that are agonists have been developed (Silva et al. 2019). These include SB-247464, a small-molecule non-peptide mimetic of murine G-CSF (Tian et al. 1998). G-CSF (like many other cytokines and growth factors) binds to the extracellular domains of its receptors and dimerises them, resulting in the activation of intracellular signalling through a pathway that leads to the phosphorylation of members of the STAT class of signalling proteins. The phosphorylated STATs form homodimers, which are translocated into the nucleus and bind to a synthetic STAT-responsive promoter that drives expression of a fluorescent reporter gene (Tian et al. 1998; Zhao et al. 2019). Screened small organic compounds selected SB-247464 (shown in Fig. 3a). In support of this model, SB-250017, a related non-symmetrical compound was developed. However, it can bind to the extracellular domain of the receptor but cannot dimerise it. Hence, SB-250017 acts as an antagonist of SB-247464 but has no effect on the activation of the receptor by G-CSF (Haberman 2012).

More so, STATs technology has been used to discover small-molecule non-peptide agonists of the human thrombopoietin (TPO) receptor (Erickson-Miller et al. 2005). Optimisation of their initial compound, SB-394725, led to the development of eltrombopag (SB-497115-GR). The small-molecule TPO agonist eltrombopag (shown in Fig. 3b) provides an alternative to recombinant human TPO, which can elicit anti-endogenous TPO antibodies in some patients, resulting in profound thrombocytopenia (Li et al. 2001). Eltrombopag is the only synthetic small-molecule direct (i.e. not allosteric) PPI modulator to reach the market. It is approved for the treatment of the rare disease, idiopathic thrombocytopenic purpura. Eltrombopag is also being tested in phase III clinical trials to treat low platelet count in patients with liver cirrhosis due to hepatitis C and in phase II trials in oncology. GSK2285921 (follow-on compound to eltrombopag) is being tested in phase II clinical trials in oncology. With respect to oncology, SB-559457 (a related compound to eltrombopag) was found to be specifically toxic to primary human myeloid leukaemia cells in culture (Kalota and Gewirtz 2010).

More recently, small-molecule agonists of the human receptors for erythropoietin (EPO) and G-CSF were discovered using STATs technology (Miller et al. 2015). STATs cellular assay technology is good for discovering such compounds, provided the chemical libraries screened in these assays contain viable hits. Research has discovered a lead series of small-molecule, selective EPO receptor agonists (including LG5640) that display partial efficacy compared with recombinant human EPO in several models of EPO-induced erythropoiesis (Haberman 2012; Miller et al. 2015; Zhao et al. 2019).

Integrins

The integrin superfamily of proteins consists of cell-surface receptors that mediate attachment between cells and either the extracellular matrix or other cells (Takada et al. 2007). Among these proteins is the leukocyte integrin lymphocyte function-associated antigen 1 (LFA-1). LFA-1 is found on such leukocytes as T cells, B cells, macrophages and neutrophils, and is involved in recruitment to sites of infection or inflammation. LFA-1 on T cells binds to the immunoglobulin superfamily member ICAM-1 on, for example, antigen-presenting cells and endothelial cells (Dustin et al. 2004). The LFA-1/ICAM-1 PPI plays an important role in processes such as T cell activation, T cell homing to peripheral lymphoid organs and sites of inflammation (Dustin et al. 2004; Graff et al. 2008). This PPI is thus a target for the discovery of drugs to treat inflammatory conditions. Using a tethering-based FBDD platform, compounds that potently inhibit both human T cell migration and T cell activation by disrupting LFA-1/ICAM-1 PPIs have so far been discovered (Haberman 2012; Zhong et al. 2012).

One of the compounds (SAR1118) showed good pharmacokinetic properties and oral availability in rodents and inhibited neutrophil migration in a murine peritonitis model (Zhong et al. 2012). A phase I clinical trial (conducted in

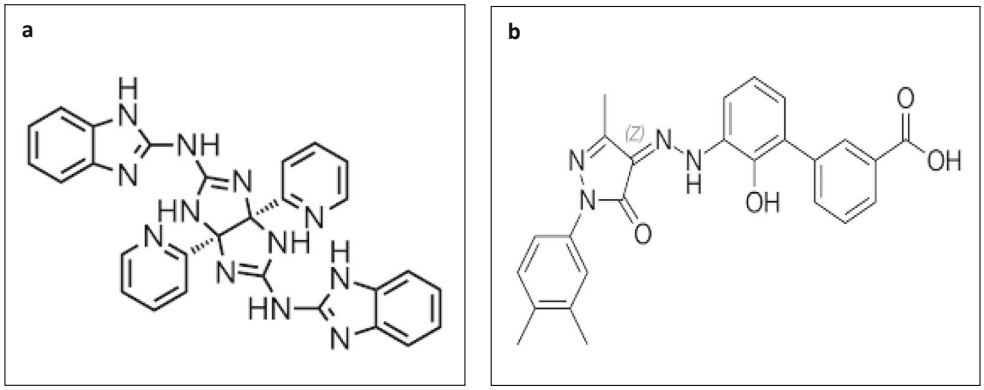


Fig. 3 Chemical structure of a SB-247464—a small-molecule cytokine non-peptide mimetic agonist of murine G-CSF developed using the proprietary cellular screening assay, STATs technology (figure taken from Grosdidier et al. 2009); and b eltrombopag—a synthetic small-molecule TPO agonist approved for the treatment of idiopathic thrombocytopenic purpura (figure taken from Susanto 2015)

2008) and phase II clinical trial (conducted in 2011) of SAR1118 observed that it was safe, well tolerated and demonstrated statistically significant improvements in tear production and visual function. A phase III study of SAR1118 observed positive ophthalmic solution in the treatment of dry eye syndrome. In addition to dry eye, SAR1118 will be tested in a broad range of ocular inflammatory conditions, including diabetic macular oedema (Haberman 2012; Zhong et al. 2012).

Chemokine receptor

Chemokines and their receptors are attractive drug targets because of their role in inflammatory diseases. Chemokines induce chemotaxis. They are members of a family of small cytokines. Chemokine receptors are members of the G protein-coupled receptor (GPCR) superfamily. They bind to natural small-molecule ligands (Guidolin et al. 2019; Husain et al. 2019). Numerous small-molecule drugs that are competitive inhibitors of these ligands already exist. GPCR antagonists represent the largest class of drugs produced (Miszta et al. 2018; Husain et al. 2019). Chemokine receptors bind to chemokines (small proteins) and thus represent a class of PPIs. Discovering small molecules that directly inhibit chemokine receptors at their chemokine binding sites has been an insurmountable task, presenting difficulties in discovering PPI modulators. However, discovering small-molecule chemokine receptor antagonists that act via an allosteric mechanism is easier, because GPCRs exert their signalling activities via complex ligand-mediated conformational changes, which may be a particularly 'allosteric' class of proteins (Miszta et al. 2018). Natural ligands and drugs that bind to 'orthosteric' sites on GPCRs induce unique GPCR conformational states that activate a discrete subset of signalling pathways and cellular behaviours. Since allosteric modulators of GPCRs work by causing conformational changes in the structures of these proteins, some of them may also give rise to functional selectivity in the actions of orthosteric natural ligands that co-bind to the GPCR (Conn et al. 2009; Husain et al. 2019; Miszta et al. 2018).

In the case of GPCRs—including chemokine receptors—whose natural ligands are peptides or proteins, allosteric sites (defined as binding sites for known allosteric modulators of these receptors) are located in distinct sites on the receptor proteins from the orthosteric peptide binding sites (Conn et al. 2009; Miszta et al. 2018). Chemokines bind specifically to orthosteric sites that are located in the extracellular domains of their receptors. Allosteric sites on chemokine receptors are located in transmembrane domains that are distant from the chemokine binding sites (Conn et al. 2009). Small-molecule allosteric modulators that bind to these sites were developed via fairly standard medicinal chemistry and high-throughput screening, augmented with structure-based drug design. Thus, although orthosteric binding sites on chemokine receptors (and on other GPCRs that have peptide ligands) have proven so far to be intractable for the discovery of small-molecule modulators, the discovery of drug-like small-molecule allosteric modulators of chemokine receptors is much more feasible (Conn et al. 2009; Haberman 2012; Miszta et al. 2018; Husain et al. 2019).

However, the discovery of small-molecule chemokine receptor antagonists has been a challenging task. Most of the agents that have been entered into clinical trials have failed, purely because the diseases addressed by these compounds have complex biology and, poor predictive animal models and target redundancy was used (Horuk 2009; Haberman 2012; Miszta et al. 2018). So far, only two small-molecule chemokine antagonists have entered the market. One example is the allosteric CCR5 antagonist maraviroc for the treatment of HIV/AIDS (shown in Fig. 4a). The other is the CXCR4 inhibitor plerixafor (shown in Fig. 4b). Plerixafor is a partial antagonist of the chemokine receptor CXCR4 and an allosteric agonist of CXCR7 (Kalatskaya et al. 2009). It is used together with G-CSF to mobilise haematopoietic stem cells to the peripheral blood for autologous transplantation in

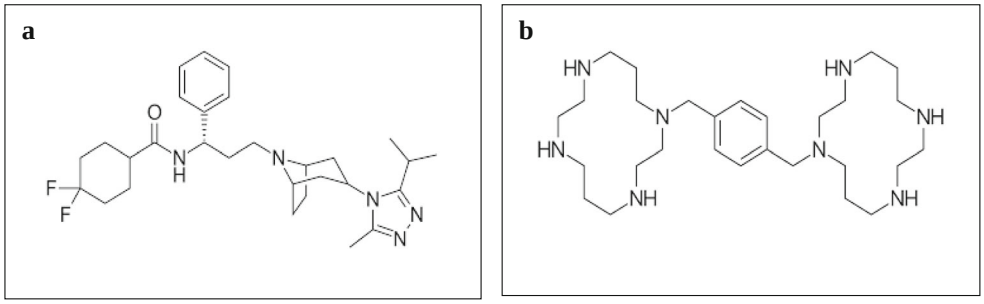


Fig. 4 Chemical structure of a maraviroc—a small-molecule allosteric CCR5 chemokine antagonist used for the treatment of HIV/AIDS (figure taken from Xu et al. 2014); and b plerixafor—a partial antagonist of the chemokine receptor CXCR4 and an allosteric agonist of CXCR7. It is used for autologous transplantation in patients with non-Hodgkin lymphoma and multiple myeloma (figure taken from Venkata Narasimha Rao et al. 2017)

patients with non-Hodgkin lymphoma and multiple myeloma (Davies et al. 2007) (Table 2).

TNF/TNFR PPI

TNF/TNFR PPIs were previously regarded as intractable. However, macrocycles that block the interaction of TNF with its receptor, TNFR, were discovered (Drahl 2009; Parveen et al. 2019). The TNF/TNFR PPIs are involved in numerous inflammatory diseases such as rheumatoid arthritis, and TNF is the target of several large-selling biologic TNF inhibitors. E-32712 is one example of an orally active TNF/TNFR inhibitory macrocycle discovered so far (Drahl 2009; Parveen et al. 2019).

Intracellular signalling pathways

Signal transduction is a process by which extracellular signals mediate changes within a cell via intracellular signalling pathways (Miszta et al. 2018; Husain et al. 2019). It begins when an extracellular signalling molecule activates a cell-surface receptor or an intracellular receptor. Once activated, receptors mediate changes in intracellular target molecules and subsequently initiate cascades of molecular changes through pathways. The end result is a physiological response such as cellular differentiation, cell growth, cell proliferation, secretion of signalling molecules (such as growth factors or cytokines), cellular motility, cellular adhesion or apoptosis. Signal transduction and intracellular signalling pathways are also fundamental in pathophysiological cues and disease progression. They often become dysregulated in metabolic diseases, immune diseases, cancer and many other diseases (Conn et al. 2009; Miszta et al. 2018; Husain et al. 2019).

Studies have targeted tractable signalling receptors such as GPCRs, nuclear receptors, growth factor receptors and cytokine receptors as drug targets (Miszta et al. 2018). The emergence of kinase inhibitors has been another very significant breakthrough. It began with the discovery of imatinib and led to the discovery of other inhibitors aimed at treating different types of cancer (Bhullar et al. 2018; Ferguson and Gray 2018). However, many intracellular signal transduction pathways remain inaccessible because they are driven by key components that have so far been intractable. Most important are the 'undruggable' PPIs, which are key components of all

Table 2 Selected chemokine receptor modulators in development (Haberman 2012)

Compound	Chemokine receptor	Comments
Maraviroc	CCR5	HIV entry inhibitor; for the treatment of HIV infection
Plerixafor	CXCR4 (partial agonist); CXCR7 (allosteric agonist)	Used in combination with G-CSF to mobilise haematopoietic stem cells to the peripheral blood for autologous transplantation in cancer patients; for treatment of HIV infection
Traficet-EN	CCR9	Phase III, Crohn's disease
Reparixin	CXCR1/CXCR2	Phase II, breast cancer
Navarixin	CXCR1/CXCR2	Phase II, chronic obstructive pulmonary disease
SB656933	CXCR2	Phase II, ulcerative colitis
Cenicriviroc	CCR5 and CCR2	Phase IIb, HIV infection
CCX140	CCR2	Phase II, diabetic nephropathy
CCX354	CCR1	Phase II, rheumatoid arthritis
PF-4136309	CCR2	Phase II, hepatitis C infection with abnormal liver enzymes

signalling pathways, for example PPIs between transcription factors and multicomponent protein complexes that are key mediators of intracellular signalling. Small-molecule modulators of PPIs that target signal transduction pathways are yet to be developed and there have been a keen interest in targeting these 'undruggable' targets (Ferguson and Gray 2018; Miszta et al. 2018; Husain et al. 2019).

Tcf/β-catenin transcription factor complex

Small-molecule inhibitors of the oncogenic Tcf/β-catenin transcription factor complex that is central to the Wnt pathway have been discovered (Lepourcelet et al. 2004; Yan et al. 2017; Jeong et al. 2018). The Wnt pathway is dysregulated in several types of cancer, which include hepatocellular carcinoma, multiple myeloma (MM), B cell chronic lymphocytic leukaemia (B-CLL) and colorectal cancer. Due to the central role of the Wnt pathway in these cancers, PPIs that are critically involved in the pathway have been targeted. One example is the 'destruction complex', a multicomponent cytoplasmic protein complex that includes among others the proteins adenomatous polyposis coli (APC) and glycogen synthase kinase 3 (GSK-3). Under normative cellular functions (when the 'destruction complex' is intact), GSK-3 phosphorylates β-catenin, a multifunctional protein that is involved both in signal transduction and in intercellular adhesion. This phosphorylation targets β-catenin for degradation in the cytoplasm (Yan et al. 2017; Jeong et al. 2018).

However, when the 'destruction complex' is disrupted via signalling from Wnt family ligands bound to their cell-surface receptor, β-catenin accumulates in the cytoplasm and moves into the nucleus. There, it binds to transcription factors of the T cell factor (Tcf) family, including Tcf4, the major Tcf expressed in stem cells of the gut and in colorectal cancer. In the absence of β-catenin, Tcf proteins are transcriptional repressors (Lepourcelet et al. 2004; Yan et al. 2017). β-Catenin binding changes Tcf proteins from repressors into transcriptional activators that activate a set of downstream genes, including the oncogene c-Myc and the cell cycle protein cyclin D1. In precancerous colonic adenomas or colorectal cancers, APC is often mutated and no destruction complex forms. This results in constitutive stabilisation of β-catenin, which can freely move into the nucleus and bind to Tcf4. In the case of other cancers caused by dysregulation of the Wnt pathway, β-catenin also becomes stabilised, via other genetic changes that do not involve ACP (Yan et al. 2017; Jeong et al. 2018).

Several small-molecule inhibitors of the human Tcf/β-catenin PPI have been developed through structural and mutagenesis studies of the PPI (which identified a hotspot), followed by assay development and screening of natural product libraries (Lepourcelet et al. 2004). Of the tested compounds, two fungal derivatives, PKF115-584 and CGP049090, gave the best results in all the assays and were tested in preclinical studies. However, the molecular mechanisms by which the compounds act on this PPI remain elusive and are not yet clear. Hence, studies have been focussed on other means to target the Wnt pathway by performing chemical genetic studies to identify novel targets that modulate the pathway (Huang et al. 2009).

BCL6/SMRT PPI in B cell lymphoma

Diffuse large B cell lymphoma (DLBCL) is the most common type of non-Hodgkin's lymphoma and accounts for about 30% of all lymphomas (Friedberg 2011). B cell lymphoma 6 (BCL6) acts as an oncogene in the majority of cases of DLBCLs (Compton and Hiebert 2010). The BTB domains of BCL6 interact with a co-repressor known as SMRT (silencing mediator for retinoid or thyroid-hormone receptors). SMRT in turn facilitates the recruitment of histone deacetylase 3 (HDAC3) to the DNA promoters bound by BCL6. This results in the repression of the genes controlled by the promoters, via removal of acetyl groups from histones of chromatin (Compton and Hiebert 2010; Friedberg 2011).

In normal lymphoid germinal centre B cell development, immunoglobulin genes undergo recombinations and somatic mutations to generate antibody diversity. Despite this genomic instability, germinal centre B cells are able to undergo rapid proliferation because BCL6 represses a set of genes that regulate the DNA damage response and cell cycle checkpoints (Compton and Hiebert 2010). Among these genes are CHK1 (checkpoint kinase 1), cyclin-dependent kinase inhibitor 1 (CDKN1A), ATR (ataxia telangiectasia and Rad3-related protein) and TP53 (which codes for p53). Once B cell clonal diversity has been achieved, BCL6 expression is downregulated. This allows restoration of cell cycle checkpoints, normal DNA damage control and B cell differentiation and maturation. Oncogenic overexpression of BCL6 through chromosomal translocation, gene amplification or promoter mutation results in continued B cell progenitor proliferation and acquisition of additional mutations. This results in an aggressive B cell lymphoma (Compton and Hebert 2010; Friedberg 2011; Haberman 2012).

Cerchietti and co-workers (Cerchietti et al. 2010) discovered small-molecule antagonists of the PPI between BCL6 and SMRT, and demonstrated that these compounds could kill DLBCL cells in vitro. Using computer-aided drug design (CADD) (to first identify putative small-molecule binding sites), screening by virtual docking of the compounds into the putative binding site and selection based on maximising chemical diversity and Lipinski's Rule of Five, a lead compound whose reproducibly inhibited BCL6/SMRT PPI in DLBCL was designed (Cerchietti et al. 2010). Furthermore, they performed structural and mutagenesis studies (which identified a hotspot), performed CADD, and used their models for virtual screening of 1,000,000 commercially available

compounds. Compound selection was based on chemical diversity, drug-likeness, immediate commercial availability and the ability to block BCL6-mediated transcriptional repression in a cellular assay (Haberman 2012).

A lead compound designated '79-6' (PubChem CID5721353, shown in Fig. 5a), which specifically killed BCL6-positive lymphoma cell lines and BCL6-positive tumour cells in xenograft models was identified. However, the five-membered ring of 79-6 contains sulphur, which is prone to oxidation and may consequently result in loss of efficacy, hence the need to optimise 79-6 in order to develop a clinical candidate for BCL6-targeted therapy (Cerchietti et al. 2010). Furthermore, epigenetic regulation was identified as a potentially important area of opportunity for drug discovery (Haberman 2012). So far, two inhibitors of class I HDACs have been approved for the treatment of cutaneous T cell lymphoma—vorinostat (Zolinza) and romidepsin (Istodax). Other agents are being tested in clinical trials for various types of cancer. However, the mechanisms by which these compounds work, and why they appear to be active against certain cancers but not others and not normal cells, remain elusive (Cerchietti et al. 2010; Haberman 2012).

HDAC inhibitors have been studied in clinical trials against DLBCL as single agents, but the results have been disappointing (Compton and Hiebert 2010). BCL6/SMRT PPI antagonist in combination with rituximab (which targets the B cell-specific cell-surface protein CD20) is a potential agent to treat DLBCL (Haberman 2012). Currently, rituximab is used in combination with chemotherapy to treat DLBCL. However, there is still a large proportion of patients with unfavourable prognosis. Hence, a drug that specifically targets the BCL6/SMRT PPI, perhaps used in combination with rituximab, may address the unmet medical needs and reduce the need for toxic chemotherapy in DLBCL patients. The discovery and development of inhibitors of the BCL6/SMRT PPI has gained wider interest in the treatment of DLBCL (Compton and Hiebert 2010; Haberman 2012).

AKAP-protein kinase A interaction

Scaffolding proteins are a tremendously important means by which the cell organises signal transduction pathways. They form PPIs with the signalling proteins (Conn et al. 2009). Thus, discovering small-molecule inhibitors of these PPIs is a potential strategy for targeting a wide array of signalling pathways (Miszta et al. 2018; Husain et al. 2019). A-kinase anchoring proteins (AKAPs) (also known as cAMP-dependent protein kinases) are scaffolding proteins that tether protein kinase A (PKA) and other signalling proteins to specific intracellular sites. They are a family of serine/threonine kinases whose activity is dependent on cellular levels of cyclic AMP (cAMP). Thus, the tethering of PKA via PPIs with an AKAP results in the compartmentalisation of cAMP signalling within the cell (Christian et al. 2011).

Small-molecule inhibitors of the AKAP/PKA interaction have been developed for the potential treatment of chronic heart failure. AKAP18δ is an AKAP isoform that serves as a scaffold for organising the adrenaline-beta-adrenoreceptor-cAMP-PKA signalling pathway in cardiac muscle (Lygren and Taskén 2008). Targeting AKAP18δ and its PPIs with PKA is a potential approach to modulate the cardiac myocyte cAMP-PKA system (Christian et al. 2011). Hence, research efforts have been centred on developing a screening assay for the disruption of the AKAP18δ/PKA PPIs by screening a library of over 20,000 'drug-like' compounds. This led to the identification of 3′,3-diamino-4,4′-dihydroxydiphenylmethane, designated FMP-API-1 (shown in Fig. 5b), which disrupt the AKAP18δ/PKA PPIs via an allosteric mechanism with micromolar dissociation constant (Christian et al. 2011). In the meantime, FMP-API-1 might be used as a tool compound to investigate the function of AKAP-PKA PPIs in cells and in animal models. However, higher affinity drug-like small-molecule AKAP/PKA PPI antagonists are still sought. Thus, scaffolding proteins provide numerous opportunities for targeting by small-molecule PPI modulators (Miszta et al. 2018; Husain et al. 2019).

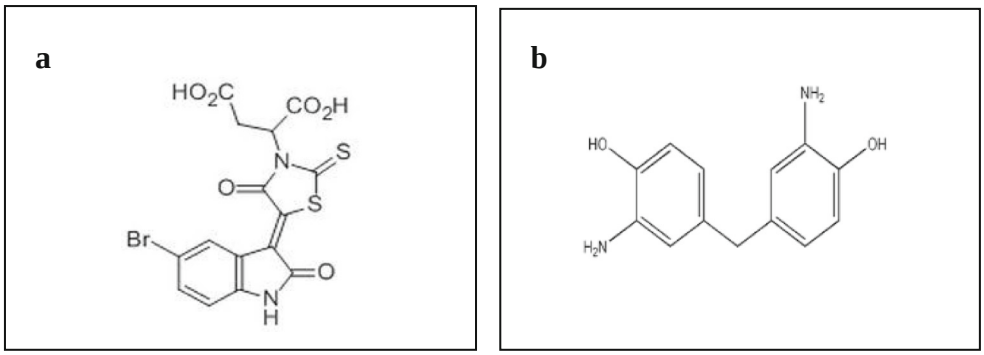


Fig. 5 Chemical structure of a '79-6' (PubChem CID5721353)—a target-specific lead compound known to kill BCL6-positive lymphoma cell lines and BCL6-positive tumour cells in xenograft models (figure taken from Yasui et al. 2017) and b FMP-API-1—a 3′,3-diamino-4,4′-dihydroxydiphenylmethane 'drug-like' compound identified to disrupt the AKAP18δ/PKA PPIs via an allosteric mechanism with a micromolar dissociation constant (figure taken from Christian et al. 2011)

The ubiquitin system

The ubiquitin (Ub) system is a vital regulatory system based on covalently linking the small (8.5 kDa) regulatory protein, Ub, to numerous specific protein targets in all eukaryotic cells (Cohen and Tcherpakov 2010; Varshavsky 2017; Wertz and Wang 2019). This system is involved in functions such as regulation of intracellular protein turnover, mitosis, innate immunity and certain protein kinases and other enzymes (Cohen and Tcherpakov 2010). The ubiquitinylation cascade is a potential target for the development of specific smart drugs. In protein degradation the Ub system works together with the proteasome in a pathway known as the Ub proteasome system (UPS) where Ub is used to tag proteins for degradation by the proteasome (Cohen and Tcherpakov 2010; Varshavsky 2017). The ubiquitinylation pathway is complex and involves several levels of mediators, which include Ub activators (E1), Ub-conjugating enzymes (E2) and Ub ligases (E3) (as shown in Fig. 6). In this pathway, Ub moves from E1s to E2s. E3s interact with ubiquitinylated E2s and substrate proteins via PPIs where Ub is transferred from E2 to the substrate. This Ub moiety is recycled when the cycle is complete, resulting in tagging of substrates with polyubiquitin chains. In humans, there are about 10 E1s, 40 E2s and over 600 E3s (Varshavsky 2017; Wertz and Wang 2019).

The Ub system is a virtually untapped area of opportunity for drug discovery and development. Drugs that target the UPS and the ubiquitinylation pathway itself are in clinical development (Wertz and Wang 2019). Examples include bortezomib, MLN4924 (which inhibits a pathway that activates one class of E3s), the Cullin RING E3 ligases (CRLs) and CC0651 (an allosteric modulator that inhibits the ubiquitinylation activity of an E2 that interacts with CRLs) (Appel 2011; Varshavsky 2017; Wertz and Wang 2019). Bortezomib targets protease activity of the proteasome. It blocks proteasomal disposal of all ubiquitinylated proteins in the cell. However, bortezomib remains a useful drug in MM, where blocking of the proteasome results in an overload of damaged proteins, which subsequently destroys the cell via apoptosis (Appel 2011). Based on bortezomib mechanism of action, it is highly nonspecific and has severe adverse effects. As a result, developing second-generation proteasome inhibitors, which include MLN9708 and MLN4924 became a necessity. MLN9708 is an oral drug, more specific and with fewer side effects than bortezomib. MLN9708 is being tested in phase I and phase II trials in MM patients (Appel 2011).

MLN4924 is another drug in clinical trials that targets an arm of the ubiquitinylation cascade itself (Deshaies 2009; Soucy et al. 2009). It is an AMP analog that inhibits NEDD8-activating enzyme (NAE). NEDD8 is a Ub-like protein and NAE is an E1 for neddylation. Inhibition of neddylation by MLN4924 at the E1 (NEDD8-activating) step thus indirectly inhibits one class of E3s, the CRLs, without the need to inhibit the PPIs between members of this class of E3s and the substrate proteins that bind to them, or the need to inhibit the catalytic activity of E2s or the PPIs between E2s and the CRL E3s. MLN4924 is thus much more specific than bortezomib in inhibiting the Ub system. However, it is not very specific with respect to targeting proteins for degradation. MLN4924 has potentially important anti-tumour effects and induces apoptosis. MLN4924 is now being tested in phase I clinical trials in patients with solid and haematological tumours (Deshaies 2009; Soucy et al. 2009).

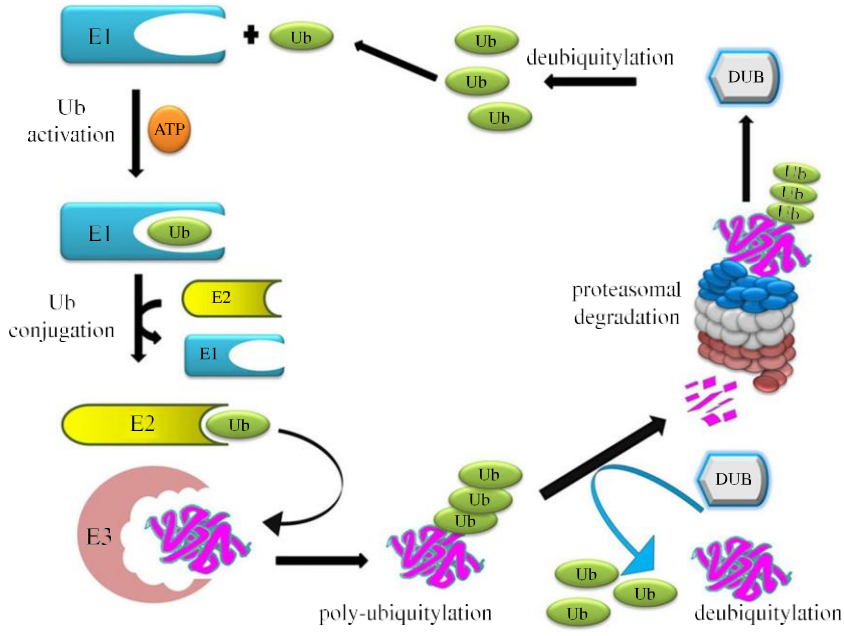


Fig. 6 A schematic illustration of the Ubiquitin-proteasome pathway (UPP). Ub is conjugated to proteins that are destined for degradation by an ATP-dependent process that involves three enzymes. A chain of five Ub molecules attached to the protein substrate is sufficient for the complex to be recognised by the 26S proteasome. In addition to ATP-dependent reactions, Ub is removed and the protein is linearised and fed into the central core of the proteasome, where it is digested to peptides. The peptides are degraded to amino acids by peptidases in the cytoplasm or used in antigen presentation (figure taken from Haq and Ramakrishna, 2017)

Another strategy to inhibit CRLs indirectly, this time at the E2 level, was developed (Ceccarelli et al. 2011). This led to the discovery of CC0651, an allosteric inhibitor of the E2 Ub-conjugating enzyme Cdc34. CC0651 analogs inhibited proliferation of human cancer cell lines. They also caused accumulation of the CDK inhibitor p27Kip, which led to uncontrolled DNA synthesis in the S-phase of the cell cycle, leading to DNA damage and induction of apoptosis. The E2 inhibitor CC0651 and its analogs appear to be more specific than MLN4924, which inhibits the activity of all CRLs. However, the target of CC0651 still ubiquitinylates hundreds of substrate proteins, so it is relatively nonspecific (Ceccarelli et al. 2011; Haberman 2012).

The best agents that target the Ub system are the PPI modulators that target E3s, which interact with their substrates via PPIs (Lecker et al. 2006; Haberman 2012). However, because of the intractability of PPIs, the development of specific agents that target the Ub system has been a major bottleneck. Research efforts have led to the development of an agent that targets one E3, the human homolog of mouse double minute 2 (MDM2) protein. HDM2 interacts with p53 via a PPI (as shown in Fig. 7) (Moll and Petrenko 2003). The p53 is often referred to as the 'guardian of the genome', controls pathways that respond to DNA damage or other stress signals by blocking cell proliferation by inducing either DNA repair or apoptosis. The p53 is mutated or inactivated in nearly all human cancers, which allows the uncontrolled proliferation of cancer cells, rendering them resistant to cytotoxic chemotherapy (Moll and Petrenko 2003). In approximately half of human cancers, p53 is inactivated via mutation. In the other half, p53 remains unmutated but is inactivated. The main means of inactivation is via HDM2, which is overexpressed in the majority of cancers with wild-type p53. HDM2 regulates p53 in three ways: inhibition of p53-induced transcription, promotion of export of p53 out of the nucleus and inducing p53 degradation by the proteasome (Moll and Petrenko 2003). The last two activities mentioned both involve HDM2's E3 Ub ligase activity, thus presenting a potential target for PPI-focused drug development (Lecker et al. 2006; Haberman 2012; Varshavsky 2017; Wertz and Wang 2019).

Currently, there are two leading drug candidates that specifically disrupt the HDM2/p53 PPI (Shangary and Wang 2009). The most advanced compound, now being tested in phase I clinical trials, is RG7112, is an analog of nutlin-3a (Vassilev et al. 2004; Shangary and Wang 2009; Cheok et al. 2011). Nutlin-3a is the active enantiomer isolated from racemic nutlin-3, which have been studied in various cell culture and preclinical animal models, as a monotherapy and in combination therapies (Shangary and Wang 2009; Cheok et al. 2011; Crane et al. 2015). These studies showed that nutlin-3 potently induced apoptosis in cell lines derived from such haematologic cancers as acute myeloid leukaemia, acute lymphoblastoid leukaemia, MM and B-CLL. These haematologic tumours and other HDM2/p53 PPI-disrupting agents are potential targets for treatment with nutlin-3, since they exhibit a high percentage of unmutated TP53 at diagnosis. These studies concur with the development of nutlins in cancer, either as single agents or in combination therapies. They led to the entry of the nutlin-3 analog RG7112 into phase I clinical trials in haematologic malignancies and advanced solid tumours (Haberman 2012).

The other compound is MI-219, analogs of which are currently being tested in advanced preclinical studies (Shangary et al. 2008; Shangary and Wang 2009). MI-219 binds to HDM2 with an inhibition constant (K_i) of 5 nM. It is designed to mimic not only phenylalanine 19, tryptophan 23 and leucine 26 in the p53 binding site for HDM2 but also a fourth residue, leucine 22, which appears to play an important role in the HDM2/p53 PPI as well (Shangary and Wang 2009). MI-219 has good pharmacological properties, a high degree of

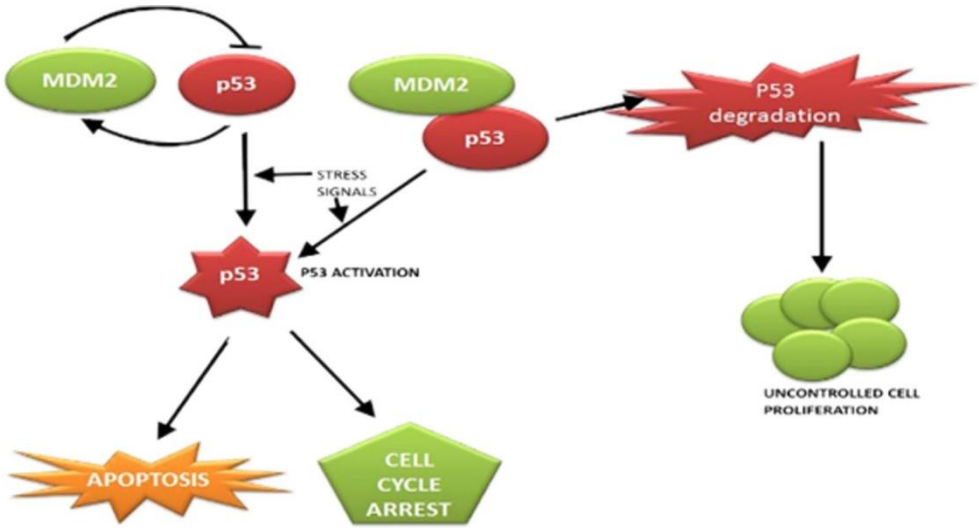


Fig. 7 A schematic diagram showing HDM2/p53 PPI. MDM2 and p53 form an auto-regulatory feedback channel. p53 stimulates the expression of MDM2; MDM2 in turn inhibits p53 activity because it stimulates its degradation in the nucleus and the cytoplasm, blocks its transcriptional activity and promotes its nuclear export. A broad range of DNA-damaging agents or deregulated oncogenes induces p53 activation (figure taken from Carry and Garcia-Echeverria 2013)

specificity for MDM2 and induced accumulation of p53, as well as inhibiting the growth of cancer cell lines with wild-type p53 with submicromolar IC50 values (Shangary and Wang 2009; Cheok et al. 2011). MI-219 analogs have also been tested and have demonstrated activity in combination therapies with etoposide, doxorubicin and cisplatin in mouse models of lung cancer, rhabdomyosarcoma and pancreatic cancer. These agents were found to have minimal toxic effects in normal cells. Given the large number and specificity of E3 Ub ligases and their important role in intracellular pathways, there is a large field of possibility for discovery of novel PPI modulators that target these biomolecules and their interactions with their substrates (Shangary and Wang 2009; Cheok et al. 2011).

JNJ-26854165, a novel tryptamine derivative is another compound in phase I clinical trials in advanced or refractory solid tumours. JNJ-26854165 is thought to be a possible HDM2/p53 inhibitor; however, it appears to work via a different mechanism of action that does not involve HDM2 (Kojima et al. 2010). It appears to work via accelerating the proteasomal degradation of p21 and to antagonise the p53-mediated transcriptional induction of p21. JNJ-26854165 also induces apoptosis in tumour cells with mutant p53 via inducing delay in the S-phase of mitosis and upregulation of expression of the transcription factor E2F1 (a key mediator of an important pathway that controls cellular proliferation). This results in apoptosis, preferentially of S-phase cells. JNJ-26854165 has similar effects on tumours to HDM2/p53 PPIs like the nutlins and MI-219 (Cheok et al. 2011; Haberman 2012).

Apoptosis regulators

Apoptosis is an ATP-dependent pathway of programmed cell death in all-multicellular animals (Danial and Korsmeyer 2004; Wyllie 2010). It is essential for normal embryonic development and for maintaining normal cellular homeostasis in adults, as well as for response to infectious agents. Apoptosis is dysregulated in several major diseases. Cancer is the major focus of studies seeking to develop drugs that modulate apoptotic pathways, since apoptosis is blocked in perhaps all cancers. This is a significant target in uncontrolled cellular proliferation in cancer (Danial and Korsmeyer 2004; Wyllie 2010). An illustration of apoptosis in ovarian cancer and some of the targeted therapeutic approaches is shown in Fig. 8.

Central to apoptotic pathways are two families of proteins, the caspases and the B cell lymphoma-2 (Bcl-2) family (Danial and Korsmeyer 2004; Wyllie 2010). Caspases are a class of serine proteases that function in apoptosis. They form a cascade that ultimately results in cell death. Bcl-2 family proteins control this process, either halting the processes that result in apoptotic cell death or allowing these processes to go forward. However, the central pathways of apoptosis are controlled by a complex system of pro-apoptotic and anti-apoptotic Bcl-2 family members, which act to ensure that apoptosis is only triggered when it is appropriate. Bcl-2 family member interactions that control apoptosis are PPIs. Thus, it has been difficult to discover agents that affect the central pathways of apoptosis and that are capable of being taken into the clinic (Danial and Korsmeyer 2004; Wyllie 2010; Haberman 2012).

The central pathways of apoptosis include an intrinsic and an extrinsic pathway (Ubanako et al. 2015). The intrinsic pathway is triggered by cellular stress (which includes among others the deprivation of growth factors needed for survival, drug treatments or ionising radiation), p53-mediated apoptotic signals triggered by DNA damage, virus infection, hypoxia and energy deprivation. The intrinsic pathway is also modulated by other signal transduction pathways, such as 'oncogene overdrive' in which the Myc oncogene may trigger apoptosis instead of hyperproliferation and malignant transformation, the Akt/PTEN pathway (which when dysregulated is a factor in several types of cancer) and the UPS pathway for degradation of unwanted or defective intracellular proteins (Boone et al. 2011; Ubanako et al. 2015). Triggers of the intrinsic pathway operate mainly via modulating members of the Bcl-2 family. Bcl-2 family proteins possess homologous domains that can enter into PPIs among the family members. Anti-apoptotic members of the Bcl-2 family, such as Bcl-2, Bcl-xL and Mcl-1, possess four conserved domains called BH1, BH2, BH3 and BH4. BH 1, 2 and 4 define a hydrophobic groove within the molecule, and BH3 is an 8-to-12 amino acid domain that binds within that groove. These anti-apoptotic proteins localise to the mitochondria, where they specifically bind and sequester pro-apoptotic multi-domain Bcl-2 family members, such as Bak and Bax (Danial and Korsmeyer 2004; Wyllie 2010; Haberman 2012; Ubanako et al. 2015).

The extrinsic apoptotic pathway is triggered by a class of cell-surface receptors of the TNFR family and their corresponding TNF family ligands (Danial and Korsmeyer 2004; Wyllie 2010; Ubanako et al. 2015). These receptor/ligand pairs include TNFR/TNF-α (tumour necrosis factor-alpha), Fas/FasL (Fas ligand) and TRAIL receptor/TRAIL (TRAIL = TNF-related apoptosis-inducing ligand). TRAIL is of great interest to cancer biologists and oncology drug developers, since it has been found to induce apoptosis selectively in cancer cells, independent of p53, which is usually inactivated in human cancers. However, the physiological role of TRAIL is not well understood (Zaba et al. 2010; Haberman 2012).

The binding of a specific ligand to TNFR family receptors on the cell surface leads to clustering of the receptors. The intracellular domains of the receptor complexes then bind to death adaptor proteins, such as Fas-associated death domain protein (Danial and Korsmeyer 2004; Ubanako et al. 2015). These in turn bind to an initiator caspase, caspase 8, which is

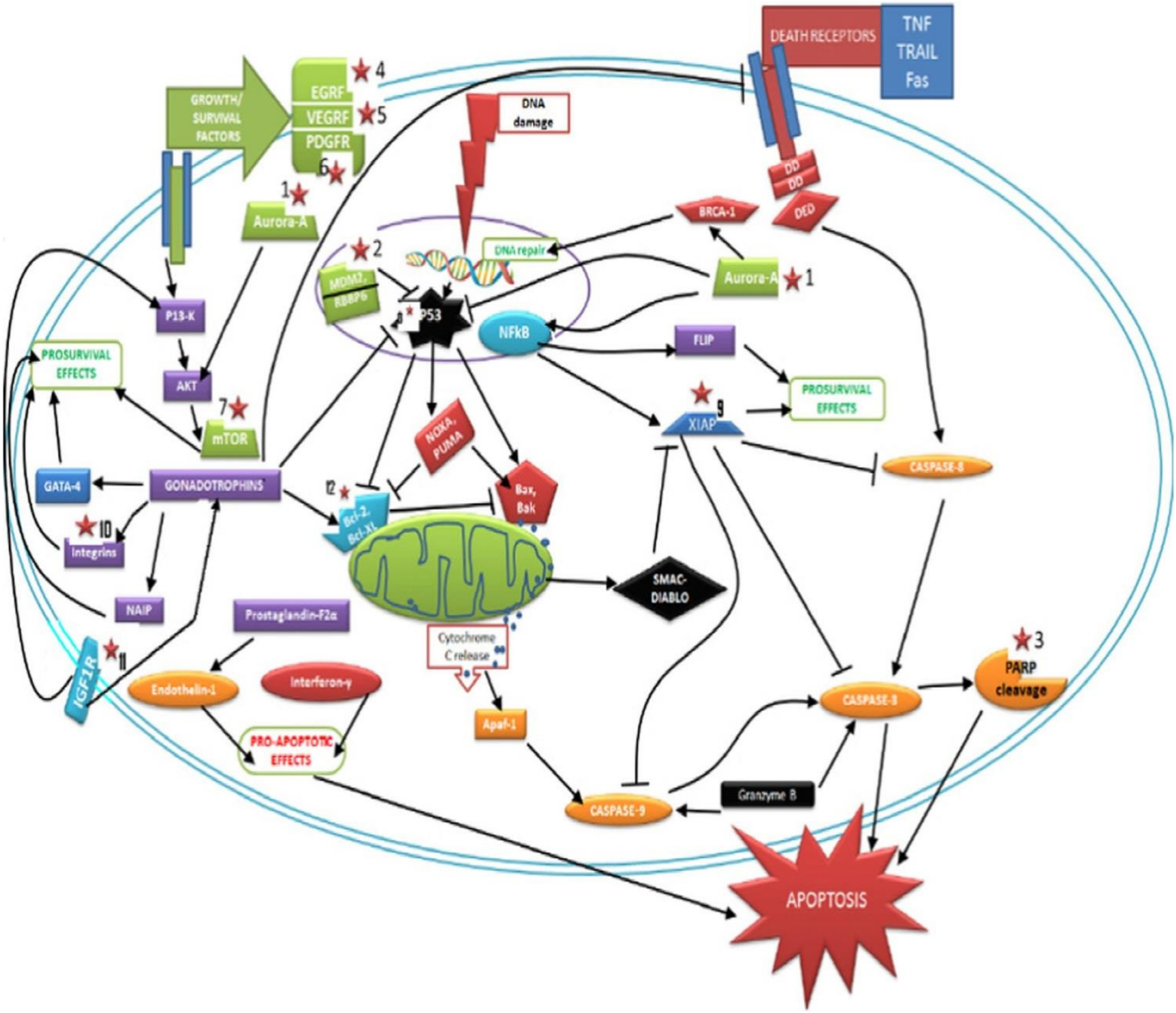


Fig. 8 A schematic mechanism of apoptosis in ovarian cancer and some current targeted molecular therapeutic approaches. Two main pathways of apoptosis have been elucidated: the death receptor (extrinsic) pathway and the mitochondrial (intrinsic) pathway. Targeted molecular therapeutic approaches include angiogenesis inhibitors, inhibitors of the epidermal growth factor receptor (EGFR), aurora kinase inhibitors, poly ADP Ribose Polymerase (PARP) inhibitors, platelet-derived growth factor (PDGF) receptor inhibitors, MTOR inhibitors, targeting Bcl-2 family in ovarian cancer and apoptosis, minimising expression of inhibitors of apoptosis (IAP) as target for ovarian cancer, therapeutic potential of TNF family members, wild-type p53: the genomic guardian target, interferons (IFN), integrins and insulin-like growth factor (IGF) (figure taken from Ubanako et al. 2015)

autocatalytically activated. Caspase 8 activates effector caspases (for example, caspase 3 and 7) leading to apoptosis. Caspase 8 also initiates the intrinsic programme of apoptosis via cleavage of the inactive p22 form of Bid, resulting in the formation of an active form of Bid. This in turn triggers the intrinsic pathway of apoptosis via mitochondria membrane pore formation by complexing of pro-apoptotic Bcl-2 family proteins (Danial and Korsmeyer 2004; Wyllie 2010; Ubanako et al. 2015).

The intention of cancer research is to develop PPI-focused small-molecule drugs specifically targeting the relevant PPIs of Bcl-2 family proteins, since all the interactions between the Bcl-2 family members are PPIs (Adams and Cory 2007). Currently, three Bcl-2 familyPPI-disrupting agents (BH3 mimetics) are being tested in clinical trials, namely navitoclax, obatoclax and ABT-199 (Venetoclax). Navitoclax (ABT-263) (shown in Fig. 9a) is the result of the fragment-based drug discovery methodology known as SAR by NMR, which led to a Bcl-2 inhibitor designated as ABT-737 (shown in Fig. 9b). ABT-737 might be useful for the treatment of lymphoma and small cell lung cancer (SCLC) as a monotherapy. It can also be used for a wide variety of cancers in combi- nation with cytotoxic agents or radiation. However, it has

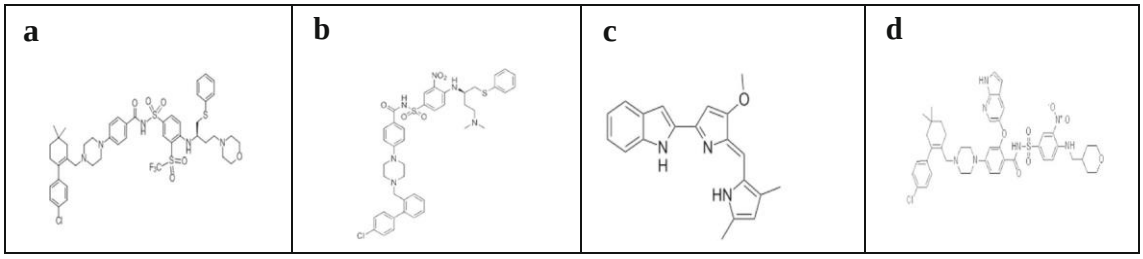


Fig. 9 Chemical structure of a ABT-263 (navitoclax)—an orally active anticancer drug which does not have off-target effects (figure taken from Tse et al. 2008). b ABT-737—an orally bioavailable, selective small molecule B cell lymphoma 2 (Bcl-2) homology 3 (BH3) mimetic, with potential pro-apoptotic and antineoplastic activities (figure taken from Tse et al. 2008). c Obatoclax mesylate (GX15-070)—an experimental drug for the treatment of various types of cancer which include leukaemia, myelofibrosis, Hodgkin's lymphoma and mantle-cell lymphoma among others (figure taken from Goard and Schimmer 2013). d ABT-199 (venetoclax)—a BH3-mimetic Bcl-2 inhibitor, does not cause Ca^{2+}-signalling dysregulation or toxicity in pancreatic acinar cells (figure taken from Souers et al. 2013)

poor physiochemical and pharmaceutical properties (Keller et al. 2006; Tse et al. 2008).

ABT-737 is deficient in oral bioavailability properties and its poor solubility makes formulation for intravenous delivery exigent. Hence, these issue rendered ABT-737 a poor prospect for clinical development (Tse et al. 2008). However, ABT-737 was later optimised to produce a second-generation compound, navitoclax, which has improved physicochemical and pharmacological properties and is orally available. Navitoclax is being tested in phase I and phase II clinical trials in various cancers, including combination therapies with drugs such as Rituxan (rituximab) and Tarceva (erlotinib), as well as with cytotoxic chemotherapies and as a single agent (Roberts et al. 2012).

Obatoclax (GX15-070) (shown in Fig. 9c) is a pan-Bcl-2 inhibitor that inhibits Mcl-1. Mcl-1 is overexpressed in several types of cancer, which include prostate cancers (Dash et al. 2010). The interest of cancer research is in the development of small-molecule PPI inhibitor drugs that target Mcl-1. One such drug, obatoclax, was discovered by using a high-throughput protein-protein interaction assay to screen natural product libraries (Shore and Viallet 2005; Nguyen et al. 2007). Studies indicated that obatoclax inhibited the Mcl-1 apoptotic pathway by disrupting the Mcl-1/BAK and the Mcl-1/BAX PPIs. It overcame apoptosis Mcl-1-mediated resistance in lymphoma cells and cultured melanoma cells (Nguyen et al. 2007). However, because of its poor solubility, obatoclax cannot be an oral drug. Nevertheless, intravenous administration of obatoclax showed single-agent anti-tumour activity in mouse xenograft models bearing several different types of human carcinomas. This suggests that obatoclax is an important Mcl-1-inhibitor in cases where Mcl-1 is involved in blocking apoptosis (Nguyen et al. 2007; Haberman 2012).

ABT-199 (Venetoclax) (shown in Fig. 9d) is a BH3 mimetic that selectively inhibits Bcl-2 (Jakubowska et al. 2018). It was initially used for the treatment of relapsed chronic lymphocytic leukaemia. Early generations of Bcl-2 inhibitors induced sustained Ca2+ retaliations in pancreatic acinar cells (PACs), inducing cell death. Therefore, BH3 mimetics are assumably toxic to the pancreas when used to treat cancer. Although ABT-199 was shown to kill Bcl-2-dependent cancer cells without affecting intracellular Ca2+ signalling, its effects on PACs remain elusive (Jakubowska et al. 2018). Hence, it is of paramount importance to assess whether this recently approved anti-leukaemic drug might potentially have pancreatotoxic effects. Inhibition of Bcl-2 via ABT-199 did not elicit intracellular Ca^{2+} signalling on its own. It did not potentiate Ca^{2+} signalling induced by physiological or pathophysiological stimuli in PACs, although ABT-199 did not affect cell death in PACs under conditions that killed ABT-199-sensitive cancer cells. Cytosolic Ca^{2+} extrusion was slightly enhanced in the presence of ABT-199. In contrast, inhibition of Bcl-xL potentiated pathophysiological Ca^{2+} responses in PACs without exacerbating cell death (Jakubowska et al. 2018).

Stapled peptides

In parallel with second-generation technologies and the discovery and development of small-molecule PPI modulators, peptides have been designed that mimic the amino acid sequence and secondary structures of protein domains that are involved in PPIs. This has led to the emergence of 'stapled peptides' comprising a small loop rigidifying the peptide conformation (Ali et al. 2019; Verhoork et al. 2019). These mimetics are resistant to degradation by proteolytic enzymes have favourable pharmacological properties and are able to penetrate cells (Schafmeister et al. 2000; Walensky et al. 2010; Kim et al. 2011). The initial application of stapled peptide technology was to the Bcl-2 family PPI system, which controls apoptosis (Verhoork et al. 2019). As discussed earlier, Bid is one of the BH3-only pro-apoptotic members of the Bcl-2 family. It appears to exert its pro-apoptotic activity by binding to the hydrophobic groove of anti-apoptotic BH-1/2/3/4 proteins such as Bcl-2 and Bcl-xL. This releases pro-apoptotic BH-1/2/3 proteins from their complexes with the anti-

apoptotic proteins, enabling them to initiate the intrinsic pathway of apoptosis. Beginning in the mid-1990s, studies to develop a Bid mimetic peptide have been ongoing until now (Kim et al. 2011; Ali et al. 2019; Verhoork et al. 2019).

The BH3 domain of Bid that triggers apoptosis possesses an α-helical structure, which was lost upon synthesis of a peptide with the amino acid sequence of the Bid BH3 domain. Additionally, this protein was ineffective in disrupting PPIs between Bcl-2 or Bcl-xL and BH-1/2/3 pro-apoptotic proteins, and was subject to degradation by serum and cellular proteases (Kim et al. 2011; Ali et al. 2019). Moreover, such peptides could not penetrate cells; hence, the stapled peptide technology was applied to produce a stable α-helical form of the peptide. To construct stapled peptides, synthesised peptides with the amino acid sequence of the domain of interest (in this case, the Bid BH3 domain) incorporate two appropriately spaced non-natural amino acids bearing olefin side chains (Walensky et al. 2004; Kim et al. 2011; Ali et al. 2019; Verhoork et al. 2019). The peptides were designed using the three-dimensional structures of the target protein and the interacting partner domain to provide the sequence of the stapled peptide. Using molecular visualisation software, they analysed the binding interface between these two biomolecules and confirmed that the partner-binding domain was α-helical. Thereafter, they selected amino acid sequences that were not directly involved in target recognition as candidates for substitution with non-natural amino acids. The synthesised peptides each contain two olefin-bearing amino acid residues 3, 4 or 7 amino acid residues apart. These modified peptides were designed to project the reactive olefin residues on the same face of the α-helix (Haberman 2012; Ali et al. 2019; Verhoork et al. 2019).

The bridge between the two non-natural amino acids (staple) was formed by using ruthenium-mediated ring-closing olefin metathesis (shown in Fig. 10). The 'staple' in a stapled peptide forms a macrocyclic ring with the amino acid residues it encompasses. The range of bridge lengths of 3, 4 or 7 amino acid residues is used to identify the optimal stapled peptide with respect to affinity for its target and other properties relevant to serving as a drug candidate (Kim et al. 2011; Haberman 2012).

Peptide stapling results in stabilisation of the α-helical form of peptides such as the Bid BH3 domain and other peptides derived from domains that are α-helical in their natural proteins. Such peptides are resistant to proteases and can also be cell-penetrant (Kim et al. 2011; Haberman 2012). Hydrocarbon staples promote effective cellular uptake via endocytic vesicle trafficking. Cellular penetration by a stapled peptide is also dependent on the net charge of the molecule; those with net positive charges show better cellular uptake (Kim et al. 2011; Haberman 2012).

In the original application of the stapled peptide technology, the constructed stapled peptide Bid BH3 mimetics was called 'stabilized alpha-helix of Bcl-2 domains' (SAHBs: Walensky et al. 2004). SAHBs were α-helical, protease-resistant and penetrated cells. One example of the SAHBs that were studied further is SAHBA, which was found to cause the same structural changes in Bcl-xL as did Bid. Intravenous administration of SAHBA consistently caused tumour regression in mouse xenograft models of human leukaemia, as well as improved survival (median survival 5 days for control animals, and 11 days for SAHBA-treated animals). SAHBA showed no overt toxicity in normal tissues. The studies designed to create a Bid mimetic using stapled peptide technology provided proof of principle for the technology and served as a starting point to build a pipeline of preclinical candidates based on stapled peptides (Walensky et al. 2004; Kim et al. 2011; Haberman 2012).

The notch pathway

The notch pathway regulates various aspects of cell proliferation, cellular differentiation, cell-cell communication and cellular survival or death. It is essential for the development of the nervous and haematopoietic systems (Huang et al. 2019). Deregulation of the notch pathway is involved in various cancers, including lung cancer, pancreatic cancer, cancer of the ovary and cancer of immature T cells (T cell acute lymphoblastic leukaemia (T-ALL)). Notch is a cell-membrane receptor just like its ligands (members of the Delta, Serrate and Lag-2 family). The notch pathway becomes activated when the extracellular domain of notch binds to a notch ligand on the

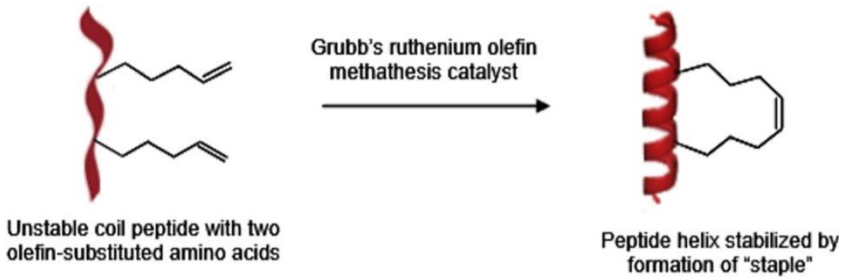


Fig. 10 Construction of a stapled peptide. A 'staple' is formed between two non-natural amino acids by using Grubb's ruthenium-mediated ring-closing olefin metathesis. The staple forms a macrocyclic ring with the amino acid residues (figure taken from Haberman 2012)

surface of an adjacent cell, leading to sequential proteolytic cleavage of the notch intracellular domain, first by an A disintegrin and metalloproteinase (ADAM) family metalloprotease such as TNF-α-converting enzyme (also known as ADAM17) and then by a γ-secretase complex (Moellering et al. 2009; Haberman 2012; Huang et al. 2019).

The free intracellular domain of notch (also known as the intracellular domain of NOTCH1 (ICN1)) translocates to the nucleus and docks with the DNA-bound transcription factor CSL (CBF1, Suppressor of Hairless and Lag-1) (Moellering et al. 2009). The CSL/ICN1 interaction creates a long shallow groove along the interface of the two proteins, which serves as a docking site for coactivator proteins of the mastermind-like (MAML) family, such as MAML1. The resulting trimolecular complex initiates specific transcription of notch-dependent target genes, which uses a stapled peptide to target the notch signal transcription pathway (Moellering et al. 2009; Haberman 2012).

As with other key signal transduction pathways, the notch pathway is centred on PPIs, especially the crucial ternary transcription factor complex CSL/ICN1/MAML1. The discovery of small-molecule drugs that modulate this pathway had thus been considered unfeasible (Moellering et al. 2009; Haberman 2012) until the application of stapled peptide technology to discover agents that modulate this pathway, in particular the key transcription factor complex (Huang et al. 2019). Previous research found that a dominant-negative fragment of MAML1, designated as dnMAML1 (consisting of amino acid residues 13-74) antagonised notch signalling and cell proliferation when expressed in T-ALL cells (Moellering et al. 2009). X-ray diffraction studies showed that the dnMAML1 polypeptide formed an α-helix, which docks with the elongated groove formed by the ICN1/CSL complex. This suggested that it might be possible to design a stapled peptide based on a portion of the sequence of dnMAML1. Such a stapled peptide might inhibit binding of MAML1 to the ICN1/CSL complex, thus blocking transcription of downstream notch-dependent target genes (Moellering et al. 2009; Haberman 2012; Huang et al. 2019).

Moellering and colleagues therefore synthesised a set of six short candidate peptides, which together encompassed the entire contact surface of dnMAML1 with the ICN1/CSL complex (Moellering et al. 2009). Functional studies with these peptides led them to select the 15-amino-acid stapled peptide SAHM1, which showed that stapling conferred a marked helical character (94% helical) to SAMH1, compared with its non-stapled counterpart. SAMH1 was found to be cell-penetrant. Biochemical studies showed that SAMH1 bound to the ICN1/CSL complex competitively with MAML1. Cell culture studies in the KOPT-K1 human T-ALL cell line with a notch-regulated fluorescent reporter gene showed that SAMH1 specifically repressed transcription of notch target genes, in a dose-dependent manner. Gene expression analysis in KOPT-K1 and HPB-ALL human T cell leukaemia cells showed that SAMH1 specifically repressed transcription of notch pathway downstream target genes (Moellering et al. 2009; Haberman 2012; Huang et al. 2019).

More so, SAMH1 markedly reduced proliferation of several T-ALL cell lines in vitro, but was ineffective against leukaemia cell lines that did not depend on the notch pathway for cell proliferation (Al-Shehabi et al. 2019; Majer et al. 2019). In sensitive T-ALL cell lines, SAMH1 treatment also triggered apoptosis. In a mouse model of T-ALL, twice-daily intraperitoneal injection of SAMH1 resulted in significant dose-dependent regression of leukaemia, compared with vehicle-treated mice. Studies of mononuclear cells from SAMH1- and vehicle-treated mice confirmed that SAMH1 treatment resulted in a significant decrease in notch target gene transcription (Al-Shehabi et al. 2019; Majer et al. 2019). These studies showed that SAMH1 is a direct transcriptional antagonist of the notch pathway. This suggests that SAHM1 can be used to determine the role of this pathway in normal physiology and development, and in disease processes. SAMH1 also provides a starting point for drug development, especially in notch-driven cancers such as T-ALL (Haberman 2012; Al-Shehabi et al. 2019; Majer et al. 2019).

Conclusions and future perspectives

The discovery and development of protein-protein interaction modulators has been difficult, because of the structure of protein-interacting interfaces, the lack of natural ligands for PPIs to serve as starting points for drug design and the unsuitability of proprietary chemical libraries for use in HTS campaigns to identify compounds that modulate PPIs (Jana et al. 2017; Stevers et al. 2018; Albert et al. 2019). Despite the difficulty of discovering and developing PPI modulatory drugs, developing PPI modulators is becoming of increasing strategic importance. However, most of the known small-molecule drugs discovered to date address only 2% of human proteins. Most of the remaining proteins are critically involved in disease pathways, lie within the category of intractable targets and are PPIs (Feng et al. 2017; Lage et al. 2018; Guidolin et al. 2019; Husain et al. 2019). Despite the difficulties in discovering small-molecule PPI modulators that are capable of being taken into clinical trials, drugs are already on the market and some drug-like candidates have reached clinical production. Central to these successes has been the determination of 'hotspots' in protein-protein interfaces. By targeting hotspots, several compounds that directly modulate PPIs have been discovered (Robertson and Spring 2018; Zhang et al. 2018). However, this has been on a sporadic 'one compound at a time' basis (Haberman 2012). Examples include eltrombopag, as well as several others currently being tested in clinical trials. In addition to these direct PPI modulators,

there are allosteric chemokine receptor modulators, which include maraviroc, plerixafor and several others now undergoing clinical trials (Kalota and Gewirtz 2010).

Following failure to meet the strategic needs to expand the number of targets that can be addressed by developable drugs using the sporadic, 'one compound at a time' approach (Haberman 2012), developing second-generation technologies designed to enable the fabrication of small-molecule and peptide PPI modulators on a more consistent basis becomes imperative. These include CS mapping, diversity-oriented synthesis of chemical libraries, DPC technology for synthesis of libraries of macrocyclic compounds and stapled-peptide technology (Drahl 2009; Franzini and Randolph 2016; Zhou et al. 2018). Responding to these new technologies, research vistas are moving back into the PPI modulator field. Whether the new suite of enabling technologies for PPI modulator discovery and development will enable this area to be successful and to meet the strategic needs for discovery and development remains an open question. Most of the compounds that have been discovered using these technologies are in clinical and preclinical stages (Taylor et al. 2018; Stevers et al. 2018; Albert et al. 2019; Fecková et al. 2019; Zhao et al. 2019). It will be necessary for some of these compounds to enter clinical production, achieve proof of concept and thereafter reach the market. Nevertheless, the PPI modulator field is an exciting area that is gaining increasing interest and investment in the PPI-focused technology development curve.

Acknowledgements Abidemi Paul Kappo is thankful to the National Research Foundation (NRF), South Africa for a Thuthuka Grant (Grant No: 107262) award and University of Zululand Research Committee for their support. Llyod Mabonga is thankful to the Department of Science and Technology (DST) and the NRF for a Doctoral Research Bursary. The contents of this manuscript are solely the responsibility of the authors and do not necessarily represent the official views of the DST, NRF and the University of Zululand.

Compliance with ethical standards

Conflict of interest Lloyd Mabonga declares that he has no conflict of interest. Abidemi Paul Kappo declares that he has no conflict of interest.

Ethical approval The article does not contain any studies with human participants or animals hence do not require ethics approval by the authors.

References

Adams JM, Cory S (2007) The Bcl-2 apoptotic switch in cancer development and therapy. Oncogene 26:1324-1337. https://doi.org/10.1038/sj.onc.1210220

Albert L, Peñalver A, Djokovic N (2019) Modulating protein-protein interactions with visible-light responsive peptide backbone photoswitches. ChemBioChem 20:1-14. https://doi.org/10.1002/cbic.201800737

Ali AM, Atmaj J, Oosterwijk NV (2019) Stapled peptides inhibitors: a new window for target drug discovery. Comput Struct Biotechnol J 17:263-281. https://doi.org/10.1016/j.csbj.2019.01.012

Alihodzić S, Bukvić M, Elenkov I et al (2018) Current trends in macrocyclic drug discovery and beyond -Ro5. Prog Med Chem 57:113-233. https://doi.org/10.1016/bs.pmch.2018.01.002

Allison M (2009) Bristol-Myers Squibb swallows last of antibody pioneers. Nat Biotechnol 27:781-783. https://doi.org/10.1038/nbt0909-781

Al-Shehabi H, Fiebig U, Kutzner J et al (2019) Human SAMHD1 restricts the xenotransplantation relevant porcine endogenous retrovirus (PERV) in non-dividing cells. J Gen Virol 100:656-661. https://doi.org/10.1099/jgv.0.001232

Appel A (2011) Drugs: more shots on target. Nature 480:S40-S42. https://doi.org/10.1038/480S40a

Arkin MR, Randal M, DeLano WL et al (2003) Binding of small molecules to an adaptive protein-protein interface. Proc Natl Acad Sci USA 100:1603-1608. https://doi.org/10.1073/pnas.252756299

Bauer RA, Wurst JM, Tan DS (2010) Expanding the range of "druggable" targets with natural product-based libraries: an academic perspective. Curr Opin Chem Biol 14:308-314. https://doi.org/10.1016/j.cbpa.2010.02.001

Basso A, Park SB, Moni L (2019) Editorial: diversity oriented synthesis. Front Chem 6:668. https://doi.org/10.3389/fchem.2018.00668

Booij TH, Price LS, Danen EHJ (2019) 3D cell-based assays for drug screens: challenges in imaging, image analysis, and high-content analysis. SLAS discovery 1-13. https://doi.org/10.1177/2472555219830087

Boone DN, Qi Y, Li Z et al (2011) Egr1 mediates p53-independent c-Myc-induced apoptosis via a noncanonical ARF-dependent transcriptional mechanism. Pro Nat Acad Sci USA 108:632-637. https://doi.org/10.1073/pnas.1008848108

Bhullar KS, Lagarón NO, McGowan EM et al (2018) Kinase-targeted cancer therapies: progress, challenges and future directions. Mol Cancer 17:48. https://doi.org/10.1186/s12943-018-0804-2

Carballo GB, Honorato JR, Farias de Lopes GP (2018) A highlight on Sonic Hedgehog pathway. Cell Commun Signal 16:11. https://doi.org/10.1186/s12964-018-0220-7

Carry JC, Garcia-Echeverria C (2013) Inhibitors of the p53/hdm2 protein-protein interaction—path to the clinic. Bioorg Med Chem Lett 23:2480-2485. https://doi.org/10.1016/j.bmcl.2013.03.034

Ceccarelli DF, Tang X, Pelletier B et al (2011) An allosteric inhibitor of the human Cdc34 ubiquitin conjugating enzyme. Cell 145:1075-1087. https://doi.org/10.1016/j.cell.2011.05.039

Cencic R, Hall DR, Robert F et al (2011) Reversing chemoresistance by small molecule inhibition of the translation initiation complex eIF4F. Proc Natl Acad Sci USA 108:1046-1051. https://doi.org/10.1073/pnas.1011477108

Cerchietti LC, Ghetu AF, Zhu X et al (2010) A small-molecule inhibitor of BCL6 kills DLBCL cells in vitro and in vivo. Cancer Cell 17:400-411. https://doi.org/10.1016/j.ccr.2009.12.050

Cheok CF, Verma CS, Baselga J et al (2011) Translating p53 into the clinic. Nat Rev Clin Oncol 8:25-37. https://doi.org/10.1038/nrclinonc.2010.174

Christian F, Szaszák M, Friedl S et al (2011) Small molecule AKAP-protein kinase a (PKA) interaction disruptors that activate PKA interfere with compartmentalized cAMP signaling in cardiac myocytes. J Biol Chem 286:9079-9096. https://doi.org/10.1074/jbc.M110.160614

Clackson T, Wells JA (1995) A hot spot of binding energy in a hormone-receptor interface. Science 267:383-386. https://doi.org/10.1126/science.7529940

Cohen P, Tcherpakov M (2010) Will the ubiquitin system furnish as many drug targets as protein kinases? Cell 143:686-693. https://doi.org/10.1016/j.cell.2010.11.016

Compton LA, Hiebert SW (2010) Anticancer therapy SMRT-ens up: targeting the BCL6-SMRT interaction in B cell lymphoma. Cancer Cell 17:315-316. https://doi.org/10.1016/j.ccr.2010.03.012

Conn PJ, Christopoulos A, Lindsley CW (2009) Allosteric modulators of GPCRs: a novel approach for the treatment of CNS disorders. Nat Rev Drug Discov 8:41-54. https://doi.org/10.1038/nrd2760

Crane EK, Kwan SY, Izaguirre DI (2015) Nutlin-3a: a potential therapeutic opportunity for TP53 wild-type ovarian carcinomas. PLoS One 10(8):e0135101. https://doi.org/10.1371/journal.pone.0135101

Danial NN, Korsmeyer SJ (2004) Cell death: critical control points. Cell 116:205-219. https://doi.org/10.1016/S0092-8674(04)00046-7

Dash R, Richards JE, Su ZZ et al (2010) Mechanism by which Mcl-1 regulates cancer-specific apoptosis triggered by mda-7/IL-24, an IL-10-related cytokine. Cancer Res 70:5034-5045. https://doi.org/10.1158/0008-5472.CAN-10-0563

Davies SL, Serradell N, Bolos J et al (2007) Plerixafor hydrochloride. Drug Today 32:123-136

Debouck C, Metcalf B (2000) The impact of genomics on drug discovery. Annu Rev Pharmacol Toxicol 40:193-208. https://doi.org/10.1146/annurev.pharmtox.40.1.193

Deshaies RJ (2009) Drug discovery: fresh target for cancer therapy. Nature 458:709-710. https://doi.org/10.1038/458709a

Díaz-Eufracio BI, JesúsNaveja J, Medina-Franco JL (2018) Protein-protein interaction modulators for epigenetic therapies. In: Donev R (ed) Advances in protein chemistry and structural biology, 1st edn. Swansea University, UK, pp 65-84. https://doi.org/10.1016/bs.apcsb.2017.06.002

Dorr P, Westby M, Dobbs S et al (2005) Maraviroc (UK-427,857), a potent, orally bioavailable, and selective small-molecule inhibitor of chemokine receptor CCR5 with broad-spectrum anti-human immunodeficiency virus type 1 activity. Antimicrob Agents Chemother 49:4721-4732. https://doi.org/10.1128/AAC.49.11.4721-4732.2005

Drahl C (2009) Big hopes ride on big rings. ACS Meeting News: constraining molecules in macrocyclic rings could help address challenges in drug discovery. Chem Eng News 87:54-57. https://doi.org/10.1021/cen-v087n036.p054

Driggers EM, Hale SP, Lee J et al (2008) The exploration of macrocycles for drug discovery - an underexploited structural class. Nat Rev Drug Discov 7:608-624. https://doi.org/10.1038/nrd2590

Du L, Grigsby SM, Yao A et al (2018) Peptidomimetics for targeting protein-protein interactions between DOT1L and MLL oncofusion proteins AF9 and ENL. ACS Med Chem Lett 9:895-900. https://doi.org/10.1021/acsmedchemlett.8b00175

Duan Z, Tu M, Zhang Q et al (2018) Novel therapeutic strategy to inhibit growth of pancreatic cancer organoids using a rational combination of drugs to induce mitotic arrest and apoptosis. J Clin Oncol 36:322-322. https://doi.org/10.1200/JCO.2018.36.4_suppl.322

Dustin ML, Bivona TG, Philips MR (2004) Membranes as messengers in T cell adhesion signaling. Nat Immunol 5:363-372. https://doi.org/10.1038/ni1057

Erlanson DA, Fesik SW, Hubbard RE et al (2016) Twenty years on: the impact of fragments on drug discovery. Nat Rev Drug Discov 15:605-619. https://doi.org/10.1038/nrd.2016.109

Erickson-Miller CL, DeLorme E, Tian SS et al (2005) Discovery and characterization of a selective, nonpeptidyl thrombopoietin receptor agonist. Exp Hematol 33:85-93. https://doi.org/10.1016/j.exphem.2004.09.006

Everts S (2008) Piece by Piece. Chem Eng News 86:15-23

Fecková B, Kimáková P, Ilkovičová L et al (2019) Methylation of the first exon in the erythropoietin receptor gene does not correlate with its mRNA and protein level in cancer cells. BMC Genet 20:1. https://doi.org/10.1186/s12863-018-0706-8

Feng Y, Wang Q, Wang T et al (2017) Drug target protein-protein interaction networks: a systematic perspective. Biomed Res Int 2017:1-13. https://doi.org/10.1155/2017/1289259

Ferguson FM, Gray NS (2018) Kinase inhibitors: the road ahead. Nat Rev Drug Discov 17:353-377. https://doi.org/10.1038/nrd.2018.21

Franzini R, Randolph C (2016) Chemical space of DNA-encoded libraries. J Med Chem 59:6629-6644. https://doi.org/10.1021/acs.jmedchem.5b01874

Friedberg JW (2011) New strategies in diffuse large B-cell lymphoma: Translating findings from gene expression analyses into clinical practice. Clin Cancer Res 108:1046-1051. https://doi.org/10.1158/1078-0432.CCR-11-1073

Galloway WR, Isidro-Llobet A, Spring DR (2010) Diversity-oriented synthesis as a tool for the discovery of novel biologically active small molecules. Nat Commun 1:80. https://doi.org/10.1038/ncomms1081

Goard CA, Schimmer AD (2013) An evidence-based review of obatoclax mesylate in the treatment of hematological malignancies. Core Evidence 8:15-26. https://doi.org/10.2147/CE.S42568

Gonzalez MW, Kann MG (2012) Protein interactions and disease. PLoS Comput Biol 8:e1002819. https://doi.org/10.1371/journal.pcbi.1002819

Gorczynski MJ, Grembecka J, Zhou Yet al (2007) Allosteric inhibition of the protein-protein interaction between the leukemia-associated proteins Runx1 and CBFbeta. Chem Biol 14:1186-1197. https://doi.org/10.1016/j.chembiol.2007.09.006

Graff JR, Konicek BW, Carter JH et al (2008) Targeting the eukaryotic translation initiation factor 4E for cancer therapy. Cancer Res 68:631-634. https://doi.org/10.1158/0008-5472.CAN-07-5635

Grosdidier S, Totrov M, Fernández-Recio J (2009) Computer applications for prediction of protein-protein interactions and rational drug design. Adv Appl Bioinforma Chem 2:101-123. https://doi.org/10.2147/AABC.S6272

Grossmann TN, Pelay-Gimeno M, Glas A et al (2015) Structure-based design of inhibitors of protein-protein interactions: mimicking peptide binding epitopes. Angew Chem Int Ed 54:8896-8927. https://doi.org/10.1002/anie.201412070

Guidolin D, Marcoli M, Tortorella C et al (2019) Receptor-receptor interactions as a widespread phenomenon: novel targets for drug development? Front Endocrinol 10:53. https://doi.org/10.3389/fendo.2019.00053

Haberman AB (2012) Advances in the discovery of protein-protein interaction modulators. SCRIP Insights Informa 2012. https://biopharmconsortium.com/2012/04/25/advances-in-the-discovery-of-protein-protein-interaction-modulators-published-by-informas-scrip-insights. Accessed 21 April 2019

Hajduk PJ, Galloway WR, Spring DR (2011) Drug discovery: a question of library design. Nature 470:42-43. https://doi.org/10.1038/470042a

Hall DR, Kozakov D, Vajda S (2012) Analysis of protein binding sites by computational solvent mapping. Methods Mol Biol 819:13-27. https://doi.org/10.1007/978-1-61779-465-02

Hansen SK, Cancilla MT, Shiau TP et al (2005) Allosteric inhibition of PTP1B activity by selective modification of a non-active site cysteine residue. Biochemistry 44:7704-7712. https://doi.org/10.1021/bi047417s

Hansen KB, Yi F, Perszyk RE et al (2018) Structure, function, and allosteric modulation of NMDA receptors. J Gen Physiol 150:1081-1105. https://doi.org/10.1085/jgp.201812032

Haq S, Ramakrishna S (2017) Deubiquitylation of deubiquitylases. Open Biol 7:170016. https://doi.org/10.1098/rsob.170016

Hitzenberger M, Schuster D, Hofer TS (2017) The binding mode of the Sonic Hedgehog inhibitor robotnikinin, a combined docking and QM/MM MD study. Front Chem 5:76. https://doi.org/10.3389/fchem.2017.00076

Horuk R (2009) Chemokine receptor antagonists: overcoming developmental hurdles. Nat Rev Drug Discov 8:23-33. https://doi.org/10.1038/nrd2734

Huang SM, Mishina YM, Liu S et al (2009) Tankyrase inhibition stabilizes axin and antagonizes Wnt signalling. Nature 461:614-620. https://doi.org/10.1038/nature08356

Huang O, Li J, Zheng J (2019) The carcinogenic role of the notch signaling pathway in the development of hepatocellular carcinoma. J Cancer 10:1570-1579. https://doi.org/10.7150/jca.26847

Husain B, Paduchuri S, Ramani SR et al (2019) Extracellular protein microarray technology for high throughput detection of low affinity receptor-ligand interactions. J Vis Exp 143:e58451. https://doi.org/10.3791/58451

Jana T, Ghosh A, Mandal SD et al (2017) PPIMpred: a web server for high-throughput screening of small molecules targeting protein-protein interaction. R Soc Open Sci 4:160501. https://doi.org/10.1098/rsos.160501

Jakubowska MA, Kerkhofs M, Martines C (2018) ABT-199 (Venetoclax), a BH3-mimetic Bcl-2 inhibitor, does not cause Ca2+-signalling dysregulation or toxicity in pancreatic acinar cells. Biochimica et Biophysica Acta (BBA) - Molecular Cell Research 1864:968-976. https://doi.org/10.1111/bph.14505

Jeong WJ, Ro EJ, Choi KY (2018) Interaction between Wnt/β-catenin and RAS-ERK pathways and an anti-cancer strategy via degradations of β-catenin and RAS by targeting the Wnt/β-catenin pathway. Precision Oncology 2:5. https://doi.org/10.1038/s41698-018-0049-y

Jochim AL, Arora PS (2010) Systematic analysis of helical protein interfaces reveals targets for synthetic inhibitors. ACS Chem Biol 5:919-923. https://doi.org/10.1021/cb1001747

Kalatskaya I, Berchiche YA, Gravel S et al (2009) AMD3100 is a CXCR7 ligand with allosteric agonist properties. Mol Pharmacol 75:1240-1247. https://doi.org/10.1124/mol.108.053389

Kalota A, Gewirtz AM (2010) A prototype nonpeptidyl, hydrazone class, thrombopoietin receptor agonist, SB-559457, is toxic to primary human myeloid leukemia cells. Blood 115:89-93. https://doi.org/10.1182/blood-2009-06-227751

Keller TH, Pichota A, Yin Z (2006) A practical view of "druggability". Curr Opin Chem Biol 10:357-361. https://doi.org/10.1016/j.cbpa.2006.06.014

Kim YW, Grossmann TN, Verdine GL (2011) Synthesis of all-hydrocarbon stapled α-helical peptides by ring-closing olefin metathesis. Nat Protoc 6:761-771. https://doi.org/10.1038/nprot.2011.324

Kojima K, Burks JK, Arts J et al (2010) The novel tryptamine derivative JNJ-26854165 induces wild-type p53- and E2F1-mediated apoptosis in acute myeloid and lymphoid leukemias. Mol Cancer Ther 9:2545-2557. https://doi.org/10.1158/1535-7163.MCT-10-0337

Kozakov D, Hall DR, Chuang GY et al (2011) Structural conservation of druggable hot spots in protein-protein interfaces. Proceedings of the National Academy of Sciences USA 108:13528-13533. https://doi.org/10.1073/pnas.1101835108

Landon M, Lancia DR, Yu J et al (2007) Identification of hot spots within druggable binding regions by computational solvent mapping of proteins. J Med Chem 50:1231-1240. https://doi.org/10.1021/jm061134b

Lage OM, Ramos MC, Calisto R et al (2018) Current screening methodologies in drug discovery for selected human diseases. Mar Drugs 16:279. https://doi.org/10.3390/md16080279

Lecker SH, Goldberg AL, Mitch WE (2006) Protein degradation by the ubiquitin-proteasome pathway in normal and disease states. J Am Soc Nephrol 17:1807-1819. https://doi.org/10.1681/ASN.2006010083

Lee CW, Grubbs RH (2001) Formation of macrocycles via Ring-closing olefin metathesis. J Org Chem 66:7155-7158. https://doi.org/10.1021/jo0158480

Lepourcelet M, Chen YN, France DS et al (2004) Small-molecule antagonists of the oncogenic Tcf/beta-catenin protein complex. Cancer Cell 5:91-102. https://doi.org/10.1016/S1535-6108(03)00334-9

Li J, Yang C, Xia Y et al (2001) Thrombocytopenia caused by the development of antibodies to thrombopoietin. Blood 98:3241-3248. https://doi.org/10.1182/blood.V98.12.3241

Li JW, Vederas JC (2009) Drug discovery and natural products: end of an era or an endless frontier? Science 325:161-165. https://doi.org/10.1126/science.1168243

Lipinski CA, Lombardo F, Dominiv BW et al (2012) Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. Adv Drug Deliv Rev 46:3-26. https://doi.org/10.1016/j.addr.2012.09.019

Luise N, Wyatt PG (2019) Diversity-oriented synthesis of bicyclic fragments containing privileged azines. Bioorg Med Chem Lett 29:248-251. https://doi.org/10.1016/j.bmcl.2018.11.046

Lygren B, Taskén K (2008) The potential use of AKAP18delta as a drug target in heart failure patients. Expert Opin Biol Ther 8:1099-1108. https://doi.org/10.1517/14712598.8.8.1099

Ma R, Wang P, Wu J et al (2016) Process of fragment-based lead discovery - a perspective from NMR. Molecules 21:E854. https://doi.org/10.3390/molecules21070854

Majer C, Schüssler JM, König R (2019) Intertwined: SAMHD1 cellular functions, restriction, and viral evasion strategies. Med Microbiol Immunol. 1-17. https://doi.org/10.1007/s00430-019-00593-x

Mella RM, Kortazar D, Roura-Ferrer M (2018) Nomad biosensors: a new multiplexed technology for the screening of GPCR ligands. SLAS Technol 23:207-216. https://doi.org/10.1177/2472630318754828

Miller JL, Church TJ, Leonoudakis D et al (2015) Discovery and characterization of nonpeptidyl agonists of the tissue-protective erythropoietin receptor. Mol Pharmacol 88:357-367. https://doi.org/10.1124/mol.115.098400

Miller JH, Field JJ, Kanakkanthara A et al (2018) Marine invertebrate natural products that target microtubules. J Nat Prod 81:691-702. https://doi.org/10.1021/acs.jnatprod.7b00964

Miszta P, Jakowiecki J, Rutkowska E (2018) Approaches for differentiation and interconverting GPCR agonists and antagonists. Methods Mol Biol 1705:265-296. https://doi.org/10.1007/978-1-4939-7465-8_12

Moellering RE, Cornejo M, Davis TN et al (2009) Direct inhibition of the NOTCH transcription factor complex. Nature 462:182-188. https://doi.org/10.1038/nature08543

Modell AE, Blosser SL, Arora PS (2016) Systematic targeting of protein-protein interactions. Trends Pharmacol Sci 37:702-713. https://doi.org/10.1016/j.tips.2016.05.008

Moll UM, Petrenko O (2003) The MDM2-p53 interaction. Mol Cancer Res 1:1001-1008

Moreira IS, Fernandes PA, Ramos MJ (2007) Hot spots - a review of the protein-protein interface determinant amino-acid residues. Proteins 6:803-812. https://doi.org/10.1002/prot.21396

Newman DJ, Cragg GM (2007) Natural products as sources of new drugs over the last 25 years. J Nat Prod 70:461-477. https://doi.org/10.1021/np068054v

Nguyen M, Marcellus RC, Roulston A et al (2007) Small molecule obatoclax (GX15-070) antagonizes MCL-1 and overcomes MCL-1-mediated resistance to apoptosis. Proceedings of the National Academy of Sciences USA 104:19512-19517. https://doi.org/10.1073/pnas.0709443104

Nielsen TE, Schreiber SL (2008) Towards the optimal screening collection: a synthesis strategy. Angew Chem Int Ed Engl 47:48-56. https://doi.org/10.1002/anie.200703073

Oltersdorf T, Elmore SW, Shoemaker AR et al (2005) An inhibitor of Bcl-2 family proteins induces regression of solid tumours. Nature 435:677-681. https://doi.org/10.1038/nature03579

Pan Y, Wang Z, Zhan W et al (2018) Computational identification of binding energy hot spots in protein-RNA complexes using an

ensemble approach. Bioinformatics 34:1473-1480. https://doi.org/10.1093/bioinformatics/btx822
Pándy-Szekeres G, Munk C, Tsonkov TM et al (2018) GPCRdb in 2018: adding GPCR structure models and ligands. Nucleic Acids Res 46: D440-D446. https://doi.org/10.1093/nar/gkx1109
Parveen A, Subedi L, Kim HW et al (2019) Phytochemicals targeting VEGF and VEGF-related multifactors as anticancer therapy. J Clin Med 8:350. https://doi.org/10.3390/jcm8030350
Roberts AW, Seymour JF, Brown JR et al (2012) Substantial susceptibility of chronic lymphocytic leukemia to BCL2 inhibition: results of a phase I study of Navitoclax in patients with relapsed or refractory disease. J Clin Oncol 30:488-496. https://doi.org/10.1200/JCO.2011.34.7898
Robertson NS, Spring DR (2018) Using peptidomimetics and constrained peptides as valuable tools for inhibiting protein-protein interactions. Molecules 23:959. https://doi.org/10.3390/molecules23040959
Robson-Tull J (2018) Biophysical screening in fragment-based drug design: a brief overview. Bioscience Horizons: The International Journal of Student Research 11:hzy01512. https://doi.org/10.1093/biohorizons/hzy015
Rüdisser S, Vangrevelinghe E, Maibaum J (2016) An integrated approach for fragment-based lead discovery: virtual, NMR, and high-throughput screening combined with structure-guided design. Application to the aspartyl protease renin. In: Erlanson DA, Jahnke W (eds) Fragment-based drug discovery lessons and outlook, 1st edn. Wiley, New York, pp 447-480. https://doi.org/10.1002/9783527683604
Schafmeister CE, Po J, Verdine GL (2000) An all-hydrocarbon cross-linking system for enhancing the helicity and metabolic stability of peptides. J Am Chem Soc 122:5891-5892. https://doi.org/10.1021/ja000563a
Schreiber SL (2009) Organic chemistry: molecular diversity by design. Nature 457:153-154. https://doi.org/10.1038/457153a
Shangary S, Qin D, McEachern D et al (2008) Temporal activation of p53 by a specific MDM2 inhibitor is selectively toxic to tumors and leads to complete tumor growth inhibition. Proc Natl Acad Sci U S A 105:3933-3938. https://doi.org/10.1073/pnas.0708917105
Shangary S, Wang S (2009) Small-molecule inhibitors of the MDM2-p53 protein-protein interaction to reactivate p53 function: a novel approach for cancer therapy. Annu Rev Pharmacol Toxicol.49:223-241. https://doi.org/10.1146/annurev.pharmtox.48.113006.094723
Shore GC, Viallet J (2005) Modulating the bcl-2 family of apoptosis suppressors for potential therapeutic benefit in cancer. Hematology Am Soc Hematol Educ Program 2005:226-230. https://doi.org/10.1182/asheducation-2005.1.226
Silva D, Yu S, Ulge UY et al (2019) De novo design of potent and selective mimics of IL-2 and IL-15. Nature 565:186-191. https://doi.org/10.1038/s41586-018-0830-7
Sinha D, Chowdhury D, Vino S (2012) Monoclonal antibodies (mAbs): the latest dimension of modern therapeutics. Int J Curr Sci 2:9-23
Song X, Lu L, Passioura T (2017) Macrocyclic peptide inhibitors for the protein-protein interaction of Zaire Ebola virus protein 24 and karyopherin alpha 5. Org Biomol Chem 15:5155-5160. https://doi.org/10.1039/c7ob00012j
Soucy TA, Smith PG, Milhollen MA et al (2009) An inhibitor of NEDD8-activating enzyme as a new approach to treat cancer. Nature 458: 732-736. https://doi.org/10.1038/nature07884
Souers AJ, Leverson JD, Boghaert ER et al (2013) ABT-199, a potent and selective BCL-2 inhibitor, achieves antitumor activity while sparing platelets. Nat Med 19:202-208. https://doi.org/10.1038/nm.3048
Stanton BZ, Peng LF, Maloof N et al (2009) A small molecule that binds Hedgehog and blocks its signaling in human cells. Nat Chem Biol 5: 154-156. https://doi.org/10.1038/nchembio.142
Stevers LM, Sijbesma E, Botta M et al (2018) Modulators of 14-3-3 protein-protein interactions. J Med Chem 61:3755-3778. https://doi.org/10.1021/acs.jmedchem.7b00574
Susanto JP (2015) The role of Eltrombopag and Romiplostim as the thrombopoietin receptor agonist (TPO-RA) in treatment of idiopathic thrombocytopenic purpura (ITP): what is TPO-RA, when TPO-RA is used and how to take TPO-RA? Folia Medica Indonesiana 51: 203-207. https://doi.org/10.20473/fmi.v51i3.2840
Takada Y, Ye X, Simon S (2007) The integrins. Genome Biol 8:215. https://doi.org/10.1186/gb-2007-8-5-215
Taylor IR, Dunyak BM, Komiyama Tet al (2018) High throughput screen for inhibitors of protein-protein interactions in a reconstituted heat shock protein 70 (Hsp70) complex. J Biol Chem 293:4014-4025. https://doi.org/10.1074/jbc.RA117.001575jbc.RA117.001575
Tian SS, Lamb P, King AG et al (1998) A small, nonpeptidyl mimic of granulocyte-colony-stimulating factor. Science 281:257-259. https://doi.org/10.1126/science.281.5374.257
Tse C, Shoemaker AR, Adickes J et al (2008) ABT-263: a potent and orally bioavailable Bcl-2 family inhibitor. Cancer Res 68:3421-3428. https://doi.org/10.1158/0008-5472.CAN-07-5836
Trinh PNH, May LT, Leach K et al (2018) Biased agonism and allosteric modulation of metabotropic glutamate receptor 5. Clin Sci 132: 2323-2338. https://doi.org/10.1042/CS20180374
Ubanako PN, Choene M, Motadi L (2015) Mechanisms of apoptosis in ovarian cancer: the small molecule targeting. Int J Med Med Sci 7: 46-60. https://doi.org/10.5897/IJMMS2014.1081
Varshavsky A (2017) The ubiquitin system, autophagy, and regulated protein degradation. Annu Rev Biochem 86:123-128. https://doi.org/10.1146/annurev-biochem-061516-044859
Vassilev LT, Vu BT, Graves B et al (2004) In vivo activation of the p53 pathway by small-molecule antagonists of MDM2. Science 303: 844-848. https://doi.org/10.1126/science.1092472
Venkata Narasimha Rao G, Ravi B, Sunil Kumar M et al (2017) Ultra performance liquid chromatographic method for simultaneous quantification of plerixafor and related substances in an injection formulation. Cogent Chemistry 3:1275955. https://doi.org/10.1080/23312009.2016.1275955
Verhoork SJM, Jennings CE, Rozatian N (2019) Tuning the binding affinity and selectivity of perfluoroaryl-stapled peptides by cysteine-editing. Chemistry 25:177-182. https://doi.org/10.1002/chem.201804163
Walensky LD, Kung AL, Escher I et al (2004) Activation of apoptosis in vivo by a hydrocarbon-stapled BH3 helix. Science 305:1466-1470. https://doi.org/10.1126/science.1099191
Walensky LD, Korsmeyer SJ, Verdine G (2010) Stabilized alpha helical peptides and uses thereof. United States Patent Number 7:469 https://patents.google.com/patent/US7723469B2/en. Accessed 24 Mar 2019
Wan H (2016) An overall comparison of small molecules and large biologics in ADME testing. ADMET & DMPK 4:1-22. https://doi.org/10.5599/admet.4.1.276
Wells JA, McClendon CL (2007) Reaching for high-hanging fruit in drug discovery at protein-protein interfaces. Nature 450:1001-1009. https://doi.org/10.1038/nature06526
Wertz IE, Wang X (2019) From discovery to bedside: targeting the ubiquitin system. Cell Chem Biol 26(2):156-177. https://doi.org/10.1016/j.chembiol.2018.10.022
Wilson CG, Arkin MR (2011) Small-molecule inhibitors of IL-2/IL-2R: lessons learned and applied. Curr Top Microbiol Immunol 348:25-59. https://doi.org/10.1007/82_2010_93
Wyllie AH (2010) "Where, O death, is thy sting?" A brief review of apoptosis biology Mol Neurobiol 42:4-9. https://doi.org/10.1007/s12035-010-8125-5
Xu GG, Guo J, Wu Y (2014) Chemokine receptor CCR5 antagonist maraviroc: medicinal chemistry and clinical applications. Curr Top Med Chem 13 : 1504 - 1514. https://doi.org/10.2174/1568026614666140827143745
Yan M, Li G, An J (2017) Discovery of small molecule inhibitors of the Wnt/β-catenin signaling pathway by targeting β-catenin/Tcf4

interactions. Exp Biol Med 242:1185-1197. https://doi.org/10.1177/1535370217708198

Yasui T, Yamamoto T, Sakai N et al (2017) Discovery of a novel B-cell lymphoma 6 (BCL6)-corepressor interaction inhibitor by utilizing structure-based drug design. Bioorg Med Chem 25:4876-4886. https://doi.org/10.1016/j.bmc.2017.07.037

Yu M, Wang C, Kyle AF et al (2011) Synthesis of macrocyclic natural products by catalysts-controlled stereoselective ring-closing metathesis. Nature 479:88-93. https://doi.org/10.1038/nature10563

Zaba LC, Fuentes-Duculan J, Eungdamrong NJ et al (2010) Identification of TNF-related apoptosis inducing ligand and other molecules that distinguish inflammatory from resident dendritic cells in patients with psoriasis. J Allergy Clin Immunol 125:1261-1268. https://doi.org/10.1016/j.jaci.2010.03.018

Zhang G, Andersen J, Gerona-Navarro G (2018) Peptidomimetics targeting protein-protein interactions for therapeutic development. Protein Pept Lett 25:1076-1089. https://doi.org/10.2174/0929866525666181101100842

Zhao F, Liu W, Yue S et al (2019) Pre-treatment with G-CSF could enhance the antifibrotic effect of BM-MSCs on pulmonary fibrosis. Stem Cells Int 2019:1726743. https://doi.org/10.1155/2019/1726743

Zhong M, Gadek TR, Bui M (2012) Discovery and development of potent LFA-1/ICAM-1 antagonist SAR 1118 as an ophthalmic solution for treating dry eye. ACS Med Chem Lett 3:203-206. https://doi.org/10.1021/ml2002482

Zhou X, Pathak P, Jayawickramarajah J (2018) Design, synthesis, and applications of DNA-macrocyclic host conjugates. Chem Commun 54:11668-11680. https://doi.org/10.1039/c8cc06716c

CHAPTER THREE

A paper already published in *International Journal of Peptide Research and Therapeutics*
2020; 26: 225-241
https://doi.org/10.1007/s10989-019-09831-5

Peptidomimetics: A Synthetic Tool for Inhibiting Protein–Protein Interactions in Cancer

Lloyd Mabonga and Abidemi Paul Kappo

Preface − About the Manuscript

The manuscript introduces peptidomimetics as one of the potential tools for inhibiting protein-protein interactions in PPI-focused smart-drug discovery in cancer. The manuscript begins by elucidating on the structural biology findings of "hotspots" regions in protein-protein binding interfaces. It then delves on the classification of peptidomimetics, flagging the four different classes established relative to the degree of their similarity to the natural peptide precursor. Examples of different PPIs that are targeted using different types of peptidomimetics in cancer pathways are expounded on and encouraging evidence is provided on the successful application of synthetic mimicries tested on different protein model systems that include apoptosis regulators, transmembrane receptors, small GTPases and transcriptional regulators. This connects well with the preceding chapter by highlighting the need to consider peptidomimetics as potential arsenals in inhibiting the SNRPG/RBBP6 PPI. The advantages of using peptidomimetics outweighs the use of their antecedent peptides in that their synthesis allows for convenient preparation of large arrays of analogs by simple replacement/insertion of alternative residues by way of commercial and/or synthetic amino acid analogs. They possess enhanced in vivo stability and often exhibit lowered toxicity as compared to their antecedent peptides.

International Journal of Peptide Research and Therapeutics
https://doi.org/10.1007/s10989-019-09831-5

Peptidomimetics: A Synthetic Tool for Inhibiting Protein–Protein Interactions in Cancer

Lloyd Mabonga[1] · **Abidemi Paul Kappo**[1]

Accepted: 20 February 2019

Abstract

Protein–protein interactions (PPI) are vital in modulating biochemical pathways in many biological processes. Inhibiting PPI is a tremendously important diagnostic and therapeutic strategy in averting pathophysiological cues and disease progression. Targeting PPI as a smart drug discovery tool has been largely overlooked over the years due to their highly dynamic and expansive interfacial areas. However in recent years, researchers have developed new technologies that have the potential to move this approach up the technology development curve and enable the regular discovery of PPI-focused smart drugs. Few drugs are already on the market and some potential drug-like candidates are in clinical trials. In this study we review the application of peptidomimetics as a valuable tool in PPI inhibition in cancer. First, we describe PPI and the general properties of the PPI interface. Next, we discuss the classification of peptidomimetics. Lastly, we focus on the application of peptidomimetics on targeted PPI in cancer pathways.

Keywords Protein–protein interactions · Peptidomimetics · Proteomimetics · Mimicries · Cancer

Introduction

Protein–protein interactions (PPI) are well recognised mediators in biological processes and are vitally important in the progression of many disease states (Du et al. 2018; Robertson and Spring 2018; Zhang et al. 2018). About 650,000 disease-relevant PPI have been so far reported in the human interactome (Bonetta 2010; Gonzalez and Kann 2012). Of which 98% of these interactions remain elusive and under-explored. Over the years PPI were regarded as prototypically "intractable" and "undruggable" due to their highly dynamic and expansive interfacial areas (flat, featureless and relatively large) (Robertson and Spring 2018; Zhang et al. 2018). However due to the improving technology expertise, PPI have now come to the spotlight as significant drug development targets. The PPI-focused drug technology presents an emerging field for drug discovery. This review will focus on the application of synthetic mimicries to target PPI in cancer diagnostics and therapeutics.

✉ Abidemi Paul Kappo
KappoA@unizulu.ac.za

[1] Biotechnology and Structural Biochemistry (BSB) Group, Department of Biochemistry and Microbiology, University of Zululand, 3886 Kwadlangezwa, South Africa

Published online: 27 February 2019

General Properties of PPI

PPI occur over a relatively large protein contact surface area of approximately 1000 to 4000 Å^2. The area is relatively larger as compared to the average contact area needed for inhibition by small molecule binding (300 to 1000 Å^2) (Jones and Thornton 1996; Conte et al. 1999). PPI contact surface area harbor certain hydrophobic regions called "hot spots". Hot spots regions contribute to the binding affinity and help to hold the two interacting proteins together (Clackson and Wells 1995; Jochim and Arora 2010). They are rich in Tyr, Trp, Leu, Ile, Phe and Arg. The amino acids Trp, Arg and Tyr are hydrophobic and form hydrogen bonds which contribute to π-interactions and the binding free energy (Bogan and Thorn 1998). In addition to that, systematic alanine scanning mutagenesis has revealed that the substitution of an amino acid residue by alanine in these hot spot regions lowers the binding affinity by at least 2 kcal/mol (Bogan and Thorn 1998).

Hot spots regions consist of two segments, a core region and a rim region (shown in Fig. 1). The rim region has an amino acid composition similar to that of the rest of the protein contact surface area. The core region consists of aromatic residues (Chakrabarti and Janin 2002; DeLano 2002; Chene 2006).

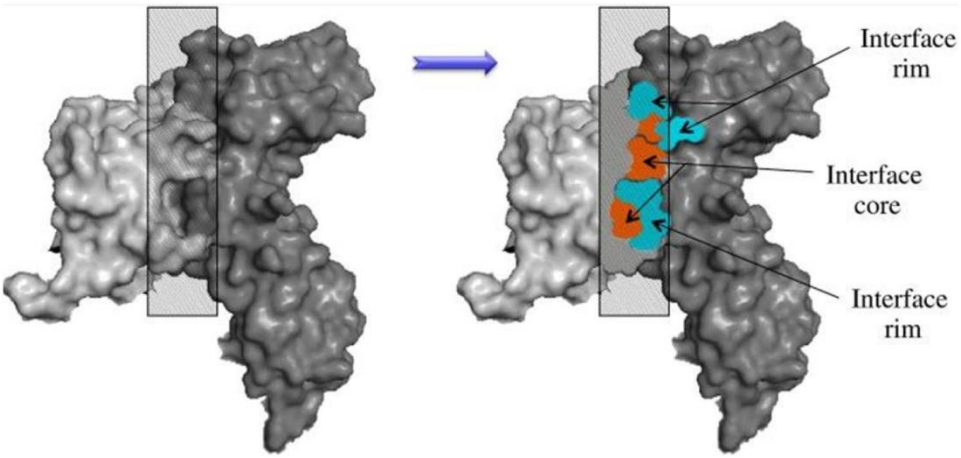


Fig. 1 Schematic diagram of core and rim interface regions. Highlighted is a cross-sectional view of a protein–protein interface. Interacting proteins are presented in light and dark gray, respectively. The interface core is presented in orange and the rim is presented in blue (David and Sternberg 2015). (Color figure online)

Many hot spots core regions are associated with the α-helix, %-sheet and %-turn protein secondary structure motifs. And of these the α-helix has been on the spotlight because they comprise more than 50% of all secondary structures in protein complexes. The α-helix actively binds into the grooves of binding partners and modulates the functioning of a large number of therapeutically relevant PPI. Of which more than 50% bind to one face of the helix (Jochim and Arora 2010; Raj et al. 2013). Peptidomimetics that mimic more than one face of an α-helix have also been reported (Lanning and Fletcher 2015; Robertson and Spring 2018). A majority of these helices contain hot spot residues on one helical face while the rest project critical functionality residues for recognition.

Helix mimicry has become a promising avenue for discovery of potent PPI inhibitors. They have been classified into two categories vizly topographical helix mimics and stabilized helices. Topographical helix mimics contain a non-peptidic scaffold which mimic more than one face of the helix to orient protein-like side chains into proper vectors and mimic the projection of side chains on α-helices. This kind of helix mimicry generally harbors low molecular weight compounds that mimic a single helix face (Bullock et al. 2011; Azzarito et al. 2013; Grossmann et al. 2015). Other than the first developed aromatic scaffold many different topographical helix mimetic scaffolds have been described to afford compounds that are less hydrophobic than the original designs (Orner et al. 2001). And these can target more than one face of a helix (Bullock et al. 2011; Lao et al. 2014a, b; Lanning and Fletcher 2015).

Stabilized helices (foldamers) often mimic 2 to 3 faces of the helix depending on the stabilization technique (Henchey et al. 2008). The side chain staples consist of lactams, thiols, triazole linkages and hydrocarbons which allows two faces for recognition and one for stabilization (Bullock et al. 2011; Azzarito et al. 2013; Grossmann et al. 2015). Both hydrogen bond surrogate (HBS) helices with peptide backbone stabilisation as well as foldamers comprised of judiciously placed α- and %-amino acids mimic proteins that require three faces for recognition (Bullock et al. 2011; Sawada and Gellman 2011; Azzarito et al. 2013; Grossmann et al. 2015). Most of the these dual faced helix peptidomimetics have been developed from single faced bis-benzamide scaffolds (Marimganti et al. 2009; Thompson et al. 2012), an amphiphilic α-helix mimetic based on a benzoylurea scaffold (Thompson and Hamilton 2012) and two-faced amphipathic α-helix mimetics based on a triazine-piperazine-triazine scaffold (Lee et al. 2016). These mimics have been successful in modulating PPI. Recent advances in helix mimicry have been extensively reviewed, and we refer the reader to these excellent reports (Azzarito et al. 2013; Milroy et al. 2014). Other drug-like proteo-mimetics based on a purine scaffold have also been reported (Lanning et al. 2015).

Despite the fundamental role of strands and sheets at protein–protein interfaces, application of %-strand or %-sheet mimics as modulators of PPI is limited. Strand designs are challenging because mimics with appropriately placed hydrogen-bonding groups tend to aggregate (Spiegel et al. 2012). An analysis of the PDB for %-strands found at PPI interfaces reveals that %-strands interact with protein partners in multiple ways: as a lone strand or a sheet, side chain recognition, and with or without engagement of backbone hydrogen bonding (Watkins and Arora 2014). A number of scaffolds for each type of these structures have been designed (Angelo and Arora 2005; Robinson 2008).

The PPI hot spots regions have been on the spotlight as potential drug targets since majority of the binding energy that contributes to interactions localises in these areas (Du et al. 2018; Robertson and Spring 2018; Bogan and Thorn 1998). The disruption of PPI targeting hot spots regions using small molecule or peptide inhibitors both diagnostic and therapeutic significance (Robertson and Spring 2018). It stirs high expectations for the development of smart drugs. Such an observation has successfully challenged the

traditional thought that PPI are "intractable and undruggable". New small molecule or peptide PPI inhibitors are already on the market and some are still in clinical trials (Whitby and Boger 2012; Grossmann et al. 2015; Robertson and Spring 2018). Several strategies and new techniques which aid in the discovery of new PPI and facilitate the discovery of small molecules and peptides inhibitors exist. And these include phage display (Ting et al. 2018), high throughput screening (Taylor et al. 2018), computational studies (Melagraki et al. 2017), crosslinking (Suchanek et al. 2005) and structural based design techniques (for an insightful review see Meireles and Mustata 2011).

Classification of Peptidomimetics

The efficient mimicking of peptides in their bioactive conformation is a long-standing goal in the design of PPI inhibitors. Advances in PPI-focused technology have facilitated a display of side chain functionalities in analogy to peptide secondary structures, yielding molecules that are generally referred to as peptidomimetics (Grossmann et al. 2015). Peptidomimetics are compounds whose essential elements (pharmacophore) mimic a natural peptide or protein in 3D space and which retain the ability to interact with the biological target and produce the same biological effect (Vagna et al. 2009). Over the years peptidomimetics have been traditionally divided into three subtypes; Types I to III. Type I mimetics were defined as short peptides which mimic the secondary structure landscape of the antecedent peptide with minor alterations to the peptide sequence. Type II mimetics were defined as non-peptidic functional molecules based on a scaffold that does not mimic the peptide secondary structure. Type III mimetics were also defined as non-peptidic molecules which match the spatial topology of key interaction motifs of the antecedent peptide (Ripka and Rich 1998; Azzarito 2013; Grossmann et al. 2015; Robertson and Spring 2018). However, these categories have been recently revised and subdivided into four different classes: Classes A–D, where Class A mimetics are the most identical to the antecedent peptide and Class D mimetics show the least similarities (Grossmann et al. 2015).

Class A mimetics, like Type I mimetics, are peptides with minimal alterations to the peptide side chains and backbone. They consist mainly of the antecedent peptide amino acid sequence with a limited number of modified amino acids incorporated to stabilize the bioactive conformation. The backbone and side chains align closely with the bioactive conformation of the antecedent peptide (Grossmann et al. 2015). Class B mimetics are modified class A mimetics with different unnatural amino acids, isolated small-molecule building blocks or major backbone alterations. While still peptidic in nature, Class B mimetics include much more dramatic backbone and side chain alterations (e.g., peptoids, %-peptides and α/%-mixed peptides). Foldamers (%- and α/%-peptides as well as peptoids) with side chains aligning topologically identical to the antecedent peptide also form part of this class (Grossmann et al. 2015).

Class C mimetics consists of highly modified structures that completely replace the entire peptide backbone with small molecule character (Grossmann et al. 2015; Robertson and Spring 2018). The central scaffold displays the substituents in comparison to the orientation of the key residues (for example hot spots) in the bioactive conformation of the antecedent peptide (Grossmann et al. 2015). The replacement of the entire peptide backbone result in molecules with improved oral bio-availability and pharmacokinetic properties. The resulting bioactive compounds are more likely to follow Lipinski's rule of five, thus rendering them promising candidates in drug development (Spiegel et al. 2012).

Class D mimetics are small molecule drugs used in the classical medicinal chemistry that mimic the mode of action of a bioactive peptide without a direct link to its side chain functionalities. The small molecule drugs bind either into the active site of a protein or at an allosteric position. However as mentioned earlier on, PPI are large and remain a challenging to target with small molecules. Nevertheless, few small molecule PPI inhibitors have been successfully developed on a "one compound at a time'' basis (Bunnage 2011; Stockwell 2011) through affinity optimization of class C molecules and screenings of compound/virtual libraries (Grossmann et al. 2015). Small molecule drugs are already on the market and some are in clinical trials (Grossmann et al. 2015).

Targeted Protein–Protein Interactions

Peptidomimetics are designed for a broad range of targets in cancer diagnostics and therapeutics. Their applicability has been tested on different protein model systems that include apoptosis regulators, transmembrane receptors, small GTPases and transcriptional regulators. Here we will discuss the application of class A–C mimetics for these cancer model systems. This section does not present all the examples of peptidomimetics comprehensively, but focuses on major target classes and recent contributions.

Apoptosis Regulation

MDM2 and MDMX

In response to cellular stress the transcription factor p53 mediates the expression of genes involved in protective processes such as DNA repair, cell cycle arrest and apoptosis (Vogelstein et al. 2000; Chene 2003). Binding of MDM2

and MDMX (also known as MDM4 and HDM4/ HDMX) to the N-terminal transactivation domain of p53 blocks the normative function of this so called "guardian of the genome". MDM2 and MDMX downregulate the tumor suppressor p53 either by acting as a direct antagonist to p53 or by mediating the ubiquitylation of p53 leading to its degradation in proteasome-dependent manner (Kubbutat et al. 1997; Toledo et al. 2006). An upregulation of MDM2 and MDMX has been detected in different types of cancers and the interactions between these proteins and p53 have become prime targets for anticancer strategies. Crystal structures of the p53-MDM2 and p53–MDMX complexes reveal an α-helical conformation of the p53 interaction domain when bound to MDM2 (Fig. 2a) (Popowicz et al. 2007). The p53 hot-spot comprise of the amino acid residues Phe19, Trp23 and Leu26 (Kussie et al. 1996). This structural information together with the crystallographic data have been vital as the starting point for a rational design of the corresponding peptidomimetics.

For some peptidomimetics, helical peptides derived from phage-display selections served as alternative starting points. For example pDi (Phan et al. 2010) and PMI (Pazgier et al. 2009). These peptides exhibit dual inhibitory effects for both the p53–MDM2 and p53–MDMX complexes. A commendable feature for efficient anticancer activity (Phan et al. 2010; Pazgier et al. 2009). For other peptidomimetics, mirror-image phage-display (MIPD) techniques coupled with native chemical ligation have provided proteolytically more-resistant D-peptide inhibitors of the p53–MDM2 interaction. However, these peptides do not feature sufficient cell permeability (Liu et al. 2010a, b, b; Zhan et al. 2012). Finally, though mRNA display has facilitated the selection of larger libraries of peptides, the proteolytic instability and/or poor cellular uptake of these peptides remain major limitations of these approaches (Shiheido et al. 2011). A variety of peptidomimetics were designed based on these peptide binders.

A modified octapeptide comprising four unnatural amino acids is an early example of a class A peptidomimetic that binds HDM2 in vitro with nanomolar affinities (Bottger et al. 1996; Garca-Echeverra et al. 2000; Sakurai et al. 2006). This peptide promotes apoptosis by mediating the accumulation of p53 in cancer cells (Chene et al. 2000). Poor cellular uptake and high proteolytic instability remain the drawbacks of this peptide. Later, Robinson and co-workers grafted the crucial residues of the p53 helix onto a cyclic %-hairpin and stabilized the %-sheet structure using the head-to-tail macrocyclization and the d-Pro-l-Pro (p-P) turn mimetic. The mimetic displayed good affinity and binds HDM2 at the p53 binding site (Fasan et al. 2004). Sequence optimization by the introduction of unnatural amino acids yielded class B mimetics with improved affinities (Fasan et al. 2006). Remarkably, this innovative approach impressively illustrates the interchangeability of secondary structures and represents one of the few examples of a stabilized %-sheet structure used as a PPI inhibitor.

The generation of class A helix mimetics for MDM2 and MDMX using the peptide-stapling technique and thiol- and triazole-based cross-links observed a prominent increase of

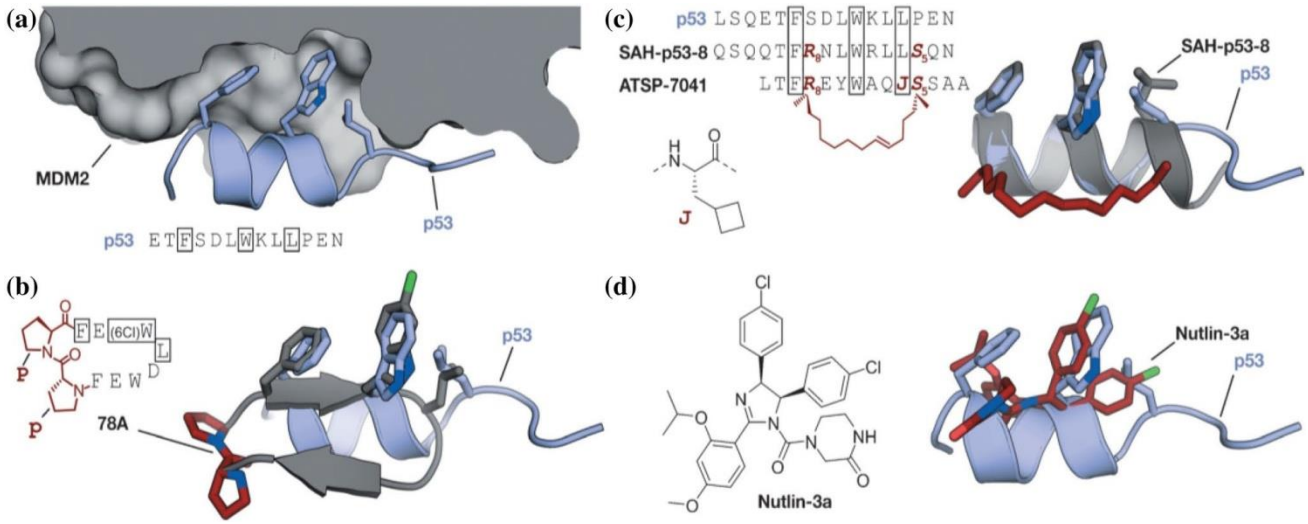


Fig. 2 MDM2–p53 interaction: **a** Crystal structures of MDM2 (gray) with the transactivation domain of p53 (blue, PDB 1YCR) (Kussie et al. 1996). **b** Superimposed crystal structures of p53 (blue, PDB 1YCR) and cyclic b-hairpin peptide 78A (gray/red, PDB 2AXI). The d-Pro-l-Pro (p-P) crosslink is highlighted in red (Fasan et al. 2004). **c** Sequences of stapled peptides (left). Superimposed crystal structures (right) of p53 (blue, PDB 1YCR) and SAHp53-8 (gray/red, PDB 3V3B). The cross-link is highlighted in red (side chains of amino acids in boxes are shown explicitly in the crystal structures) (Baek et al. 2012). **d** Superimposed crystal structures of p53 (blue, PDB 1YCR) and Nutlin-3a (red, PDB 4HG7) (Anil et al. 2013) All superimposed structures were obtained from structures of complexes with MDM2 or MDMX. (Color figure online)

cellular uptake, thus alleviating the challenges regarding cellular permeability and proteolytic instability (Madden et al. 2011). The incorporation of a bisaryl cross-link at positions i and i + 7 of the pDi sequence enhances the α-helicity and bioactivity of the class A mimetics. A cross-linking based on addition of photo induced 1, 3-dipolar cyclo produced high affinities peptides for MDM2 and MDMX. The peptides displayed dual inhibitory activity and improved cellular uptake after the incorporation of positively charged amino acids (Madden et al. 2011). A D,L-dicysteine-linked 6,6'bis(bromomethyl)-3 3'-bipyridine (Bpy) crosslink observed additional MDMX contacts allowing more affine binders (Muppidi et al. 2011). A double triazole tethering approach based on a single p53-derived sequence enables the synthesis of several cross-linked peptides using a set of modified linkers. The resulting cross-linked peptides observed improved proteolytic stability and affinities. And the subsequent incorporation of Arg moieties in the linker resulted in cell penetrating peptides, thus eliminating the need for additional sequence variations (Lau et al. 2014a, b). Metallopeptides and HBS stabilized helices also produced MDM2 affine binders (Henchey et al. 2010a, b; Zaykov and Ball 2011).

Stapled α-helical p53-derived peptides (SAHp53, Fig. 2c) with i and i + 7 cross-linking positions showed increased α-helicity and improved binding affinity for MDM2. The peptides observed enhanced proteolytic stability as compared to the wild-type p53 peptide. Neutral and positively charged stapled peptides following the substitution of negatively charged amino acids induced apoptosis, cell permeability and suppressed tumor growth features in vivo (Bernal et al. 2007, 2010). The direct involvement of the hydrocarbon cross-link in MDM2 binding following the incorporation of a staple explains the increase in the binding affinity (Baek et al. 2012). Furthermore, Aileron Therapeutics reported another series of stapled peptides based on phage-display-derived peptide pDi which include ATSP-7041-peptides and another candidate currently in clinical trials (Chang et al. 2013). The ATSP-7041-peptides observed improved pharmacokinetic properties with high specificity and affinity for both MDMX and MDM2. They bind to mutated forms of MDM2 that are inaccessible for small-molecule p53–MDM2 inhibitors of the Nutlin family (Fig. 2c) (Brown et al. 2013; Wei et al. 2013).

The development of class B mimetics using foldamers as validated scaffolds proved useful for the generation of p53–MDM2 inhibitors. The p53-MDM2 interactions active residues (Phe19, Trp23, and Leu26) were integrated into the recognition face of a 14-helix %-peptide. The helical structure was coerced using the electrostatic macrodipole strategy to proffer micromolar binders. (Kritzer et al. 2005). Several techniques to synthesize and evaluate %-peptides targeting MDM2 have been reported however their relatively poor binding affinities suggest that the 14-helix may not reproduce the p53–MDM2 interaction suitably (Murray et al. 2005). Non-natural side chains were introduced into these %-peptides and moderate improvements were observed in the biological activity (Michel et al. 2009). The cellular uptake of these %-peptides was increased by conjugation to cell-penetrating peptides (Hintersteiner et al. 2009) and by the introduction of side chain to side chain cross-links or %-homoarginines (Bautista et al. 2010). HBS α/%-peptides harboring the ααα% pattern and α-amino acids hot spots yielded affine MDM2 binders with improved conformational rigidity (Patgiri et al. 2012). And rationally designed achiral peptoids with high conformational flexibility observed moderate inhibitory activity of the p53–MDM2 complex (Hara et al. 2006).

Class C structural mimetics were used to inhibit the interaction between p53 and MDM2. Hamilton and group developed a series of trisubstituted terphenyls scaffolds (3,2',2''-terphenyl compounds) that mimic an α-helix face in order to target PPI. The terphenyls' aryl cores adopt a wobbled dihedral conformation (59.1° and 120.7°) to mimic the helix residues (i, i + 3, i + 4 and i + 7) through the ortho positions of the scaffold (Orner et al. 2001). The Hamilton group also developed other extended α-helix scaffolds such as terephthalamides (Yin et al. 2005), 4,4'-dicarboxamines (Rodriguez et al. 2009), 5-6-5-imidazole-phenylthiazoles (Cummings et al. 2009), trispyridylamines (Ernst et al. 2003) and enaminones (Adler et al. 2012). Sterically enforced terphenyls (27) with large aromatic substituents at the central position and aliphatic groups at the termini were used to mimic the binding epitope of p53. These mimetics proved to be active in cell-based assays and exhibit highest affinity for MDM2 Notably, these compounds also and the best selectivity when binding between MMD2 and BCL-2 family proteins (Yin et al. 2005; Chen et al. 2005). Wilson and group described a solid-phase synthesis for an α-helix mimetic with N-alkylated oligobenzamides as well as hybrids which act as inhibitors of p53-HDM2 complex in vitro (Campbell et al. 2010; Long et al. 2013; Barnard et al. 2014, 2015; Azzarito et al. 2015). Notably, the same group reported an orthogonal chemical functionalisation of non-peptidic helix mimetics using a copper-mediated 'click' technique (Barnard et al. 2014).

Spiroligomers disrupted the p53–HDM2 complex and trigger HDM2 accumulation in cells assumably by preventing proteolytic degradation (Brown et al. 2012). Covalently constrained OHMs were also able to bind MDM2 in vitro (Lao et al. 2014a, b). Furthermore, cell permeable pyrrolopyrimidines were also utilised to disrupt both the p53–MDM2 and the p53–MDMX complexes, thereby facilitating p53-dependent apoptosis in cultured cancer cells (Lee et al. 2011). The other groups have developed scaffolds to mimic amino acid side chains

on α-helices. And these include phenyl-piperazine-triazines (Moon et al. 2014), pyrazines (Van Mileghem et al. 2017), 2-O-alkylated picolinamides (Yap et al. 2012), 3-O-alkylated oligobenzamides (Plante et al. 2009; Prabhakaran et al. 2013) and pyridazines (Biros et al. 2007; Londregan et al. 2016). Although less formidable, the hydrogen-bond-guided mimetics also act as inhibitors of p53-HDM2 complex in vitro (Barnard et al. 2015).

Class D peptidomimetics are small molecule p53–MDM2 inhibitors developed from lead structures obtained from high-throughput screening of synthetic chemical libraries (Wade et al. 2013). The benzodiazapinediones (Johnson and Johnson Pharmaceuticals) along with the Nutlin family of small molecule compounds (Hoffmann-La Roche) inhibit the p53/mDM2 PPI in the initiation of cancer (Vassilev et al. 2004; Grasberger et al. 2005). Nutlins bind to MDM2 through the p53 binding site and, induce cell-cycle arrest and apoptosis in a p53-dependent manner. They are highly potent and selective compounds. Their rigid scaffold enables efficient mimicking of p53 binding (Fig. 2d). Nutlins are also known to inhibit tumor growth in human xenograft models (Vassilev et al. 2004; Grasberger et al. 2005).

BCL-2 Family Proteins

BCL-2 family proteins play a key role in apoptosis regulation. Members of the BCL-2 family participate in a complex network of PPI in either pro-apoptotic manner (e.g. BAK, BAX, BID, BIM, NOXA, HRK, PUMA, BAD) or anti-apoptotic manner (e.g. BCL-xL, BCL-2, BCL-w, MCL-1, A1) (Youle and Strasser 2008; Moldoveanu et al. 2014). The interactions between members of the two classes of proteins are directed in the sensing of cellular stress thereby modulating induced cell death by apoptosis. Pro-apoptotic proteins are classified into three, namely; effectors, direct activators and de-repressors/sensitizers. Anti-apoptotic proteins, just like the effectors (e.g. BAK, BAX), have four BCL-2 homology domains (BH1–BH4) that harbors a shared folding motif which creates a hydrophobic groove called the BC groove. The BC groove mediates the binding to an α-helical stretch of BH3-only proteins, including direct activators (BID, BIM, and PUMA) and de-repressors/sensitizers (BAD, NOXA and HRK). This binding interaction involves highly conserved hydrophobic and polar residues that closely interact with the BC groove (Fig. 3a). The specificity required for the interactions within the BCL-2 family members is precisely orchestrated through the variations in the remaining BH3 sequence (Youle and Strasser 2008; Moldoveanu et al. 2014).

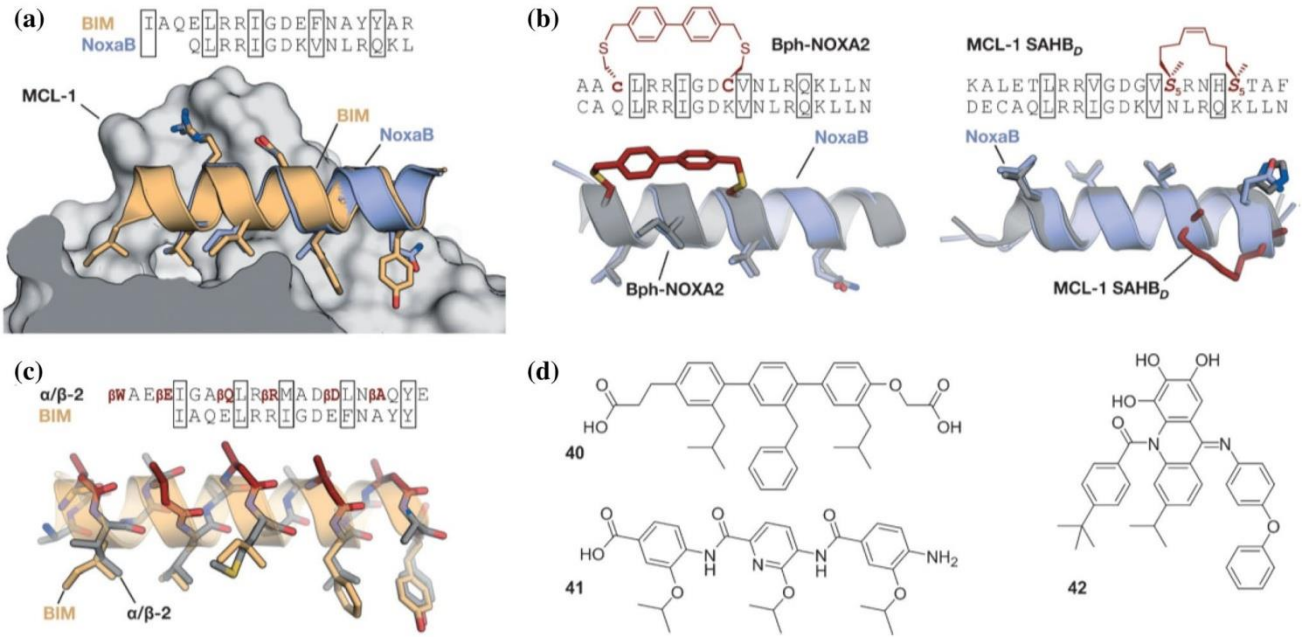


Fig. 3 PPI involving proteins of the BCL-2 family: **a** Superimposed crystal structures of BIM (orange, PDB 2L9) and NoxaB (blue, PDB2NLA) bound to MCL-1 (Czabotar et al. 2007). **b** Superimposed crystal structures of NoxaB (blue, PDB 2NLA) with (left) bisaryl cross-linked peptide Bph-Noxa2 (gray, PDB 4G35, c = d-cysteine) (Muppidi et al. 2012) and (right) stapled peptide MCL-1 SAHBD (gray, PDB 3MK8) (Stewart et al. 2010). Cross-links are highlighted in red (side chains of amino acids in boxes are shown explicitly). **c** Superimposed crystal structures of BIM (orange, PDB 2L9) and α/%-peptide a/b-2 (gray/red, PDB 4BPI) (Smith et al. 2013). %-Amino acids are highlighted in red (%E, %Q, %R, %D, and %A are %3-amino acids that correspond to E, Q, R, D, and A, respectively). **d** Structural mimetics of helical MCL-1 binding peptides (Li et al. 2014). (Color figure online)

Proteins of the BCL-2 family are considered high-interest targets in the field of drug development and their modulation has been widely addressed using different strategies (Moldoveanu et al. 2014). The relevant PPI between BCL-2 family members have been targeted for inhibition using class A peptidomimetics. For example peptides stabilized by thiol-based cross-links, hydrocarbon-stapling approaches and hydrogen-bond surrogates. Class B mimetics such as α/%-peptides and class C mimetics such as sterically constrained and hydrogen-bond-guided structural mimetics have been used as inhibitors of these interactions (Grossmann et al. 2015).

The cross-linking of D-Cys(c) and L-Cys using a bisaryl moiety at positions i and i + 7 stabilized the NOXA derived peptide and provided a selective binder of MCL-1. The NOXA derived peptide-MCL-1 complex (Fig. 3b) observed the presence of edge-to-face π-π interactions between the MCL-1 and the aryl cross-link (Muppidi et al. 2012). Utilising this structure as a foundation for further modifications and to improve the cellular activity, the hydrophobicity was increased by the introduction of backbone N-methylation and replacing non-interacting charged amino acids with Ala. A number of "stabilized a-helices of BCL-2 domains" (SAHBs) have been produced by using the peptide-stapling technique. The incorporation of the hydrocarbon cross-link increased the resistance to proteolysis, cellular uptake and helicity of these BH3-derived peptides. However, not all of these peptides were efficient enough to inhibit the PPI between BCL-2 family members. In leukemia cells, the SAHB from the BH3 domain of the BID protein proved to induce apoptosis in vitro and in vivo (Walensky et al. 2004). The MCL-1-derived SAHB inhibits the BAK-MCL-1 complex formation and subsequently induce cell death by caspase-dependent apoptosis. The direct participation of the staple in target binding is proved in the complex of the crystal structure of this stapled peptide with MCL-1 (Fig. 3b). However synthesis, testing and screening of several stapled peptides was required to select efficient PPI inhibitors (Stewart et al. 2010). SAHB peptides also provide valuable insights into the molecular regulation of proteins of the BCL-2 family (Gavathiotis et al. 2008).

Class B peptidomimetics which include heterogeneous (e.g. α/% peptides) and chimeric foldamers (e.g. α/% + α) provided the desired inhibitors to disrupt PPI between proteins of the BCL-2 family. Pure %-peptides did not inhibit these PPI. The ααα% pattern was used to mimic the BIM BH3 helix and provide binders of BCL-xL and MCL-1 proteins (Boersma et al. 2012). The αα%ααα% backbone was used to a PUMA BH3 derived peptide to obtain foldamers with high binding affinity for the same proteins (Horne and Gellman 2008). Surprisingly the selectivity for the targets in both cases is highly dependent on the number and position of α-to-%3 replacements.

According to Sadowsky et al. (2007), a chimeric peptide (α/% + α) with a 9-mer α/%-peptide at the N-terminus and a 6-mer α-peptide at the C-terminus proved more potent than the natural BAK 16-mer. The chimeric peptide efficiently inhibited formation of the BAK-BCL-xL complex by binding at the same position targeted by the natural peptide (Sadowsky et al. 2007). The N-terminal fragment features an equivalent ratio of α- and % amino acids and projects a new helical disposition called the 14/15 helix (Hayen et al. 2004). The proteolytic stability and selectivity within the BCL-2 family members increased and a subsequent release of cytochrome c was observed in cell lysates. The optimization provided foldamers with increased proteolytic stability but insignificant cellular uptake (Sadowsky et al. 2007).

Using terphenyls as sample models of class C peptidomimetics to mimic the location of hot-spot residues of helical BH3 peptides (Yin et al. 2005), a number of terphenyls inhibited the interaction of BCL-2, BCL-xL and MCL-1 with either BAX or BAK, or with BAD or BIM in cultured cells. This inhibition triggers cell induced death by apoptosis in a caspase-dependent manner (Kazi et al. 2011). A number of heterocyclic scaffolds containing pyridazine also inhibited the the BAK–BCL-xL complex formation in vitro (Biros et al. 2007). Biphenyls, benzoylureas, and trispyridylamides, derived from hydrogen-bond-guided scaffolds were also reported to inhibit the formation of BAK–BCL-xL complex in vitro (Rodriguez et al. 2009). The oligoamide scaffolds made from different combinatorial ratios of pyridine and phenyl rings observed that the molecules containing a higher percentage of phenyl rings inhibits the BAK–BCL-xL complex more efficiently because of their increased hydrophobicity and flexibility.

However, this trend was not observed in cell-based assays probably because of potential off-target effects and differences in cell permeability (Yap et al. 2012). The scaffold with one pyridine ring and two phenyl rings inhibits the formation of BAK–BCL-xL and BAK–MCL-1 complexes thereby mediating cell induced death by apoptosis in cancer cell lines. Noteworthy, the compound also exhibits inhibitory effects on tumor growth in mouse models (Cao et al. 2013). In human cell culture, terephthalamides also disrupted the BAK–BCL-xL complex formation (Rodriguez et al. 2009). Computational studies and NMR spectroscopy observed binding to the same cleft as the BAK BH3 peptide. Finally, BIM–BCL-2 and BIM–MCL-1 PPI were addressed using cross-acridine scaffolds (Li et al. 2014).

Class D mimetics of small molecule have also been used to inhibit the same interactions. ABT-236 and ABT-737 are examples of small molecule Class D peptidomimetic which inhibit the Bcl-xL/Bak PPI in the apoptosis pathway (Lee et al. 2007, Tse et al. 2008). ABT-263 binds with high affinity to anti-apoptotic proteins BCL-2 and BCL-XL and with lower affinity to BCL-w (Tse et al. 2008). ABT-263

has demonstrated impressive single agent activity against lymphoid malignancies and small cell lung cancer (SCLC). Phase I/II trials report that ABT-263 was either effective as a single agent (Roberts et al. 2012) or in combination with other drugs in refractory chronic lymphocytic leukaemia (CLL) (Kipps et al. 2017). ABT-737 binds to Bcl-2/Bcl-xL (but not Mcl-1) with high affinity and disrupts their interaction with pro-apoptotic Bax/Bak, thus enhancing apoptosis (Parrondo et al. 2013).

Inactivating the anti-apoptotic BCL-2 family proteins with small molecule BH3 mimetic drugs is one potential 'push' as displacing the active but sequestered pro-apoptotic proteins results in mitochondrial outer membrane permeabilization (Leber et al. 2010; Del Gaizo Moore and Letai 2013; Shamas-Din et al. 2013). Small molecule BH3 mimetics, like ABT-263 (Navitoclax) and ABT-199 (Venetoclax), mimic the binding of BH3 peptides to the hydrophobic BH3 domain-binding groove of anti-apoptotic proteins and thus displace BH3-only proteins and active BAX/BAK from anti-apoptotic proteins (Tse et al. 2008; Souers et al. 2013). By binding to the BH3 domain-binding grooves of anti-apoptotic proteins, ABT-263 inhibits BCL-2, BCL-XL and BCL-W, whereas ABT-199 only inhibits BCL-2. ABT-199 is approved for use in chronic lymphocytic leukemia and both drugs are being used in dozens of clinical trials as single agents and in combination with other therapies (Delbridge et al. 2016). Other BH3 mimetics are emerging (Ashkenazi et al. 2017). They include another BCL-2-specific inhibitor (Servier's S55746); the BCL-X_L-specific WEHI-539 (Lessene et al. 2013) and its more potent derivatives A-1155463 and A-1331852 (Leverson et al. 2015).

Some cancers depend primarily on MCL-1 for survival (Grabow et al. 2014; Zhang et al. 2011; Xiang et al. 2010) and others acquire resistance to drugs that target BCL-2/BCL-XL/BCL-W by upregulating MCL-1 (Adams and Cory 2018). The small molecule MCL-1 inhibitor, S63845 shows promise as a therapeutic (Kotschy et al. 2016). S63845 was efficacious in killing multiple cancer-derived cell lines in vitro and had potent anti-tumor activity in pre-clinical mouse models of hematological malignancies in vivo while sparing normal tissues. Another cancer treatment strategy would be small molecule activation of BAX and/or BAK. The BAK BH3 helix was stabilized using the hydrogen-bond surrogates thereby increasing the helicity and proteolysis resistance. However, it resulted in the loss of binding affinity as compared to the wild-type peptide. A subsequent sequence optimization provided a peptide with improved affinity (Adams and Cory 2018).

Transmembrane Receptors

Transmembrane receptors are vital in signaling processes that connect extracellular events with intracellular responses. Their impaired functioning is implicated with numerous pathogenic states that include cancer (Yarden and Pines 2012). Receptors are activated by the binding of effectors (protein ligands, peptide hormones and small molecules) and may at times require additional cofactors, highlighting the complexity of the signaling networks. Over the years several PPI inhibitors that are recognized by receptors have been designed from peptide sequences. Examples of such PPI inhibitors include α/%-peptides that target the receptor binding site of vascular endothelial growth factor (Haase et al. 2012) and helical % peptides that inhibit the interaction between the high-density lipoprotein and scavenger receptor B (Werder et al. 1999).

A hyperactivity of epidermal growth factor receptor (EGFR) tyrosine kinase is implicated in the tumorigenesis and tumor development different types of cancer (Yarden and Pines 2012). PPI inhibitors of EGFR target the interaction between EGFR and cofactor Grb2 (growth factor receptor bound protein 2) (Furet et al. 1998), the intracellular adenosine triphosphate binding site (Yarden and Pines 2012) or the extracellular receptor binding site (Li et al. 2005). The dimerization mediated by a coiled coil structure is implicated as a vital step for the receptor activity (Jura et al. 2009). And to inhibit this dimerization, peptides which proved active in cell-based assays were developed. These include all-hydrocarbon-stapled peptides and corresponding peptides with an open cross-link bearing the two olefin side chains (Walensky and Bird 2014; Grossmann et al. 2015). Hanold and coworkers also introduced a non-helical, triazolyl-bridged peptide targeting EGFR dimerization (Hanold et al. 2015).

G-protein coupled receptors (GPCR) are a large family of transmembrane receptors that are activated by a number of different ligands which include but not limited to peptide hormones. Several inhibitors of peptide ligand/receptor interactions have been developed and have been extensively reviewed (Ruiz-Gomez et al. 2010). The similarity observed in the interactions between receptors and peptides or proteins suggests that the application of PPI inhibitor concept to interfere with peptide–receptor interactions may also be a potential drug target. The incorporation of benzodiazepines into Angiotensin II and the use of a glucose scaffold presenting Somatostatin side chains in a %-turn conformation are good examples. The final peptidomimetics observed affinity for AT1 and AT2 receptors (Gallo-Payet et al. 2011) and a potent agonist of the Somatostatin receptor (Hirschmann et al. 1993), respectively. The trans-pyrollidine-3,4-dicarboxamide scaffold led to high-affinity ligands for human opioid receptors.

The glucose and the trans-pyrollidine-3,4-dicarboxamide highlighted above are both class C structural-turn mimetics (Whitby et al. 2011). GPCR protein effectors are also known to interact with agouti (ASP) and agouti-related protein

(AGRP) probably because of the C-terminal binding site that reveals a cysteine knot presenting three crucial residues in a turn structure as observed by the NMR (McNulty et al. 2001). The isolated binding motif can be chemically stabilized by addition of a lactam bridge on the position initially occupied by a disulfide bridge (Thirumoorthy et al. 2001). Other GPCRs recognize binding partners through their helical interaction domains. Helical α/%-peptides inhibit the interaction between parathyroid hormone and the parathyroid hormone-related peptide receptor (Cheloha et al. 2014), whereas hydro carbon stapled peptides with enhanced agonist potency are agonists of vaso-active intestinal peptide receptor 2 (VPAC2) (Giordanetto et al. 2013).

Integrins play a vital role in the interaction of extracellular matrix protein with the cell surface and in cell– cell adhesion in vertebrates. Misregulation of certain integrin receptors is linked to several diseases including cancer (Desgrosellier and Cheresh 2010). Integrins are composed of an α- and a %-subunit; many of which recognize binding partners through an Arg-Gly-Asp (RGD) sequence (Fig. 4a) (Pierschbacher and Ruoslahti 1984). Haubner and co-workers integrated the RGD sequence into cyclic pentapeptides and increased their activity and bioavailability (Haubner et al. 1997). Optimization efforts resulted in the identification of the macrocyclic inhibitor cyclo (RGDfV) called Cilengitide (Dechantsreiter et al. 1999) and the cyclic pentapeptide cyclo (RGDf-N(Me)V). The latter combines high receptor affinity and selectivity with improved biostability and oral availability (Fig. 4b) (Conibear et al. 2014).

Small GTPases

Small GTPases are switch-like proteins that exist in two distinct conformational states and are defined by their binding

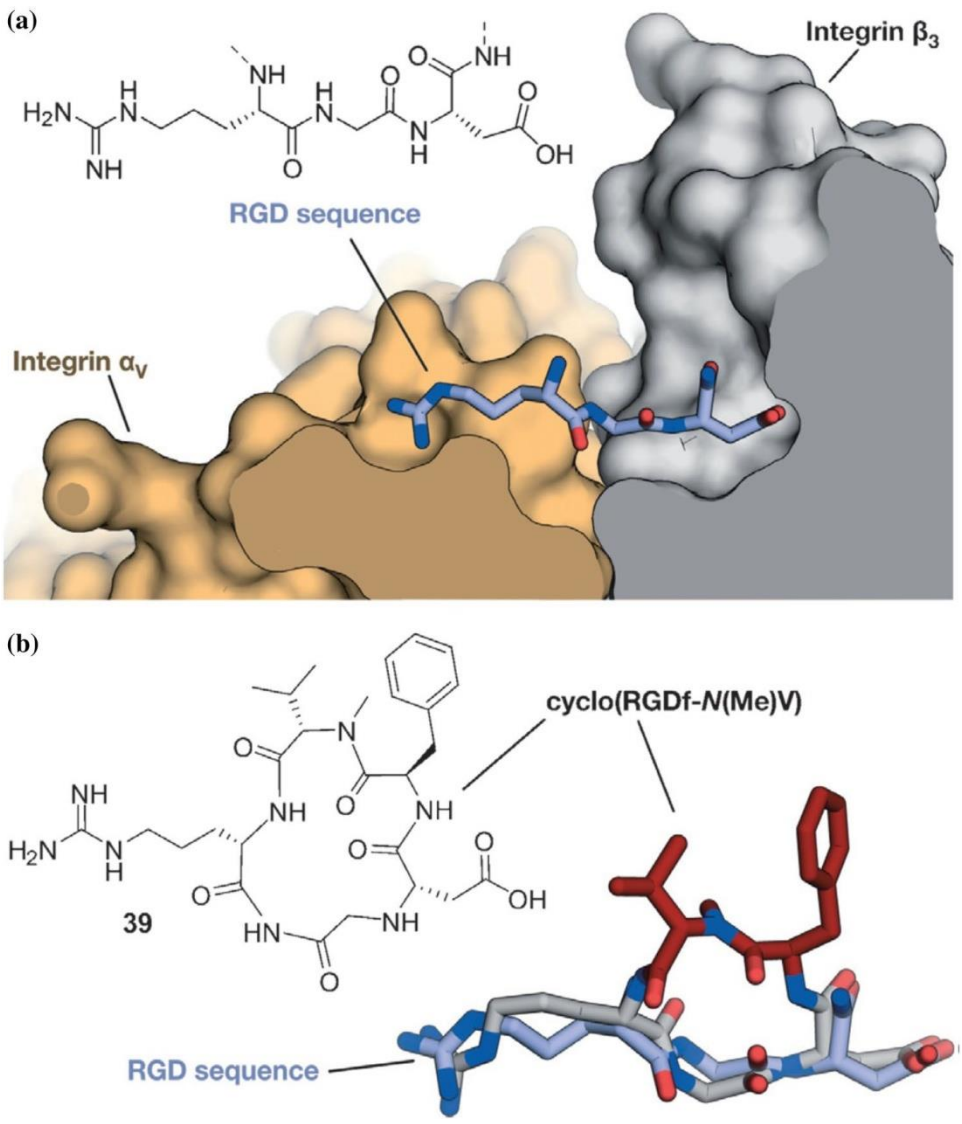


Fig. 4 RGD–integrin interaction: **a** Crystal structure of the RGD sequence from fibronectin bound to the αV (orange) and %3-subunit (gray) of the integrin receptor (PDB 4MMX). **b** Chemical structure of the cyclic pentapeptide cyclo (RGDf-N(Me)V) and crystal structures (gray/red, PDB 1L5G) (Marelli et al. 2014) superimposed with fibronectin RGD (gray; red = constraining amino acids; f = D-phenylalanine). (Color figure online)

to triphosphate (GTP) or guanosine diphosphate (GDP) (Bourne et al. 1990). When bound to GTP, they adopt an active conformation that facilitates binding to effector proteins thereby triggering downstream signaling events. The nucleotide binding state is regulated by PPI. Guanine nucleotide exchange factors (GEF) mediate a GDP to GTP exchange while GTPase-activating proteins (GAP) promote hydrolysis of bound GTP to GDP. Malfunctioning of GTPase regulation has implications in cancer formation and propagation. One good example is the proto-oncogene Ras which gives its name to a subfamily of related proteins such as Rab (Ras related in brain) and Rho (Ras homology) proteins (Spiegel et al. 2014a). Their targeting has proved extremely challenging because of the involvement of numerous PPI in small GTPase regulation and signal propagation (Spiegel et al. 2014b). The use of an HBS-stabilized a-helix derived from a GEF protein of Ras (Sos) is one successful example. The modified peptide HBS3 binds the GDP bound form of Ras with micro molar affinity and is capable of inhibiting the nucleotide exchange by Sos in vitro and in cell culture (Patgiri et al. 2011).

Hydrocarbon peptide stapling was used to stabilize an α-helix of the Rab6-interacting protein, an effector of Rab GTPases. Most strikingly class A mimetics i, i + 4 stapled peptide StRIP3 showed micro molar affinity for the active form of Rab8a and was able to compete with effector binding in vitro (Spiegel et al. 2014a). Hamilton and co-workers reported a class C mimetic based on a 5-6-5 imidazole-phenyltriazole scaffold to target Cdc42, a member of the Rho GTPase family. By mimicking three residues (Leu, Lys, Gln) of the GEF protein Dbs, the compound was able to inhibit the Dbs-promoted nucleotide exchange in vitro (IC50 = 67 mm) (Cummings et al. 2009). However, despite extensive efforts, clinically relevant compounds that directly target small GTPases remain elusive.

Transcriptional Regulation

Protein–protein interactions are key in transcriptional regulation pathways that include the NOTCH, Wnt, and Hedgehog signaling cascades. Impaired modulation of such pathways has strong implications in the genesis and progression of various types of cancer (Katoh 2007). Verdine and Bradner research groups designed peptidomimetics aimed at targeting transcription factor complexes (Moellering et al. 2009). They reported the development of hydrocarbon-stapled peptides for the inhibition of NOTCH signaling. The binding of protein ligands to NOTCH transmembrane receptors facilitates the activation of NOTCH target genes which triggers proteolytic cleavage of the intracellular domain of NOTCH (ICN) (Bray 2006). The ICN activates transcription by forming a trimeric complex with the coactivator proteins of the mastermind-like (MAML) family and DNA bound transcription factor CSL in the nucleus. On the basis of the α-helical binding domain of MAML, the i, i + 4 stapled peptide SAHM1 was developed. The peptide observed robust cellular uptake and potent inhibition of the trimer formation in vitro (Moellering et al. 2009). Cell-based assays confirmed the inhibition of NOTCH dependent gene expression. SAHM1 treatment observed specific antiproliferative effects in a mouse model of NOTCH driven T-cell acute lymphoblastic leukemia (Moellering et al. 2009).

Based on the a-helical b-catenin binding epitopes of Axin and BCL9 hydrocarbon-stapled peptides were used to target the Wnt signaling cascade (Hahne and Grossmann 2013). The Wnt signaling is activated by the binding of extracellular Wnt protein ligands to a receptor complex, which subsequently leads to intracellular inhibition of a multiprotein destruction complex consisting of scaffolding proteins such as protein kinases and Axin. In the absence of the Wnt ligand the complex facilitates the degradation of the protein %-catenin. The inhibition of the destruction complex in the presence of Wnt ligand triggers accumulation of %-catenin and its translocation into the nucleus. In the nucleus it binds to transcription factors of the LEF/TCF family and co-activators such as B-cell lymphoma 9 protein (BCL9) enabling the activation of transcription of the Wnt target genes (Katoh 2007).

The direct targeting of %-catenin has been a long standing goal (Hahne and Grossmann 2013). The i, i + 4 stapled peptides StAx-35R (Grossmann et al. 2012) and SAH-BCL9B (Takada et al. 2012) were developed. In cell-based assays the StAx-35R prevents the formation of a complex between LEF/TCF transcription factors and % catenin thereby inhibiting target genes under the control of Wnt signaling (Grossmann et al. 2012). The correct subcellular localization is essential for efficient inhibition of the signaling cascade (Cui et al. 2013). The SAH-BCL9B prevents the interaction between co-activator BCL9 and % catenin thereby inhibiting a subset of Wnt target genes that control stem-cell-like behavior in some forms of cancer. SAH-BCL9B reduces tumor growth, metathesis and invasion in mouse xenograft models (Takada et al. 2012).

Hydrocarbon- stapled peptides were also used to modulate other facets of gene expression. In histone methylation processes the complex between EED (embryonic ectoderm development and suppressor of zeste 12 homologue) and EZH2 (enhancer of zeste homologue 2) is vital. The complex formation was inhibited using EZH2-derived stapled peptides (Kim et al. 2013). Stapled peptides have also been used to target protein–protein complexes involved in the regulation of mRNA transcription (Lama et al. 2013) and DNA protection mechanisms (Frank et al. 2014). Estrogen receptors are activated by steroid hormones and regulated by co-activator proteins. The hyperactivation of these transcription factors has been implicated in the development

of cancer (Darnell 2002). Co-activator proteins bind the receptor through a nuclear receptor box (NR-box) consisting of a LXXLL motif. Upon binding the motif adopts an α-helical secondary structure (Fig. 5a). Attempts to stabilize the binding motif in its active conformation were done using disulfide (PERM-1, Fig. 5b), lactam, or thioether side chain to side chain cross-links (Galande et al. 2004). A lactam cross-linked peptide was further modified by incorporating unnatural amino acids, increasing the selectivity between receptor subtypes (Geistlinger and Guy 2003). Stapled peptides (i, i + 4) were designed using the crystal structure of nuclear receptor co-activator (NRCA) peptide 2 bound to ERα (Fig. 5a).

Structural studies revealed significant variations in the binding mode, affinity and selectivity. Notably, one of the crucial Leu amino acids was replaced by a building block in the formation of the marcocycle (Sp2; Fig. 5b). In this case, the hydrophobic cross-link is involved in the binding, thereby leaving the remaining residues of the stabilized peptide in good alignment with the wild-type peptide (Phillips et al. 2011). A structural mimetic was designed based on pyridylpyridone derivatives with substitutions in the 2-pyridyl and 1,5-pyridone positions (e.g. 44) to provide compounds that compete with the natural binding sequence in vitro (Hamilton and coworkers designed). The crystal structure aligns well with the Leu side chains of the helical LXXLL motif (Fig. 5c) (Becerril and Hamilton 2007). Another PPI with implications in the occurrence of cancer is the interaction between hypoxia-inducible transcription factors (HIFs) and p300/CBP coactivator proteins. HIFs are expressed under the cellular state of reduced oxygen levels.

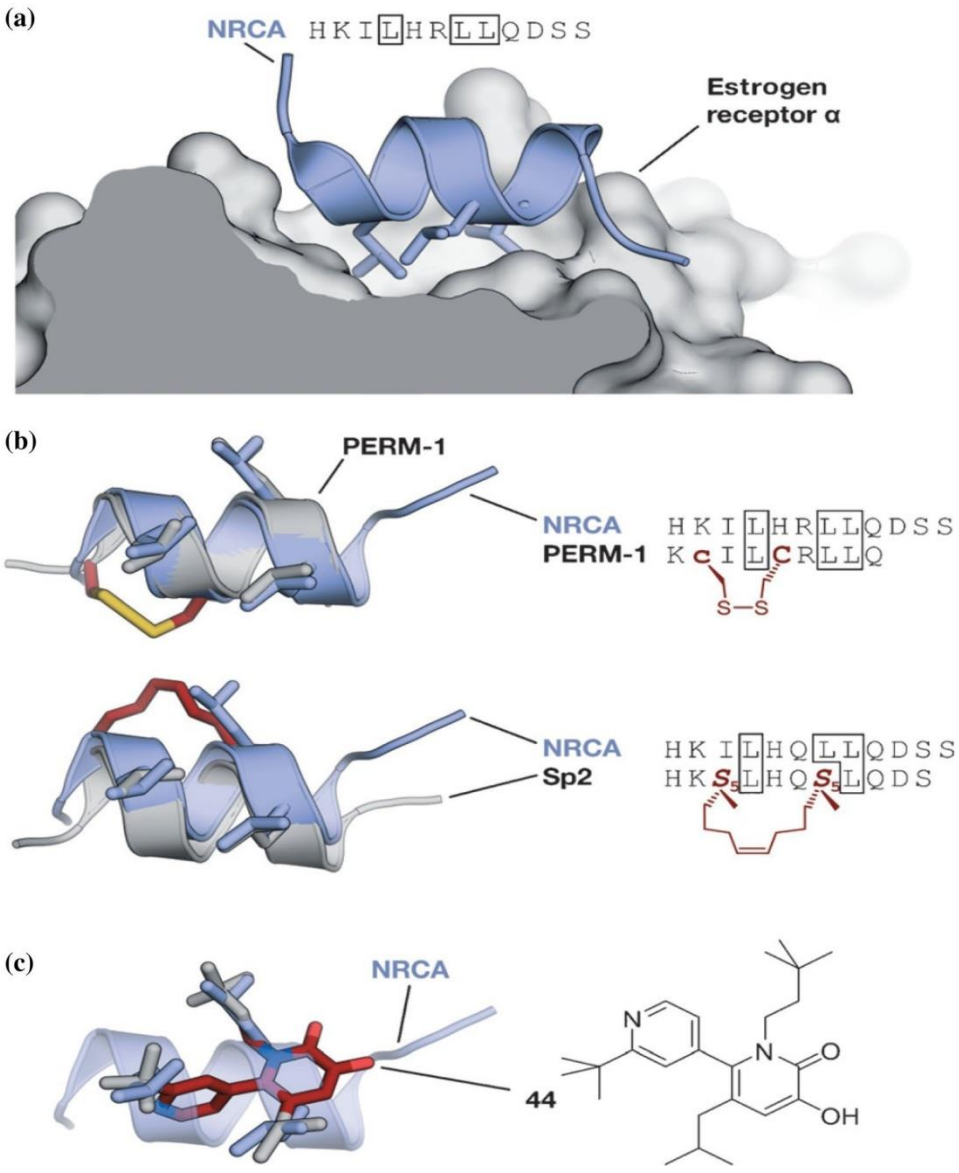


Fig. 5 Estrogen receptor (ER) coactivator interaction: **a** Coactivator peptide NRCA bound to ERa (gray; PDB 2QGT); **b** top: superimposed crystal structures of NRCA (blue, PDB 2QGT) and disulfide cross-linked PERM-1 (gray, PDB 1PCG; left). Cys and d-Cys **c** are highlighted in red, the disulfide bridge in yellow; sequences of cross linked peptide (right). Bottom: Superimposed crystal structures of NRCA (blue, PDB 2QGT) and stapled peptide Sp2 (gray, PDB 2YJA; left). The cross-link is highlighted in red. Sequences of stapled peptide (right). Selected side chains are shown explicitly and highlighted in sequence. **c** Superimposed crystal structures of NRCA (blue, PDB 2QGT) and 6-(2-tert-butyl-4-pyridyl)-3-hydroxy-5-isobutyl-1-(3,3-dimethylbutyl) 1H-pyridin-2-one (44, gray/red, CCDC: 636896) (Grossmann et al. 2015). (Color figure online)

In cancer cells, the interaction between HIFs and its co-activators can trigger the expression of genes that promote invasion, angiogenesis, and a modified metabolism (Hirota and Semenza 2006). The interaction between HIF-1α and p300/CBP is mediated by two short α helices in HIF-1α. Arora and co-workers designed a number of different peptidomimetics based on these peptide sequences. Initial efforts focused on hydrogen-bond surrogates to yield modified peptides inhibit complex formation (Henchey et al. 2010a, b; Kushal et al. 2013). Stabilized peptides showed inhibitory effects both in murine tumor xenografts and cancer-cell-based assays. Class C peptidomimetics observed potency in inhibiting the HIF-1ap300/CBP interaction (Lao et al. 2014a, b; Burslem et al. 2014). Aromatic oligoamides showed inhibitory effects in vitro (Burslem et al. 2014). Class C oligooxopiperazine helix mimetics (OHM, 37) compete with HIF-1α binding in vitro, reduces the expression of hypoxia-inducible genes in cell-based assays and is active in murine tumor xenografts. These results highlight the remarkable potential of α-helix mimetic based on oligooxopiperazine scaffolds (Lao et al. 2014a, b).

Conclusions and Future Perspectives

The discovery of PPI as potential drug targets for therapeutics has been an impressive journey to fame over the years. The continuously improving technology expertise in PPI-focused drug approach has brought the once intractable and undruggable approach on the spotlight as significant drug development strategy. The PPI-focused drug technology presents an emerging field for drug discovery and researchers have siphoned new technologies that have the potential to move this field further up the technology development curve and enable the regular discovery of PPI modulators. Peptidomimetics tend to mimic peptide side chains to take advantage of the binding affinity of a number of hot spot residues. The use of peptidomimetics has recently come of age with new drugs going into clinical trials. We envision that research in peptidomimetics will continue to be an indispensable tool to target PPI in drug discovery for the foreseeable future.

Acknowledgements Research reported in this manuscript was supported by the Department of Science and Technology (DST) South African, National Research Foundation (NRF) and University of Zululand. Its contents are solely the responsibility of the authors and do not necessarily represent the official views of the Department of Science and Technology (DST) South Africa, National Research Foundation (NRF) South Africa and University of Zululand.

Funding This research was funded by the Department of Science and Technology (DST) Innovation and Priority Research Areas Doctoral Scholarships programme, Reference No. SFH170530234909.

Compliance with Ethical standards

Conflict of Interest The authors declare no conflict of interest.

References

Adams JM, Cory S (2018) The BCL-2 arbiters of apoptosis and their growing role as cancer targets. Cell Death Differ 25(1):27–36

Adler MJ, Scott RTW, Hamilton AD (2012) Enaminone-based mimics of extended and hydrophilic α-helices. Chemistry 18:12974–12977

Angelo NG, Arora PS (2005) Nonpeptidic foldamers from amino acids: synthesis and characterization of 1,3-substituted triazole oligomers. J Am Chem Soc 127(49):17134–17135

Anil B, Riedinger C, Endicott JA, Noble ME (2013) The structure of an MDM2-Nutlin-3a complex solved by the use of a validated MDM2 surface-entropy reduction mutant. Acta Crystallogr Sect D 69:1358–1366

Ashkenazi A, Fairbrother WJ, Leverson JD, Souers AJ (2017) From basic apoptosis discoveries to advanced selective BCL-2 family inhibitors. Nat Rev Drug Discov 16(4):273–284

Azzarito V, Long K, Murphy NS, Wilson AJ (2013) Inhibition of α-helix-mediated protein–protein interactions using designed molecules. Nat Chem 5(3):161–173

Azzarito V, Miles JA, Fisher J, Edwards TA, Warriner S, Wilson A (2015) Stereocontrolled protein surface recognition using chiral oligoamide proteomimetic foldamers. Chem Sci 6:2434–2443

Baek S, Kutchukian PS, Verdine GL, Huber R, Holak TA, Lee KW, Popowicz GM (2012) Structure of the stapled p53 peptide bound to Mdm. J Am Chem Soc 134 2:103–106

Barnard A, Long K, Yeo DJ, Miles JA, Azzarito V, Burslem GM, Prabhakaran P, Edwards AT, Wilson AJ (2014) Orthogonal functionalisation of α-helix mimetics. Org Biomol Chem 12:6794–6799

Barnard A, Miles JA, Burslem GM, Barker AM, Wilson AJ (2015) Multivalent helix mimetics for PPI-inhibition. Org Biomol Chem 2015, 13:258–264

Bautista AD, Appelbaum JS, Craig CJ, Michel J, Schepartz A (2010) Bridged ℀3-peptide inhibitors of p53-hDM2 complexation–correlation between affinity and cell permeability. J Am Chem Soc 132(9):2904–2906

Becerril J, Hamilton AD (2007) Helix mimetics as inhibitors of the interaction of the estrogen receptor with coactivator peptides. Angew Chem Int Ed Engl 46:4471–4473; Angew. Chem 119:4555–4557

Bernal F, Tyler AF, Korsmeyer SJ, Walensky LD, Verdine GL (2007) Reactivation of the p53 tumor suppressor pathway by a stapled p53 peptide. J Am Chem Soc 129:2456–2457

Bernal F, Wade M, Godes M, Davis TN, Whitehead DG, Kung AL, Wahl GM, Walensky LD (2010) A stapled p53 helix overcomes HDMX-mediated suppression of p53. Cancer Cell 18:411–422

Biros SM, Moisan L, Mann E, Carella A, Zhai D, Reed JC, Rebek J Jr (2007) Heterocyclic α-helix mimetics for targeting protein–protein interactions. Bioorg Med Chem Lett 17:4641–4645

Boersma MD, Haase HS, Peterson-Kaufman KJ, Lee EF, Clarke OB, Colman PM, Smith BJ, Horne WS, Fairlie WD, Gellman SH (2012) Evaluation of diverse α/% backbone patterns for functional α-helix mimicry: analogues of the Bim BH3 domain. J Am Chem Soc 134:315–323

Bogan AA, Thorn KS (1998) Anatomy of hot spots in protein interfaces. J Mol Biol 280:1–9

Bonetta L (2010) Protein-protein interactions: interactome under construction. Nature 468:851–854

Bottger V, Bottger A, Howard SF, Picksley SM, Chene P, Garcia-Echeverria C, Hochkeppel HK, Lane DP (1996) Identification of novel mdm2 binding peptides by phage display. Oncogene 13:2141–2147

Bourne HR, Sanders DA, McCormick F (1990) The GTPase superfamily: a conserved switch for diverse cell functions. Nature 348:125–132

Bray SJ (2006) Notch signalling: a simple pathway becomes complex. Nat Rev Mol Cell Biol 7:678–689

Brown ZZ, Akula K, Arzumanyan A, Alleva J, Jackson M, Bichenkov E, Sheffield JB, Feitelson MA, Schafmeister CE (2012) A spiroligomer α-helix mimic that binds HDM2, penetrates human cells and stabilizes HDM2 in cell culture. PLoS ONE 7:e45948

Brown CJ, Quah ST, Jong J, Goh AM, Chiam PC, Khoo KH, Choong ML, Lee MA, Yurlova L, Zolghadr K, Joseph TL, Verma CS, Lane DP (2013) stapled peptides with improved potency and specificity that activate p53. ACS Chem Biol 8:506–512

Bullock BN, Jochim AL, Arora PS (2011) Assessing helical protein interfaces for inhibitor design. J Am Chem Soc 133(36):14220–14223

Bunnage ME (2011) Getting pharmaceutical R and D back on target. Nature Chemical Biology 7(6):335–339

Burslem GM, Kyle HF, Breeze AL, Edwards TA, Nelson A, Warriner SL, Wilson AJ (2014) Small-molecule proteomimetic inhibitors of the HIF-1α-p300 protein–protein interaction. ChemBioChem 15:1083–1087

Campbell F, Plante JP, Edwards TA, Warriner SL, Wilson AJ (2010) N-alkylated oligoamide alpha-helical proteomimetics. Org Biomol Chem 8:2344–2351

Cao X, Yap JL, Newell-Rogers MK, Peddaboina C, Jiang W, Papaconstantinou HT, Jupitor D, Rai A, Jung KY, Tubin RP, Yu W, Vanommeslaeghe K, Wilder PT, MacKerell AD, Fletcher S Jr, Smythe RW (2013) The novel BH3 α-helix mimetic JY-1-106 induces apoptosis in a subset of cancer cells (lung cancer, colon cancer and mesothelioma) by disrupting Bcl-xL and Mcl-1 protein-protein interactions with Bak. Mol Cancer 12:42

Chakrabarti P, Janin J (2002) Dissecting protein-protein recognition sites. Proteins 47:334–343

Chang YS, Graves B, Guerlavais V, Tovar C, Packman K, To KH, Olson KA, Kesavan K, Gangurde P, Mukherjee A, Baker T, Darlak K, Elkin C, Filipovic Z, Qureshi FZ, Cai H, Berry P, Feyfant E, Shi XE, Horstick J, Annis DA, Manning AM, Fotouhi N, Nash H, Vassilev LT, Sawyer TK (2013) Stapled α-helical peptide drug development: a potent dual inhibitor of MDM2 and MDMX for p53-dependent cancer therapy. Proc Natl Acad Sci USA 110:E3445–E3454

Cheloha RW, Maeda A, Dean T, Gardella TJ, Gellman SH (2014) Backbone modification of a polypeptide drug alters duration of action in vivo. Nat Biotechnol 32:653–655

Chen L, Yin H, Farooqi B, Sebti S, Hamilton AD. Chen J (2005) p53 alpha-Helix mimetics antagonize p53/MDM2 interaction and activate p53. Mol Cancer Ther 4:1019–1025

Chene P (2003) Inhibiting the p53-MDM2 interaction: an important target for cancer therapy. Nat Rev Cancer 3:102–109

Chene P (2006) Drugs targeting protein–protein interactions. Chem Med Chem 1:400–411

Chene P, Fuchs J, Bohn J, Garca-Echeverra C, Furet P, Fabbro D (2000) A small synthetic peptide, which inhibits the p53-hdm2 interaction, stimulates the p53 pathway in tumour cell lines. J Mol Biol 299:245–253

Clackson T, Wells JA (1995) A hot spot of binding energy in a hormone-receptor interface. Science 267:383–386

Conibear AC, Bochen A, Rosengren KJ, Stupar P, Wang C, Kessler H, Craik DJ (2014) The cyclic cystine ladder of theta-defensins as a stable, bifunctional scaffold: a proof-of-concept study using the integrin-binding RGD motif. ChemBioChem 15:451–459

Conte LL, Chothia C, Janin J (1999) The atomic structure of protein-protein recognition sites. J Mol Biol 285:2177–2198

Cui HK, Zhao B, Li Y, Guo Y, Hu H, Liu L, Chen YG (2013) Design of stapled α-helical peptides to specifically activate Wnt/%-catenin signaling. Cell Res 23:581–584

Cummings CG, Ross NT, Katt WP, Hamilton AD (2009) Synthesis and biological evaluation of a 5-6-5 imidazole-phenyl-thiazole based alpha-helix mimetic. Org Lett 11:25–28

Czabotar PE, Lee EF, van Delft MF, Day CL, Smith BJ, Huang DCS, Fairlie WD, Hinds MG, Colman PM (2007) Structural insights into the degradation of Mcl-1 induced by BH3 domains. Proc Natl Acad Sci USA 104:6217–6222

Darnell JE (2002) Transcription factors as targets for cancer therapy. Nat Rev Cancer 2:740–749

David A, Sternberg MJE (2015) The contribution of missense mutations in core and rim residues of protein–protein interfaces to human disease. J Mol Biol 427(17):2886–2898

Dechantsreiter MA, Planker E, Matha B, Lohof L, Lohof E, Hçlzemann G, Jonczyk A, Goodman SL, Kessler H (1999) N-Methylated cyclic RGD peptides as highly active and selective alpha(V)beta(3) integrin antagonists. J Med Chem 42:3033–3040

Del Gaizo Moore V, Letai A (2013) BH3 profiling–measuring integrated function of the mitochondrial apoptotic pathway to predict cell fate decisions. Cancer Lett 332:202–205

DeLano WL (2002) Unraveling hot spots in binding interfaces: progress and challenges. Curr Opin Struct Biol 12:14–20

Delbridge AR, Grabow S, Strasser A, Vaux DL (2016) Thirty years of BCL-2: translating cell death discoveries into novel cancer therapies. Nat Rev Cancer 16:99

Desgrosellier JS, Cheresh DA (2010) Integrins in cancer: biological implications and therapeutic opportunities. Nat Rev Cancer 10:9–22

Du L, Grigsby SM, Yao A, Chang Y, Johnson G, Sun H, Nikolovska-Coleska Z (2018) Peptidomimetics for targeting protein–protein interactions between DOT1L and MLL oncofusion proteins AF9 and ENL. ACS Med Chem Lett 9(9):895–900

Ernst JT, Becerril J, Park HS, Yin H, Hamilton AD (2003) Design and application of an alpha-helix-mimetic scaffold based on an oligoamide-foldamer strategy: antagonism of the Bak BH3/Bcl-xL complex. Angew Chem Int Ed Engl 42:535–539

Fasan R, Dias RLA, Moehle K, Zerbe O, Vrijbloed JW, Obrecht D, Robinson JA (2004) Using a beta-hairpin to mimic an alpha-helix: cyclic peptidomimetic inhibitors of the p53-HDM2 protein–protein interaction. Angew Chem Int Ed Engl 116:2161–2164

Fasan R, Dias RL, Moehle K, Zerbe O, Obrecht D, Mittl PR, Grutter MG, Robinson JA (2006) Structure-activity studies in a family of beta-hairpin protein epitope mimetic inhibitors of the p53-HDM2 protein–protein interaction. Chem Bio Chem 7:515–526

Frank AO, Vangamudi B, Feldkamp MD, Souza- Fagundes EM, Luzwick JW, Cortez D, Olejniczak ET, Waterson AG, Rossanese OW, Chazin WJ, Fesik SW (2014) Discovery of a potent stapled

helix peptide that binds to the 70N domain of replication protein A. J Med Chem 57:2455–2461

Furet P, Gay B, Caravatti G, Garca-Echeverra C, Rahuel J, Schoepfer J, Fretz H (1998) Structure-based design and synthesis of high affinity tripeptide ligands of the Grb2–SH2 domain. J Med Chem 41:3442–3449

Galande AK, Bramlett KS, Burris TP, Wittliff JL, Spatola AF (2004) Thioether side chain cyclization for helical peptide formation: inhibitors of estrogen receptor-coactivator interactions. J Pept Res 63:297–302

Gallo-Payet N, Guimond M, Bilodeau L, Wallinder C, Alterman M, Hallberg A (2011) Angiotensin II, a neuropeptide at the frontier between endocrinology and neuroscience: is there a link between the angiotensin ii type 2 receptor and alzheimer's disease? Front Endocrinol (Lausanne) 2:17

Garca-Echeverra C, Chene P, Blommers MJJ, Furet P (2000) Discovery of potent antagonists of the interaction between human double minute 2 and tumor suppressor p53. J Med Chem 43:3205–3208

Gavathiotis E, Suzuki M, Davis ML, Pitter K, Bird GH, Katz SG, Tu HC, Kim H, Cheng EH, Tjandra N, Walensky LD (2008) BAX activation is initiated at a novel interaction site. Nature 455:1076–1081

Geistlinger TR, Guy RK (2003) Novel selective inhibitors of the interaction of individual nuclear hormone receptors with a mutually shared steroid receptor coactivator 2. J Am Chem Soc 125:6852–6853

Giordanetto F, Revell JD, Knerr L, Hostettler M, Paunovic A, Priest C, Janefeldt A, Gill A (2013) Stapled vasoactive intestinal peptide (VIP) derivatives improve VPAC2 agonism and glucose-dependent insulin secretion. ACS Med Chem Lett 2013, 4:1163–1168

Gonzalez MW, Kann MG (2012) Chap. 4: Protein interactions and disease. PLoS Comput Biol 8:e1002819

Grabow S, Delbridge AR, Valente LJ, Strasser A (2014) MCL-1 but not BCL-XL is critical for the development and sustained expansion of thymic lymphoma in p53-deficient mice. Blood 124:3939–3946

Grasberger BL, Lu T, Schubert C, Parks DJ, Carver TE, Koblish HK, Cummings MD, LaFrance LV, Milkiewicz KL, Calvo RR et al (2005) Discovery and co-crystal structure of benzodiazepinedione HDM2 antagonists that activate p53 in cells. J Med Chem 2005 48:909–912

Grossmann TN, Yeh JTH, Bowman BR, Chu Q, Moellering RE, Verdine GL (2012) Inhibition of oncogenic Wnt signaling through direct targeting of %-catenin. Proc Natl Acad Sci USA 109:17942–17947

Grossmann TN, Pelay-Gimeno M, Glas A, Koch O (2015) Structure-Based design of inhibitors of protein–protein interactions: mimicking peptide binding epitopes. Angew Chem Int Ed Engl 54:8896–8927

Haase HS, Peterson-Kaufman KJ, Lan Levengood SK, Checco JW, Murphy WL, Gellman SH (2012) Extending foldamer design beyond α-helix mimicry: α/%-peptide inhibitors of vascular endothelial growth factor signaling. J Am Chem Soc 134:7652–7655

Hahne G, Grossmann TN (2013) Direct targeting of beta-catenin: Inhibition of protein-protein interactions for the inactivation of Wnt signalling. Bioorg Med Chem 21:4020–4026

Hanold LE, Oruganty K, Ton NT, Beedle AM, Kannan N, Kennedy EJ (2015) Inhibiting EGFR dimerization using triazolyl-bridged dimerization arm mimics. PLoS ONE 10(3):e0118796. https://doi.org/10.1371/journal.pone.0118796

Hara T, Durell SR, Myers MC, Appella DH (2006) Probing the structural requirements of peptoids that inhibit HDM2–p53 interactions. J Am Chem Soc 128:1995–2004

Haubner R, Finsinger D, Kessler H (1997) Stereoisomeric peptide libraries and peptidomimetics for designing selective inhibitors of the αv%3 integrin for a new cancer therapy. Angew Chem Int Ed Engl 109:36–1374; 1440–1456

Hayen A, Schmitt MA, Ngassa FN, Thomasson KA, Gellman SH (2004) Two helical conformations from a single foldamer backbone: "split personality" in short alpha/beta-peptides. Angew Chem Int Ed Engl 43:505–510; Angew Chem Int Ed Engl 116:511–516

Henchey LK, Jochim AL, Arora PS (2008) Contemporary strategies for the stabilization of peptides in the alpha-helical conformation. Curr Opin Chem Biol 12(6):692–697

Henchey LK, Porter JR, Ghosh I, Arora PS (2010a) High specificity in protein recognition by hydrogen bond surrogate α-helices: selective inhibition of the p53/MDM2 complex. ChemBioChem 11:2104–2107

Henchey LK, Kushal S, Dubey R, Chapman RN, Olenyuk BZ, Arora PS (2010b) High specificity in protein recognition by hydrogen bond surrogate α-helices: selective inhibition of the p53/MDM2 complex. J Am Chem Soc 132:941–943

Hintersteiner M, Kimmerlin T, Garavel G, Schindler T, Bauer R, Meisner NC, Seifert JM, Uhl V, Auer M (2009) A highly potent and cellularly active %-peptidic Inhibitor of the p53/hDM2 interaction. ChemBioChem 10:994–998

Hirota K, Semenza GL (2006) Regulation of angiogenesis by hypoxia-inducible factor 1. Crit Rev Oncol Hematol 59:15–26

Hirschmann R, Nicolaou KC, Pietranico S, Leahy EM, Salvino J, Arison B, Cichy MA, Spoors PG, Shakespeare WC, Sprengeler PA, Hamley P, Smith AB, Reisine T, Raynor K, Maechler L, Donaldson C, Vale W, Freidinger RM, Cascieri MA, Strader CD (1993) J Am Chem Soc 115:12550–12568

Horne WS, Gellman SH (2008) Foldamers with heterogeneous backbones. Acc Chem Res 41:1399–1408

Jochim AL, Arora PS (2010) Systematic analysis of helical protein interfaces reveals targets for synthetic inhibitors. ACS Chem Biol 5:919–923

Jones S, Thornton JM (1996) Principles of protein-protein interactions. Proc Natl Acad Sci USA 93:13–20

Jura N, Endres NF, Engel K, Deindl S, Das R, Lamers MH, Wemmer DE, Zhang X, Kuriyan J (2009) Mechanism for activation of the EGF receptor catalytic domain by the juxta membrane segment. Cell 137:1293–1307

Katoh M (2007) Networking of WNT, FGF, Notch, BMP, and Hedgehog signaling pathways during carcinogenesis. Stem Cell Rev 3:30–38

Kazi A, Sun J, Doi K, Sung SS, Takahashi Y, Yin H, Rodriguez JM, Becerril J, Berndt N, Hamilton AD, Wang HG, Sebti SM (2011) The BH3 α-helical mimic BH3-M6 disrupts Bcl-XL, Bcl-2, and MCL-1 protein-protein interactions with Bax, Bak, Bad, or Bim and induces apoptosis in a Bax- and Bim-dependent manner. J Biol Chem 286:9382–9392

Kim W, Bird GH, Neff T, Guo G, Kerenyi MA, Walensky LD, Orkin SH (2013) Targeted disruption of the EZH2-EED complex inhibits EZH2-dependent cancer. Nat Chem Biol 9:643–650

Kipps TJ, Stevenson FK, Wu CJ, Croce CM, Packham G, Wierda WG, O'Brien S, Gribben J, Rai K (2017) Chronic lymphocytic leukaemia. Nat Rev Dis Primers 3:16096

Kotschy A, Szlavik Z, Murray J, Davidson J, Maragno AL, Le Toumelin-Braizat G et al (2016) The MCL1 inhibitor S63845 is tolerable and effective in diverse cancer models. Nature 538(7626):477–482

Kritzer JA, Stephens OM, Guarracino DA, Reznik SK, Schepartz A (2005) %-Peptides as inhibitors of protein–protein interactions. Bioorg Med Chem 13:11–16

Kubbutat MHG, Jones SN, Vousden KH (1997) Regulation of p53 stability by Mdm2. Nature 387:299–303

Kushal S, Lao B, Henchey LK, Dubey R, Mesallati H, Traaseth NJ, Olenyuk BZ, Arora PS (2013) Protein domain mimetics as in vivo modulators of hypoxia-inducible factor signaling. Proc Natl Acad Sci USA 110:15602–15607

Kussie PH, Gorina S, Marechal V, Elenbaas B, Moreau J, Levine AJ, Pavletich NP (1996) Structure of the MDM2 oncoprotein bound to the p53 tumor suppressor transactivation domain. Science 274:948–953

Lama D, Quah ST, Verma CS, Lakshminarayanan R, Beuerman RW, Lane DP, Brown CJ (2013) Rational optimization of conformational effects induced by hydrocarbon staples in peptides and their binding interfaces. Sci Rep 3:3451

Lanning M, Fletcher S (2015) Multi-facial, non-peptidic α-helix mimetics. Biology 4:540–555

Lanning ME, Wilder PT, Bailey H, Drennen B, Cavalier M, Chen L, Yap JL, Raje M, Fletcher S (2015) Towards more drug-like proteomimetics: two-faced, synthetic α-helix mimetics based on a purine scaffold. Org Biomol Chem 13:8642–8646

Lao BB, Drew K, Guarracino DA, Brewer TF, Heindel DW, Bonneau R, Arora PS (2014a) Rational design of topographical helix mimics as potent inhibitors of protein–protein interactions. J Am Chem Soc 136:7877–7888

Lao BB, Grishagin I, Mesallati H, Brewer TF, Olenyuk BZ, Arora PS (2014b) In vivo modulation of hypoxia-inducible signaling by topographical helix mimetics. Proc Natl Acad Sci USA 111:7531–7536

Lau YH, de Andrade P, McKenzie GJ, Venkitaraman AR, Spring DR (2014a) Linear aliphatic dialkynes as alternative linkers for double-click stapling of p53-derived peptides. ChemBioChem 15:2680–2683

Lau YH, de Andrade P, Quah ST, Rossmann M, Laraia L, Skold N, Sum TJ, Rowling PJE, Joseph TL, Verma C, Hyvonen M, Itzhaki LS, Venkitaraman AR, Brown CJ, Lane DP, Spring DR (2014b) Functionalised staple linkages for modulating the cellular activity of stapled peptides. Chem Sci 5:1804–1809

Leber B, Geng F, Kale J, Andrews DW (2010) Drugs targeting Bcl-2 family members as an emerging strategy in cancer. Expert Rev Mol Med 12:e28

Lee EF, Czabotar PE, Smith BJ, Deshayes K, Zobel K, Colman PM, Fairlie WD (2007) Crystal structure of ABT-737 complexed with Bcl-xL: Implications for selectivity of antagonists of the Bcl-2 family. Cell Death Differ 14:1711–1713

Lee JH, Zhang Q, Jo S, Chai SC, Oh M, Im W, Lu H, Lim HS (2011) Novel pyrrolopyrimidine-based α-helix mimetics: cell-permeable inhibitors of protein–protein. Interactions J Am Chem Soc 133:676–679

Lee JH, Oh M, Kim HS, Lee H, Im W, Lim HS (2016) Converting one-face α-helix mimetics into amphiphilic α-helix mimetics as potent inhibitors of protein–protein interactions. ACSComb Sci 18:36–42

Lessene G, Czabotar PE, Sleebs BE, Zobel K, Lowes KN, Adams JM, Baell JB, Colman PM, Deshayes K, Fairbrother WJ, Flygare JA, Gibbons P, Kersten WJ, Kulasegaram S, Moss RM, Parisot JP, Smith BJ, Street IP, Yang H, Huang DC, Watson KG (2013) Structure-guided design of a selective BCL-X(L) inhibitor. Nat Chem Biol 9(6):390–397

Leverson JD, Phillips DC, Mitten MJ, Boghaert ER, Diaz D, Tahir SK, Belmont LD, Nimmer P, Xiao Y, Ma XM, Lowes KN, Kovar P, Chen J, Jin S, Smith M, Xue J, Zhang H, Oleksijew A, Magoc TJ, Vaidya KS, Albert DH, Tarrant JM, La N, Wang L, Tao ZF, Wendt MD, Sampath D, Rosenberg SH, Tse C, Huang DC, Fairbrother WJ, Elmore SW, Souers AJ (2015) Exploiting selective BCL-2 family inhibitors to dissect cell survival dependencies and define improved strategies for cancer therapy. Sci Transl Med 7(279):279ra40

Li X, Wang Z, Feng Y, Song T, Su P, Chen C, Chai G, Yang Y, Zhang Z (2014) Two-face, two-turn α-helix mimetics based on a cross-acridine scaffold: analogues of the Bim BH3 domain. ChemBioChem 15:1280–1285

Liu M, Pazgier M, Li C, Yuan W, Li C, Lu W (2010a) A left-handed solution to peptide inhibition of the p53–MDM2 interaction. Angew Chem Int Ed Engl 49:3649–3652; Angew. Chem. 122:3731–3734

Liu M, Li C, Pazgier M, Li C, Mao Y, Lv Y, Gu B, Wei G, Yuan W, Zhan C, Lu WY, Lu W (2010b) D-Peptide inhibitors of the p53–MDM2 interaction for targeted molecular therapy of malignant neoplasms. Proc Natl Acad Sci USA 107:14321–14326

Londregan AT, Piotrowski DW, Wei L (2016) Synthesis of pyridazine-based α-helix mimetics. ACSComb Sci 18:651–654

Long K, Edwards TA, Wilson AJ (2013) Microwave assisted solid phase synthesis of highly functionalized N-alkylated oligobenzamide α-helix mimetics. Bioorg Med Chem 21:4034–4040

Madden MM, Muppidi A, Li Z, Li X. Chen J, Lin Q (2011) Synthesis of cell-permeable stapled peptide dual inhibitors of the p53-Mdm2/Mdmx interactions via photoinduced cycloaddition. Bioorg Med Chem Lett 21:1472–1475

Marelli UK, Frank AO, Wahl B, La Pietra V, Novellino E, Marinelli L, Herdtweck E, Groll M, Kessler H (2014) Receptor-bound conformation of cilengitide better representedby its solution-state structure than the solid-state structure. ChemEur J 20:14201–14206

Marimganti S, Cheemala MN, Ahn JM (2009) Novel amphiphilic α-helix mimetics based on a bis-benzamide scaffold. Org Lett 11:4418–4421

McNulty JC, Thompson DA, Bolin KA, Wilken J, Barsh GS, Millhauser GL (2001) High-resolution NMR structure of the chemically-synthesized melanocortin receptor binding domain AGRP(87–132) of the agouti-related protein. Biochemistry 40:15520–15527

Meireles LMC, Mustata G (2011) Discovery of modulators of protein–protein interactions: current approaches and limitations. Curr Top Med Chem 11:248–257

Melagraki G, Ntougkos E, Rinotas V, Papaneophytou C, Leonis G, Mavromoustakos T, Kontopidis G, Douni E, Afantitis A, Kollias G (2017) Cheminformatics-aided discovery of small-molecule protein–protein interaction (PPI) dual inhibitors of tumor necrosis factor (TNF) and receptor activator of NF-B ligand (RANKL). PLoS Comput Biol 13:e1005372

Michel J, Harker EA, Tirado-Rives J, Jorgensen WL, Schepartz A (2009) In Silico Improvement of %3-peptide Inhibitors of p53•hDM2 and p53 Hdmx. J Am Chem Soc 131:6356–6357

Milroy LG, Grossmann TN, Hennig S, Brunsveld L, Ottmann C (2014) Modulators of protein–protein interactions. Chem Rev 114(9):4695–4748

Moellering RE, Cornejo M, Davis TN, Del Bianco C, Aster JC, Blacklow SC, Kung AL, Gilliland DG, Verdine GL, Bradner JE (2009) Direct inhibition of the NOTCH transcription factor complex. Nature 462:182–188

Moldoveanu T, Follis AV, Kriwacki RW, Green DR (2014) Many players in BCL-2 family affairs. Trends Biochem Sci 39:101–111

Moon H, Lee WS, Oh M, Lee H, Lee JH, Im W, Lim HS (2014) Design, solid-phase synthesis, and evaluation of a phenyl–piperazine–triazine scaffold as α-helix mimetics. ACS Comb Sci 16:695–701

Muppidi A, Wang Z, Li X, Chen J, Lin Q (2011) Achieving cell penetration with distance-matching cysteine cross-linkers: a facile route to cell-permeable peptide dual inhibitors of Mdm2/Mdmx. Chem Commun 47:9396–9398

Muppidi A, Doi K, Edwardraja S, Drake EJ, Gulick AM, Wang HG, Lin Q (2012) Rational design of proteolytically stable, cell-permeable peptide-based selective Mcl-1 inhibitors. J Am Chem Soc 134:14734–14737

Murray JK, Farooqi B, Sadowsky JD, Scalf M, Freund WA, Smith LM, Chen J, Gellman SH (2005) Efficient synthesis of a %-peptide combinatorial library with microwave irradiation. J Am Chem Soc 127:13271–13280

Orner BP, Ernst JT, Hamilton AD (2001) Toward proteomimetics: terphenyl derivatives as structural and functional mimics of extended regions of an α-helix. J Am Chem Soc 123:5382–5383

Parrondo R, de las Pozas A, Reiner T, Perez-Stable C (2013) ABT-737, a small molecule Bcl-2/Bcl-xL antagonist, increases antimitotic-mediated apoptosis in human prostate cancer cells. PeerJ 1:e144

Patgiri A, Yadav KK, Arora PS, Bar-Sagi D (2011) An orthosteric inhibitor of the Ras–Sos interaction. Nat Chem Biol 7:585–587

Patgiri A, Joy ST, Arora PS (2012) Nucleation effects in peptide foldamers. J Am Chem Soc 134:11495–11502

Pazgier M, Liu M, Zou G, Yuan W, Li C, Li C, Li L, Monbo J, Zella D, Tarasov SG, Lu W (2009) Structural basis for high-affinity peptide inhibition of p53 interactions with MDM2 and MDMX. Proc Natl Acad Sci USA 106:4665–4670

Phan J, Li Z, Kasprzak A, Li B, Sebti S, Guida W, Schçnbrunn E. Chen J (2010) Structure-based design of high affinity peptides inhibiting the interaction of p53 with MDM2 and MDM. J Biol Chem 285:2174–2183

Phillips C, Roberts LR, Schade M, Bazin R, Bent A, Davies NL, Moore R, Pannifer AD, Pickford AR, Prior SH, Read CM, Scott A, Brown DG, Xu B, Irving SL (2011) Design and structure of stapled peptides binding to estrogen receptors. J Am Chem Soc 133:9696–9699

Pierschbacher MD, Ruoslahti E (1984) Cell attachment activity of fibronectin can be duplicated by small synthetic fragments of the molecule. Nature 309:30–33

Plante JP, Burnley T, Malkova B, Webb ME, Warriner SL, Edwards TA, Wilson AJ (2009) Oligobenzamide proteomimetic inhibitors of the p53-hDM2 protein–protein interaction. Chem. Commun. 5091–5093

Popowicz GM, Czarna A, Rothweiler U, Szwagierczak A, Krajewski M, Weber L, Holak TA (2007) Molecular basis for the inhibition of p53 by Mdmx. Cell Cycle 6:2386–2392

Prabhakaran P, Barnard A, Murphy NS, Kilner CA, Edwards TA, Wilson AJ (2013) Aromatic oligoamide foldamers with a "Wet Edge" as inhibitors of the α-Helix-Mediated p53–hDM2 protein–protein interaction. Eur J Org Chem 2013:3504–3512

Raj M, Bullock BN, Arora PS (2013) Plucking the high hanging fruit: A systematic approach for targeting protein–protein interactions. Bioorg Med Chem 21:4051–4057

Ripka AS, Rich DH (1998) Peptidomimetic design. Curr Opin Chem Biol 2:441–452

Roberts AW, Seymour JF, Brown JR, Wierda WG, Kipps TJ, Khaw SL, Carney DA, He SZ, Huang DC, Xiong H, Cui Y, Busman TA, McKeegan EM, Krivoshik AP, Enschede SH, Humerickhouse R (2012) Substantial susceptibility of chronic lymphocytic leukemia to BCL2 inhibition: results of a phase I study of navitoclax in patients with relapsed or refractory disease. J Clin Oncol 10(5):488–496 30(

Robertson NS, Spring DR (2018) Using peptidomimetics and constrained peptides as valuable tools for inhibiting protein–protein interactions. Molecules 23:959

Robinson JA (2008) Beta-hairpin peptidomimetics: design, structures and biological activities. ACC Chem Res 41(10):1278–1288

Rodriguez JM, Nevola L, Ross NT, Lee G, Hamilton AD (2009) Synthetic inhibitors of extended helix-protein interactions based on a biphenyl 4,40-dicarboxamide scaffold. Chem Biochem 10:829–833

Ruiz-Gomez G, Tyndall JDA, Pfeiffer B, Abbenante G, Fairlie DP (2010) Update 1 of: over one hundred peptide-activated G protein-coupled receptors recognize ligands with turn structure. Chem Rev 110:PR1–PR41

Sadowsky JD, Murray JK, Tomita Y, Gellman SH (2007) Exploration of backbone space in foldamers containing alpha- and beta-amino acid residues: developing protease-resistant oligomers that bind tightly to the BH3-recognition cleft of Bcl-xL. ChemBioChem 8:903–916

Sakurai K, Schubert C, Kahne D (2006) Crystallographic analysis of an 8-mer p53 peptide analogue complexed with MDM2. J Am Chem Soc 128:11000–11001

Sawada T, Gellman SH (2011) Structural mimicry of the α-helix in aqueous solution with an isoatomic α/%/η-peptide backbone. J Am Chem Soc 133(19):7336–7339

Shamas-Din A, Kale J, Leber B, Andrews DW (2013) Mechanisms of action of Bcl-2 family proteins. Cold Spring Harb Perspect Biol 5:a008714

Shiheido H, Takashima H, Doi N, Yanagawa H (2011) mRNA Display selection of an optimized MDM2-binding peptide. PLoS ONE 6:e17898

Smith BJ, Lee EF, Checco JW, Evangelista M, Gellman SH, Fairlie WD (2013) Structure-guided rational design of α/%-peptide foldamers with high affinity for BCL-2 family prosurvival proteins. ChemBioChem 14:1564–1572

Souers AJ, Leverson JD, Boghaert ER, Ackler SL, Catron ND, Chen J et al (2013) ABT-199, a potent and selective BCL2 inhibitor, achieves antitumor activity while sparing platelets. Nat Med 19:202–208

Spiegel J, Mas-Moruno C, Kessler H, Lubell WD (2012) Cyclic azapeptide integrin ligand synthesis and biological activity. J Org Chem 77:5271–5278

Spiegel J, Cromm PM, Itzen A, Goody RS, Grossmann TN, Waldmann H (2014a) Direct targeting of Rab-GTPase-effector interactions. Angew Chem Int Ed Engl 53:2498–2503; Angew Chem Int Ed Engl 126:2531–2536

Spiegel J, Cromm PM, Zimmermann G, Grossmann TN, Waldmann H (2014b) Small-molecule modulation of Ras signaling. Nat Chem Biol 10:613–622

Stewart ML, Fire E, Keating AE, Walensky LD (2010) The MCL-1 BH3 helix is an exclusive MCL-1 inhibitor and apoptosis sensitizer. Nat Chem Biol 6:595–601

Stockwell BR (2011) The quest for the cure: the science and stories behind the next generation of medicines. Columbia University Press, New York

Suchanek M, Radzikowska A, Thiele C (2005) Photo-leucine and photo-methionine allow identification of protein-protein interactions in living cells. Nat Methods 2:261–268

Takada T, Zhu D, Bird GH, Sukhdeo K, Zhao JJ, Mani M, Lemieux M, Carrasco DE, Ryan J, Horst D, Fulciniti M, Munshi NC, Xu W, Kung AL, Shivdasani RA, Walensky LD, Carrasco DR (2012) Targeted disruption of the BCL9/%-catenin complex inhibits oncogenic Wnt signaling. Sci Transl Med 4:148ra117

Taylor IR, Dunyak BM, Komiyama T, Shao H, Ran X, Assimon VA, Kalyanaraman C, Rauch JN, Jacobson MP, Zuiderweg ERP et al (2018) High throughput screen for inhibitors of protein-protein interactions in a reconstituted heat shock protein 70 (Hsp70) complex. J Biol Chem 293:4014–4025

Thirumoorthy R, Holder JR, Bauzo RM, Richards NGJ, Edison AS, Haskell-Luevano C (2001) Novel agouti-related-pro-tein-based melanocortin-1 receptor antagonist. J Med Chem 44:4114–4124

Thompson S, Hamilton AD (2012) Amphiphilic α-helix mimetics based on a benzoylurea scaffold. Org Biomol Chem 10:5780–5782

Thompson S, Vallinayagam R, Adler MJ, Scott RTW, Hamilton AD (2012) Double-sided α-helix mimetics. Tetrahedron 68:4501–4505

Ting JP, Tung F, Antonysamy S, Wasserman S, Jones SB, Zhang FF, Espada A, Broughton H, Chalmers MJ, Woodman ME et al (2018) Utilization of peptide phage display to investigate

hotspots on IL-17A and what it means for drug discovery. PLoS ONE 13:e0190850

Toledo F, Krummel KA, Lee CJ, Liu CW, Rodewald LW, Tang M, Wahl GM (2006) Mouse mutants reveal that putative protein interaction sites in the p53 proline-rich domain are dispensable for tumor suppression. Cancer Cell 9:273–285

Tse C, Shoemaker AR, Adickes J, Anderson MG, Chen J, Jin S, Johnson EF, Marsh KC, Mitten MJ, Nimmer P et al (2008) ABT-263: A potent and orally bioavailable Bcl-2 family inhibitor. Cancer Res 68:3421–3428

Van Mileghem S, Egle B, Gilles P, Veryser C, Van Meervelt L, De Borggraeve WM (2017) Carbonylation as a novel method for the assembly of pyrazine based oligoamide alpha-helix mimetics. Org Biomol Chem 15:373–378

Vassilev LT, Vu BT, Graves B, Carvajal D, Podlaski F, Filipovic Z, Kong N, Kammlott U, Lukacs C, Klein C et al (2004) In vivo activation of the p53 pathway by small-molecule antagonists of MDM2. Science 303:844–848

Vogelstein B, Lane D, Levine AJ (2000) Surfing the p53 network. Nature 408:307–310

Wade M, Li YC, Wahl GM (2013) MDM2, MDMX and p53 in oncogenesis and cancer therapy. Nat Rev Cancer 2013 13:83–96

Walensky LD, Bird GH (2014) Hydrocarbon-stapled peptides: principles, practice, and progress. J Med Chem 57:6275–6288

Walensky LD, Kung AL, Escher I, Malia TJ, Barbuto S, Wright RD, Wagner G, Verdine GL, Korsmeyer SJ (2004) Activation of apoptosis in vivo by a hydrocarbon-stapled BH3 helix. Science 305(5689):1466–1470

Watkins AM, Arora PS (2014) Anatomy of %-strands at protein-protein interfaces. *ACS* Chem Biol 9(8):1747–1754

Wei SJ, Joseph T, Chee S, Li L, Yurlova L, Zolghadr K, Brown C, Lane D, Verma C, Ghadessy F (2013) Inhibition of nutlin-resistant HDM2 mutants by stapled peptides. PLoS ONE 8:e81068

Werder M, Hauser H, Abele S, Seebach D (1999) %-Peptides as inhibitors of small-intestinal cholesterol and fat absorption. Helv Chim Acta 82:1774–1783

Whitby LR, Ando Y, Setola V, Vogt PK (2011) Design, synthesis, and validation of a %-turn mimetic library targeting protein–protein and peptide–receptor interactions. J Am Chem Soc. 133(26):10184–10194

Whitby LR, Boger DL (2012) Comprehensive peptidomimetic libraries targeting protein–protein interactions. Acc Chem Res 45(10):1698–1709

Xiang Z, Luo H, Payton JE, Cain J, Ley TJ, Opferman JT et al (2010) Mcl1 haplo insufficiency protects mice from Myc-induced acute myeloid leukemia. J Clin Invest 120:2109–2118

Yap JL, Cao X, Vanommeslaeghe K, Jung KY, Peddaboina C, Wilder PT, Nan A, MacKerell AD, Smythe WR, Fletcher S (2012) Relaxation of the rigid backbone of an oligoamide-foldamer-based α-helix mimetic: Identification of potent Bcl-xL inhibitors. Org Biomol Chem 10:2928–2933

Yarden Y, Pines G (2012) The ERBB network: at last, cancer therapy meets systems biology. Nat Rev Cancer 12:553–563

Yin H, Lee GI, Sedey KA, Rodriguez JM, Wang HG, Sebti SM, Hamilton AD (2005) Terephthalamide derivatives as mimetics of helical peptides: disruption of the Bcl-x(L)/Bak interaction. J Am Chem Soc 127:5463–5468

Youle RJ, Strasser A (2008) The BCL-2 protein family: opposing activities that mediate cell death. Nat Rev Mol Cell Biol 9:47–59

Zaykov AN, Ball ZT (2011) A general synthesis of dirhodium metallopeptides as MDM2 ligands. Chem Commun 47:10927–10929

Zhan C, Zhao L, Wei X, Wu X, Chen X, Yuan W, Lu WY, Pazgier M, Lu W (2012) An ultrahigh affinity d-peptide antagonist of MDM2. J Med Chem 55:6237–6241

Zhang H, Guttikonda S, Roberts L, Uziel T, Semizarov D, Elmore SW et al (2011) Mcl-1 is critical for survival in a subgroup of non-small-cell lung cancer cell lines. Oncogene 30:1963–1968

Zhang G, Andersen J, Gerona-Navarro G (2018) Peptidomimetics targeting protein-protein interactions for therapeutic development. Protein Pept Lett 31:1076-1089. https://doi.org/10.2174/0929866526566181101100842

CHAPTER FOUR

A paper already published in *American Journal of Translational Research* 2019;11 (11):6702-6716
ISSN: 1943-8141/AJTR0099816

The Oncogenic Potential of Small Nuclear Ribonucleoprotein Polypeptide G: A Comprehensive and Perspective View

Lloyd Mabonga and Abidemi Paul Kappo

Preface − About the Manuscript

The manuscript gives a comprehensive and perspective overview on the oncogenic potential of the small nuclear ribonucleoprotein polypeptide G (SNRPG). It provides in-depth explorations on the milestones covered in identifying the critical roles played by SNRPG in the biogenesis of spliceosomal uridyl-rich small nuclear ribonucleoprotein particles (U snRNPs; U1, U2, U4 and U5), which are precursors of both the major and minor spliceosome. The script introduces the overlooked aspects of how splicing associated Sm proteins, particularly SNRPG, can be relevant therapeutic targets to unlock the mysteries behind finding permanent solutions against cancer. For a long time splicing associated Sm proteins have been sidelined, overlooked, underexplored and branded undruggable. Very little is known about these proteins. However, splicing proteins such as SNRPG are indispensable in mRNA processing and they have been implicated in different types of cancers. Therefore, Sm proteins remain central in tumorigenesis and tumor development. Thus, the manuscript aims to put the overlooked SNRPG protein on a limelight, and prompt scientists to reconsider the critical role it may play in PPI-focused drug discovery. Perhaps investigating the mundus operandi of the core-splicing protein might unveil new avenues in designing and developing new diagnostics and therapeutic drugs in cancer.

Am J Transl Res 2019;11(11):6702-6716
www.ajtr.org /ISSN:1943-8141/AJTR0099816

Review Article

The oncogenic potential of small nuclear ribonucleoprotein polypeptide G: a comprehensive and perspective view

Lloyd Mabonga*, Abidemi Paul Kappo*

*Molecular Biophysics and Structural Biology (MBSB) Group, Department of Biochemistry and Microbiology, University of Zululand, KwaDlangezwa 3886, South Africa. *Equal contributors.*

Received July 18, 2019; Accepted October 19, 2019; Epub November 15, 2019; Published November 30, 2019

Abstract: Small nuclear ribonucleoprotein polypeptide G (SNRPG), often referred to as Smith protein G (SmG), is an indispensable component in the biogenesis of spliceosomal uridyl-rich small nuclear ribonucleoprotein particles (U snRNPs; U1, U2, U4 and U5), which are precursors of both the major and minor spliceosome. SNRPG has attracted significant attention because of its implicated roles in tumorigenesis and tumor development. Suggestive evidence of its varying expression levels has been reported in different types of cancers, which include breast cancer, lung cancer, prostate cancer and colon cancer. The accumulating evidence suggests that the splicing machinery compo- nent plays a significant role in the initiation and progression of cancers. SNRPG has a wide interaction network, and its functions are predominantly mediated by protein-protein interactions (PPIs), making it a promising anti-cancer therapeutic target in PPI-focused drug technology. Understanding its roles in tumorigenesis and tumor developmentis an indispensable arsenal in the development of molecular-targeted therapies. Several antitumor drugs linked to splicing machinery components have been reported in different types of cancers and some have already entered the clinic. However, targeting SNRPG as a drug development tool has been an overlooked and underdeveloped strat- egy in cancer therapy. In this article, we present a comprehensive and perspective view on the oncogenic potential of SNRPG in PPI-focused drug discovery.

Keywords: DEAD-box helicase 20, mRNA splicing, RBBP6, SNRPG, protein-protein interactions, transforming acidic coiled-coil containing protein 1, tumorigenesis

Introduction

Protein-protein interactions (PPIs) are indis- pensable in normative cellular processes and are tremendously important mediators in the progression of many disease states [1-3]. More than 600,000 disease-relevant PPIs have so far been reported in the human interactome [4, 5], most of which remain elusive and underex- plored. Optimizing the integration of PPIs with conventional and targeted cytotoxic therapies may lead to greatly protracted remissions and even curative therapies for several diseases, including cancer [6-8]. Over the years PPIs have been regarded as prototypically "intractable" and "undruggable" owing to their highly dynam-ic and expansive interfacial areas [2, 3].However, owing to improving technology exper- tise, the advent of PPI-focused smart-drug technology presents a notable advance in dis- ease diagnostics and therapeutic studies [7-9]. PPIs have emerged as significant arsenals in the drug development armory and inhibiting PPIs using small molecules or peptides modu- lates biochemical pathways and has therapeu- tic significance [4, 7, 8, 10].

The emergence of PPI-focused drug technology has prompted scientists to consider targeting splicing machinery components as possible tar- gets in alleviating the existing cancer challeng- es [5-7]. The strategy ushered in a new dawn in the field of drug discovery. A few drugs arealready on the market and some potential drug-like candidates are in clinical trials [2, 3, 11-13]. Nevertheless, targeting Smith (Sm) proteins as PPI drug development tools has been an over- looked and underdeveloped strategy in cancer

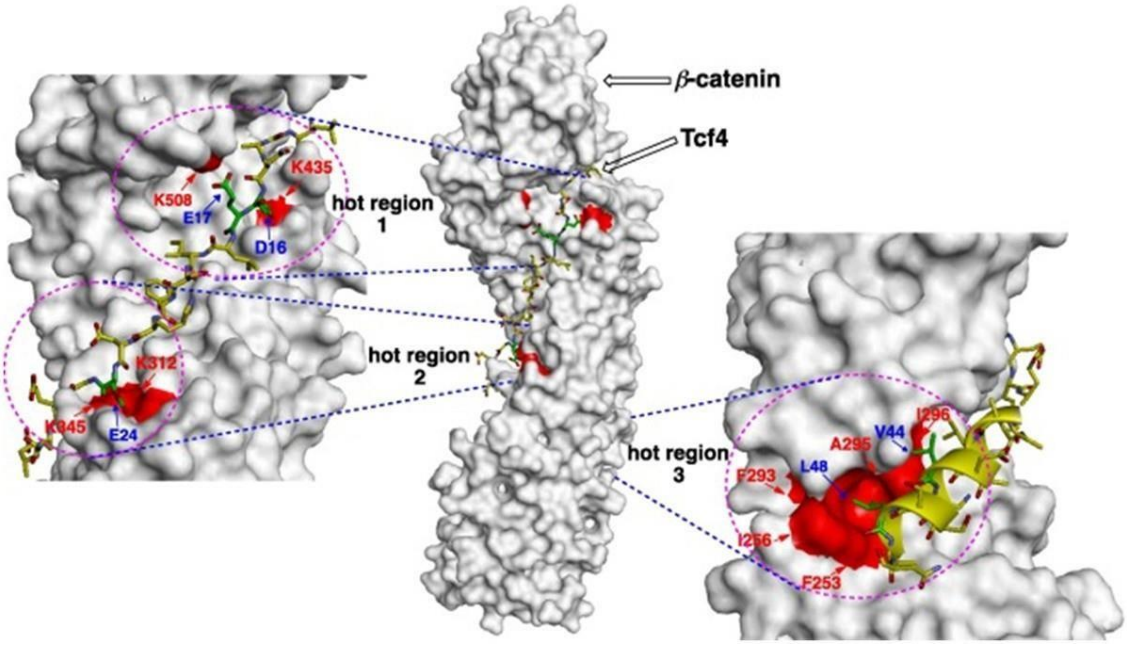


Figure 1. Cartoon representation of p-catenin/T-cell factor (Tcf) PPl interface showing three hot regions. Hot region 1 (K435 and K508 of p-catenin and D16 and E17 of Tcf4), hot region 2 (K312 and K345 of p-catenin and E24 and E29 of Tcf4) and hot region 3 (F253, l256, F293, A295, and l296 of p-catenin and V44 and L48 of Tcf4). The coop- erative contributions of hot spots within one hot region stabilize PPls (Figure taken from [31]).

therapy [7, 14]. Varying expression levels have been reported in different types of cancers, which include breast cancer, lung cancer, pros- tate cancer and colon cancer [15-20]. However, very little is known about their putative interac-tions in cancer-cell protein networks and their roles in different types of cancers.

Understanding their functional implications may lead to new avenues to design and devel- op PPI-focused therapeutic drugs in cancer [2, 3, 5-7, 21, 22]. In this article, we present a com- prehensive view and perspective on the onco- genic potential of SNRPG in PPI-focused drug technology.

PPI interfaces

PPIs occur over a relatively large interfacial area of approximately 1000 to 4000 Å2. The area is relatively prodigious in comparison to the mean contact area obligated for inhibition by small molecule inhibitors (300 to 1000 Å2) [23, 24]. The interfacial area of PPIs harbors incontrovertible hydrophobic regions called “hot spots”. These hydrophobic regions contrib- ute to the binding affinity and help to hold the two interacting proteins together [25, 26].

Typically, hot spot density on the protein-pro- tein interface composes 10% of the binding site residues. The amount of the structurally conserved residues (energetic hot spots) increases with the expansion of the interacting surface area [27].

Hot spot regions usually occur in clusters and within each cluster tightly packed hot spots form a network of conserved interactions called hot regions (shown in Figure 1) [28]. The coop- erative contributions of hot spots within one hot region stabilize PPIs. Hot regions are net- worked; their energetic contributions can be additive or cooperative and contribute domi- nantly to the stability of PPIs [29, 30]. Hot spot pockets for PPls are distinguishable from the other regions of the protein surface owing to their concave topology, combined with a pat- tern of hydrophobic and polar functionality. This combination of properties confers on concave hot regions a tendency to bind other proteins and small organic compounds possessing some polar functionality decorating a largely hydrophobic scaffold [31].

Hot spot regions are rich in hydrogen bonding and hydrophobic amino acids (Trp, Arg and Tyr),

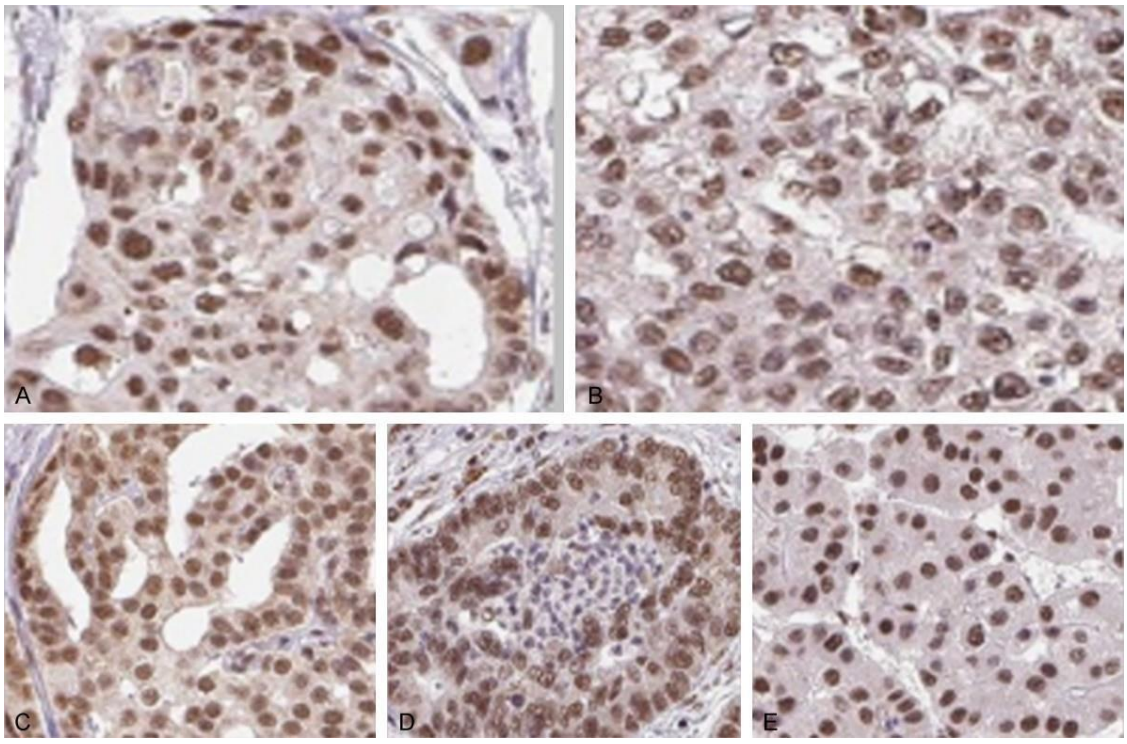


Figure 2. Antibody staining of five standard cancer tissues samples highlighting the localization of SNRPG in tumor cells. A. Colorectal Cancer. B. Breast Cancer. C. Prostate Cancer. D. Lung Cancer. E. Liver Cancer. Antibodies are labeled with DAB (3,3'-diaminobenzidine) and the resulting brown staining indicates where an antibody has bound to its corresponding antigen (SNRPG). Staining: Medium, Intensity: Moderate, Quantity: > 75%, Location: Nuclear, Magnification: 40 × (Figure taken from [18]).

which contribute to π-interactions and the bind- ing of free energy [32]. Systematic alanine scanning mutagenesis has revealed that exchanging amino acid residues for alanine in these hot spot regions reduces the binding affinity by at least 2 kcal/mol [32]. Further anal- ysis observed that hot spot regions comprise a core and a rim region. The rim region has an amino acid composition similar to the whole interfacial area, whereas the core region con- sists solely of aromatic residues [33, 34]. The core region fosters the a-helix, p-sheet and p-turn motifs, with the a-helix having a higher ratio in most of the secondary protein struc- tures. The a-helix actively binds into the grooves of binding partners and modulates the func- tioning of a large number of the disease-rele- vant PPIs [26, 35].

Comprehensive view

SNRPG is an approximately 8.5 kDa core-splic- ing and cancer-implicated Sm protein whose functions are predominantly mediated by PPIs [16, 36, 37]. The SNRPG protein coding gene is found on chromosome 2p13.3 and is made up of 8 exons. The gene comprises 455 nucleo- tides with an open reading frame encoding a predicted protein of 76 amino acid residues. An important paralog of this gene is LSM7. SNRPGprotein has a theoretical pI of 8.9 and is trans- lated *in vitro* from a single SNRPG mRNA that migrates as a doublet on high-TEMED SDS- PAGE [38]. The two bands represent conforma- tional isomers of the same protein. However, several transcript variants encoding different isoforms have been found for this gene. Northern blot analysis revealed that the SNRPG gene is expressed as an approximately 0.5-kb mRNA in HeLa cells [39].

SNRPG is a bona fide component of survival of motor neurons (SMN)-Sm protein complex, U1 snRNP, U2 snRNP, U12 type spliceosomal com- plex, U4 snRNP, U5 snRNP, spliceosomal tri-snRNP complex, catalytic step 2 spliceosome, Cytosol, methylosome, nucleoplasm, small nuclear ribonucleoprotein complex and spliceo- somal complex [21]. Among its related path- ways are the mRNA splicing-minor pathway and

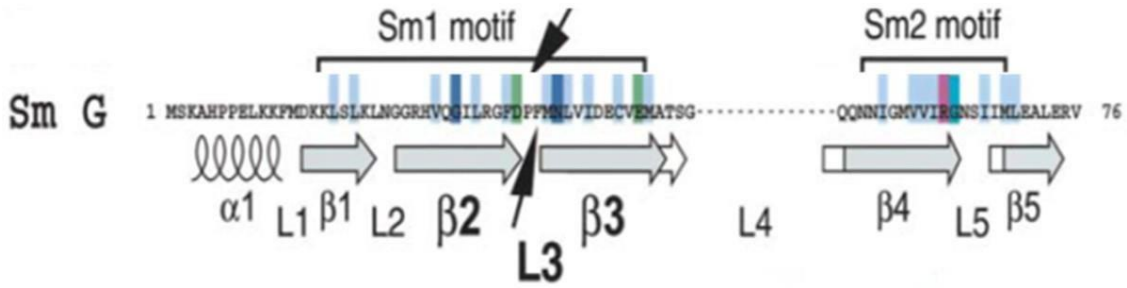


Figure 3. Human SNRPG protein primary structure alignment showing Sm1 and Sm2 motifs. Conserved aminoacids are highlighted as follows: Light blue (uncharged hydrophobic residues), green (acidic amino acids), purple (basic amino acids), dark blue (100% conserved amino acids) and turquoise (80% conserved glycine). Arrows mark the cross-linked amino acids in the protein sequences as identified by N-terminal sequencing, for example Phe37, Met38 and Asn39. The cross-linking sites are located within loop L3 of the Sm1 motif (Figure taken from [39]).

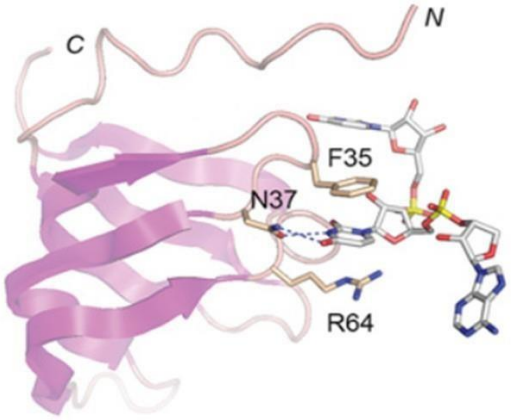


Figure 4. Stereo view of the human SNRPG (depictedas a cartoon trace with magenta p strands) and its interactions with the Sm site in U1 snRNA. Selected amino acids are shown as stick models and num- bered according to their positions in the SNRPG poly-peptide. Atomic contacts are indicated by dashed lines (Figure taken from [14]).

transport of the SLBP independent mature mRNA. The protein may also be a part of the U7 small nuclear ribonucleoprotein (U7 snRNP) complex, which participates in the processing of the 3' end of histone transcripts [21]. How- ever, it plays a yet uncharacterised role in link- ing core pre-mRNA splicing proteins to various cancers.

As shown in Figure 2, varying expression levels of SNRPG have been reported in different typesof cancers, which include colorectal cancer, breast cancer, lung cancer, prostate cancer and liver cancer [15-20]. According to Blijlevens and co-workers, increased expression levels of SNRPG protein in different types of cancers show a positive correlation with disease initiation, progression and severity [40]. The varying expression levels of SNRPG in different types of cancers may be explained by the overexpres- sion of the protein, the mislocalisation of unassembled protein or the mislocalisation of misassembled protein [41]. Thus, SNRPG may contribute significantly to the initiation and progression of cancers [14, 16, 37, 42-46].

SNRPG, like other Sm proteins, is characterisedby the presence of a conserved motif called the Sm motif. As shown in Figure 3, the Sm motif consists of two conserved regions that are separated by a non-conserved linker region, Sm1 and Sm2. The conserved motif comp- rises an antiparallel p sheet of p5‡•p1↓•p2‡• p3↓•p4‡ topology [39]. Several of the Sm sub- units are decorated by additional unstructured C terminal extensions and secondary stru- cture elements. The Sm motif encodes for a common folding domain (Sm domain) that is responsible for mediating PPIs between Sm proteins through the antiparallel p strands [47]. Moreover, SNRPG possesses two solvent- exposed hydrophobic interaction surfaces that are prone to nonspecific interactions under physiological conditions [47-52]. According to Stark and co-workers SNRPG has a wide inter- action network comprising more than 138 interactions with more than 115 identified interactors [21]. Its functions are mediated by both the specific and non-specific PPIs.

Prior to their involvement in the splicing cycle, SNRPG together with the other Sm proteins initially undergo translation in the cytoplasm and follow a hierarchical maturation pathway in which they interact independently of snRNA (shown in Figure 5) [53]. The activity is mediat-

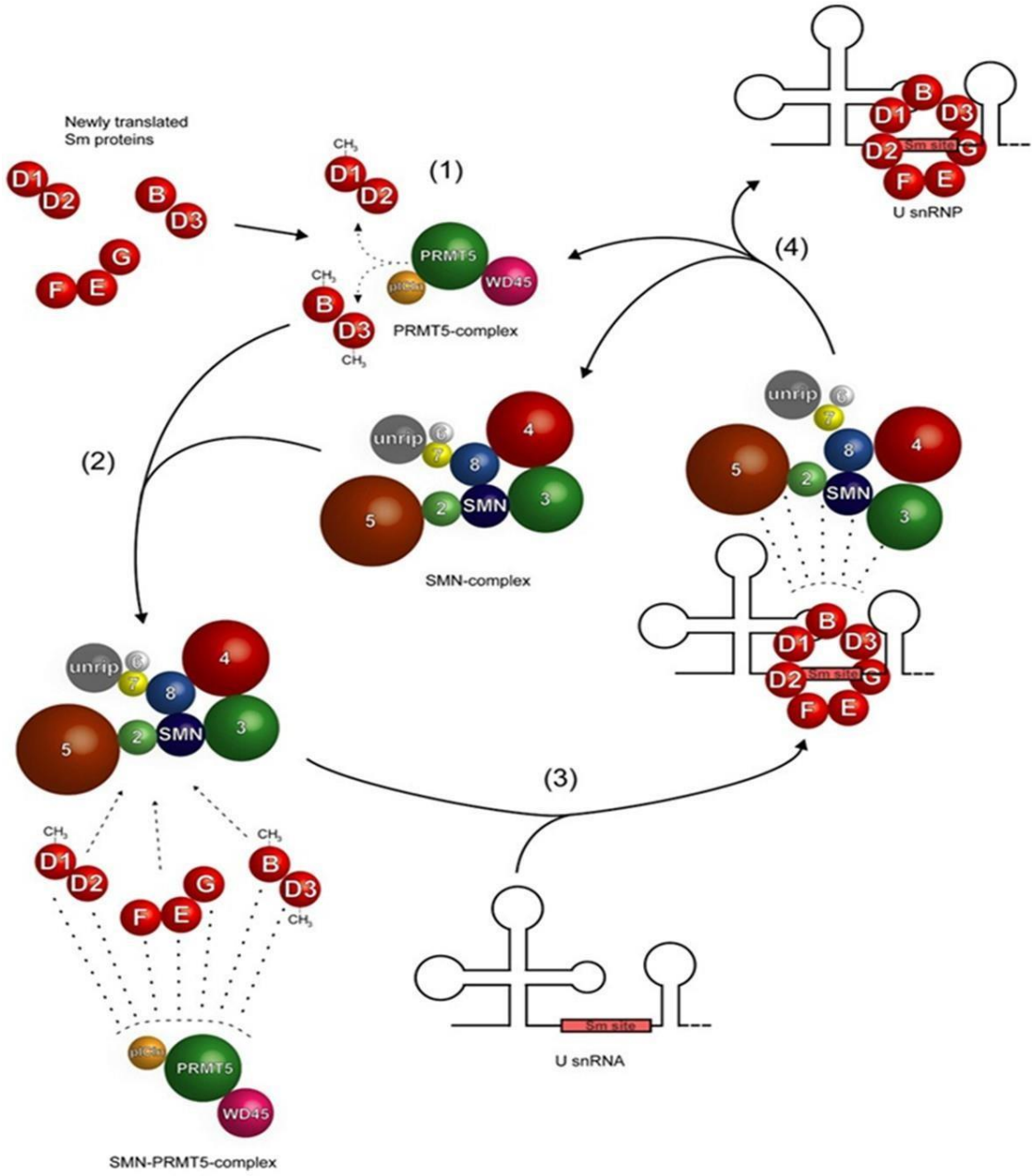


Figure 5. Model of assisted assembly of U snRNPs. Sm proteins are initially translated in the cytoplasm and se- questered by the PRMT5-complex, consisting of the Type ll methyltransferase PRMT5, WD45 (also termed Mep50) and pICln, which promotes symmetric dimethylation of arginines on Sm proteins B/B0, D1 and D3 (step 1). Next, the SMN-complex interacts with the PRMT5-complex to form an SMN-PRMT5-complex in which the Sm proteins are transferred onto the SMN-complex (step 2). These Sm proteins are assembled onto the "Sm-site" of U snRNAs to form U snRNPs (step 3). Finally, the U snRNP, the SMN-complex and PRMT5-complex dissociate and the latter two engage in a new round of U snRNP (Figure extracted from [53]).

ed predominantly by the assembly chaperone pICln, which inhibits the pre-mature binding of Sm proteins onto U snRNA and recruits all newly synthesized Sm proteins to the protein arginine methyltransferase 5 (PRMT5) complex forming three hetero-oligomers, D3/B, D1/D2 and E/F/G [22, 41, 52, 54]. The PRMT5-complex (comprising the Type II methyltransferase PRMT5, WD45 and pICln) promotes symmetric dimethylation of arginines on Sm pro-

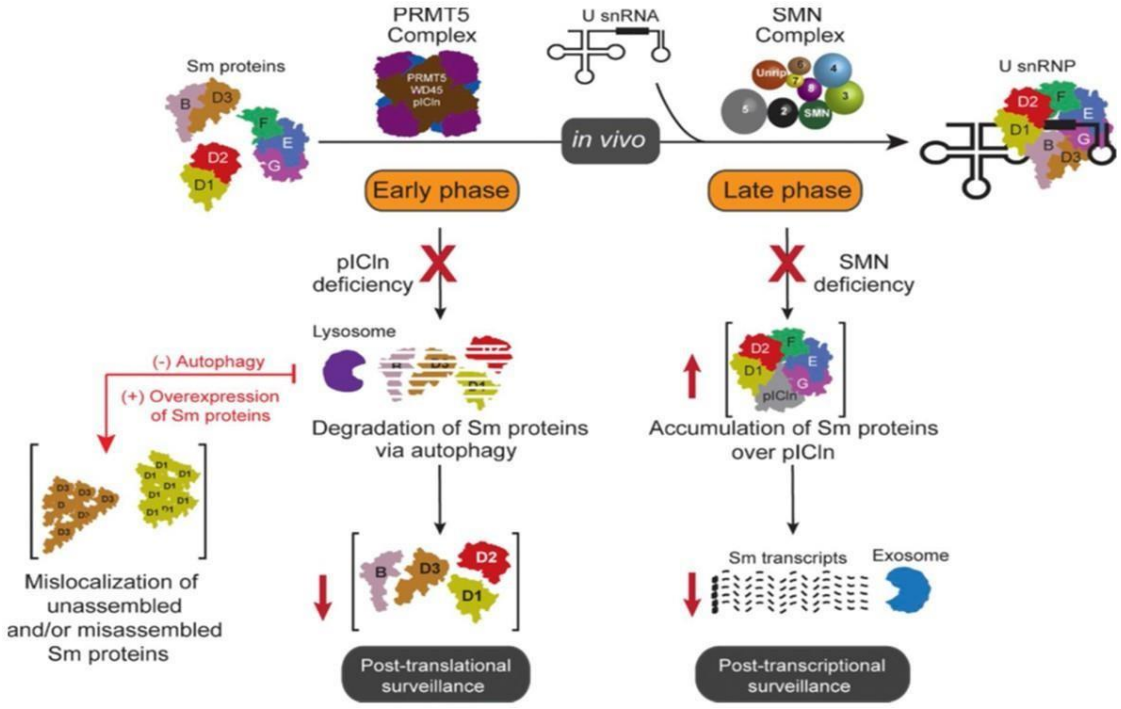


Figure 6. Schematic representation of the dysregulatory events in the homeostasis of U snRNPs. Dysregulation of Sm proteins during U snRNP assembly causes cellular proteotoxicity. Early phase plCln deficiency leads to degrada- tion of Sm proteins via autophagy and mislocalisation of unassembled and/or misassembled Sm proteins. Late phase SMN deficiency leads to accumulation of Sm proteins over plCln (Figure taken from [53]).

teins B/B', D1 and D3. The SMN-complex interacts with the PRMT5-complex facilitating the assembly of Sm proteins onto the "Sm-site" of U snRNAs forming the U snRNPs, as shown in Figure 4 [41]. After additional modification and processing steps, the U snRNP is targeted to its nuclear site of function, where it ultimately accumulates in interchromatin region struc- tures known as splicing speckles [22, 41, 53].

Very little is known about the manner in which the Sm proteins recognize and interact with the RNA-Sm site element. However, SNRPG has been highlighted to play a critical role in the direct recognition of the Sm site in the U snRNPs assembly [55]. According to Heinrichs and co-workers a direct contact between the SNRPG and the 5' part of the Sm site element within HeLa U1 snRNP particles was demon- strated by cross-linking approaches [55]. As indicated in Figure 2, the cross-linking sites are located within loop L3 of the Sm1 motif. The cross-linked amino acids are identified in the protein sequences as identified by N-terminal sequencing: Phe37, Met38 and Asn39 [55]. The cross-linking observed for the SNRPG is an outstanding feature and provides the first line of evidence that SNRPG plays a yet uncharac- terized and pivotal role in the functional activi- ties of Sm proteins.

In this context, cells are engineered to use a plethora of PPl networks to provide a therapeu- tically tractable way of tweaking and manipulat- ing the interplay of Sm proteins in order to maintain and address the normative cellular functions and progression of many disease states, including cancer [1, 56]. As shown in Figure 6, deficiency of plCln due to pathophysi- ological cues and disease progression has been linked to mislocalization of the unassem- bled and/or misassembled Sm proteins and their subsequent degradation via autophagy [41]. SMN deficiency has been linked to accu- mulation of Sm proteins over the pICln [41, 51, 57]. The reduced expression of functional SMN caused by genomic mutations has been linked to the debilitating human disorder spinal mus- cular atrophy [58].

In case of abnormality or impairment in the regulation of the assembly pathway, the cell

activates fail-safe measures, including target- ed autophagosome-mediated Sm protein degradation and exosome-processed Sm-encoding transcript degradation [41, 51, 57, 58]. These measures refine cellular quality control mechanisms to prevent proteotoxicity during imbalances in UsnRNP assembly [41]. Therefore, the regulation of Sm proteins during U snRNP assembly is tremendously important to prevent cellular proteotoxicity in disease progression [22, 41, 53].

Perspective view

Recent studies have shown significant evi- dence that deregulation of spliceosomal Sm proteins is linked to pathophysiological cues and disease states, such as cancer [40, 59]. The interference of Sm protein expression has been shown to induce apoptosis in non-small- cell lung cancer (NSCLC) cells [40]. According toBlijlevens and co-workers Sm proteins are fre- quently upregulated in NSCLC and their inc- reased expression shows positive correlation with disease severity [40]. Despite their inabili- ty to induce apoptosis in non-malignant cells, Sm proteins represent a particularly useful novel target for selective treatment of NSCLC [40]. However, their functional basis remains elusive and yet to be fully understood.

ln another study, Jin and co-workers investigat- ed the effects of silencing SNRPN expression on cell growth using the Daoy human medulloblastoma cell line *in vitro* [60]. The study observed that the knockdown of SNRPN markedly reduced the proliferation and colony-forming ability of Daoy medulloblastoma cells. The results indicate that SNRPN may be a potential novel target for the development of pharmacological therapeutics in human medulloblasto- ma [60]. Variable methylation of SNRPN has also been linked to germ cell tumors and acutemyeloid leukemia [61, 62]. SNRPN depletion inhibits the proliferation and colony formation of BxPC-3 pancreatic adenocarcinoma cells, leading to S phase cell cycle arrest and cell accumulation at the sub G1 phase [63]. However, the signalling pathway of SNRPN involved in the BxPC-3 cell proliferation and tumorigenesis remains unclear and yet to be fully elucidated. The results suggest that SNRPN may promote pancreatic adenocarcino- ma cell growth via regulation of the cell cycle and apoptosis, and lentivirus mediated SNRPN knockdown may be a potential therapeutic method for the treatment of pancreatic cancer [60].

Using semi-quantitative RT-PCR, Anchi and co-workers also reported the involvement of SNRPE in cell proliferation and progression of high-grade prostate cancer through the regula- tion of androgen receptor expression [64]. SNRPE overexpression promoted prostate can- cer cell proliferation in high-grade prostate can- cer cells compared with normal prostatic epi- thelial cells, indicating its oncogenic effects. lts knockdown expression by short interfering RNA (siRNA) resulted in the marked suppres- sion of prostate cancer cell proliferation [64]. Furthermore, the study observed that the regu- lation of androgen receptor expression bySNRPE is essential for cell proliferation and progression of high-grade prostate cancer.Thus, SNRPE may present a novel molecular target for cancer drugs [64].

SNRPG is one good example of proteins that have been implicated in cancer and whose functions are predominantly mediated by PPIs [16, 36, 37]. According to Johnson and co-work- ers most cancer-implicated proteins possess structural domains that have a higher ratio of infidelity compared to non-cancer-implicated proteins, making them more prone to interac- tion with a wide diversity of proteins [65]. Cancer-implicated proteins have a large num- ber of interacting proteins and occupy a central position in cancer-cell protein networks [66- 68]. PPIs between cancer-implicated proteins have a higher probability of being related to thecancer processes than non-interacting proteins [1, 56, 65, 69]. The accumulative and sugges- tive evidence of the varying expression levels inmore than 20 different types of cancers makes SNRPG a promising anti-cancer therapeutic tar- get in PPI-focused drug technology [16].

SNRPG and retinoblastoma binding protein 6 (RBBP6)

Putative PPIs between SNRPG and RBBP6 have been suggested. RBBP6 is a 250 kDa splicing-associated human protein initially known to bind to the retinoblastoma gene product, pRB [15, 17]. The RBBP6 gene is known to possess six different domains (shown schematically in Figure 7) that have been characterized and

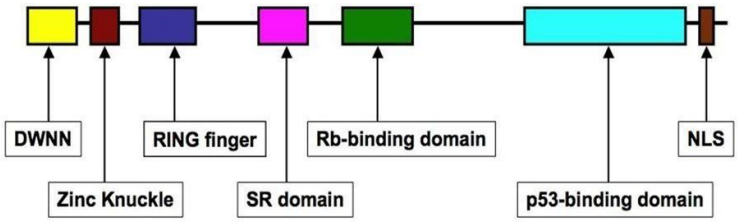


Figure 7. The domain organisation in human RBBP6. RBBP6 has three well-conserved N-terminal domains namely (i) "Domain with no name" (DWNN), (ii) Zinc knuckle and (iii) RlNG (really interesting new gene) finger domain, and three C-terminal domains, viz (i) proline-rich SR domain, (ii) Rb-binding domain and (iii) p53-binding domain (Figure taken from [76]).

linked with different types of cancer [15]. RBBP6 has three well-conserved N-terminal domains ("domain with no name" (DWNN), zinc knuckle and "really interesting new gene" (RlNG) finger domain) and three C-terminal domains (proline-rich SR domain, Rb-binding domain and p53-binding domain) [15, 70, 71]. Even though the functions of the N-terminal domains are not yet fully understood, it is well understood that RBBP6 is linked to tumorigen- esis and tumor development, and its functions are predominantly mediated by PPIs [17]. RBBP6 has been characterized and linked with14 different types of cancer at varying expres- sion levels [15, 19, 42, 44, 46, 71-75]. It inter- acts with the two-prototypical tumor suppres- sor proteins p53 and pRB [15, 17, 19].

Among other oncogenic functions, RBBP6 facil-itates interaction between p53 and its negative regulator, MDM2, leading to enhanced p53 ubiquitination and degradation [17, 19]. lt also interferes with the binding of p53 to DNA and facilitates the ubiquitination of pRb [17]. RBBP6 interacts directly with the pro-proliferative tran-scription factor Y-box-binding protein-1 (YB-1). Its overexpression in cultured mammalian cells leads to suppression of the anti-apoptotic YB-1 in a proteasome-dependent manner [71]. However, because it down-regulates both the pro-apoptotic p53 and the anti-apoptotic YB-1, the effect of RBBP6 on tumorigenesis is likely to be highly complex [45].

Accumulative evidence has shown that RBBP6 interacts with core splicing Sm proteins, SNRPB [15] and SNRPG [43, 45]. Using immunoblot analysis, Simons and co-workers observed that the N-terminal domain of RBBP6 interacts with Sm proteins [15]. The result points to a possi-ble connection between tu- mor suppressor proteins and the splicing machinery com-ponents. The putative effects of the N-terminal domain of RBBP6 on Sm proteins is an interesting finding that may catapult investigations to see whether tumor suppressor proteins can directly influence Sm proteins in pre-mRNA splicing. Using a yeast 2-hy-brid (Y2H) technique, Chibi and co-workers postulated that RBBP6 may perhaps interact with the core splicing SNRPG protein through its N-terminal domains, which is a crucial component of the RNA processing machinery in the cell [43]. These suggestions substantiate the possible involvement of RBBP6 in pathways linked to the pre-mRNA splicing machinery. However, the precise mechanisms involved remain elusiveand yet to be characterised.

Furthermore, Kappo and co-workers identified two copies of SNRPG (conformational isomers of the same protein) as part of the five sub- strates that bind to the N-terminal domain of RBBP6 [45]. The Y2H findings support the results that there might be a link between RBBP6 and the Sm proteins in the initiation andprogression of cancers [15, 43]. Considering the critical role played by SNRPG in the forma- tion of the hetero-oligomer E/F/G, Sm protein assembly and the subsequent assembly of Sm proteins onto the "Sm-site" of U snRNAs form- ing the U snRNPs, the results may suggest a possible strong link between SNRPG, pRb/p53 pathways and tumorigenesis [37]. Although many aspects of the above model remain to be proven and the mechanisms and functional basis of the links have yet to be fully under- stood, the findings have strong and interesting implications that prompt further investigations into the oncogenic potential of the core splicing SNRPG protein in the initiation and progressionof cancers.

The possible connection between SNRPG and the N-terminal domains of RBBP6 relates to features that suggest that these proteins may be the "forgotten link" connecting the cellular pre-mRNA splicing mechanism to tumorigene- sis and tumor development. First, the abundant

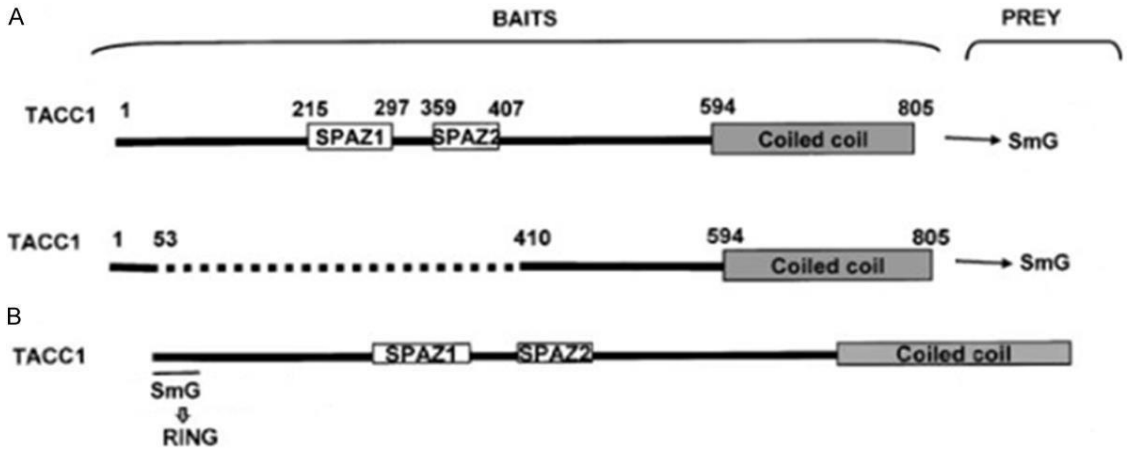


Figure 8. Mapping of TACC1/SNRPG interactions using the two-hybrid method in yeast. A. Schematic representationof TACC1 protein showing its three different regions: N-terminus, central serine/proline-rich region with two SPAZ motifs and coiled-coil C-terminus. Below are the different fragments generated as baits. B. Results of the mapping showing SNRPG binding to the N-terminus region of TACC1 in yeast two-hybrid experiments (Figure taken from [16]).

localization of SNRPG and RBBP6 in the nucleolus and nuclear speckles in tumor cells indicates a close connection between RBBP6 and several pre-mRNA splicing components [77]. Second, the association of RBBP6 with the Sm antigens in nuclear extracts, as shown by coimmunoprecipitation, suggests that RBBP6 is associated with these splicing factors in the liv-ing cells. Third, RBBP6 cDNA encodes an SR region that has biochemical properties similar to known SR splicing proteins, mediates the ubiquitination of p53 and pRb and interacts with SNRPB and SNRPG through its DWNN. It has been speculated that RBBP6 may link mRNA 3'-end processing to pRb/p53 pathways and tumorigenesis through its DWNN [37]. However, the question of how this happens remains unanswered.

The physiological relevance of the interactions and the functional basis of the association between SNRPG and DWNN still remain obscure. Very little is known about the putative PPIs between SNRPG and the N-terminal domains of RBBP6. The two proteins play an as yet uncharacterised role in linking splicing machinery components to tumorigenesis and tumor development in various cancers. Perhaps understanding the binding events between RBBP6's N-terminal domains and SNRPG may lead to new knowledge on how the two proteins relate in regulating splicing, tumorigenesis and tumor development. Inhibiting these PPIs may present a potential drug target in cancer diagnostics and therapeutic studies. Thus, new avenues to design and develop new therapeu- tic drugs may be established.

SNRPG and transforming acidic coiled coil containing protein 1 (TACC1)

The TACC1 gene is in region p12 of chromosome 8. lts mRNA is ubiquitously expressed and encodes a protein with an apparent molec- ular mass of 125 kDa [16]. The TACC1 protein's subcellular localization is within the cell cytoplasm and especially concentrated in the perinuclear area [16]. TACC1 was first identified asa potential oncogene; it is sometimes included in the amplification of the 8p12 region in breast cancers and can transform fibroblasts [78]. Based on the differential expression assay of chromosome 8p11-21 genes, researchers identified the TACC1 gene, whose mRNA is reduced or absent in breast carcinomas [79]. TACC1 mRNA gene expression is downregulat- ed in various types of tumors. Using immuno- histochemistry of tumor tissue-microarrays and sections, the level of TACC1 protein is down- regulated in breast cancer [16]. Furthermore, using the two-hybrid screen in yeast, GST pull- downs and co-immunoprecipitations, Conteand co-workers identified SNRPG as one of the two potential binding partners for TACC1 in breast cancer (shown in Figure 8) [16]. The findings suggested that TACC1 might play a role in the control of mRNA metabolism. Thus, Conte and co-workers speculated that down-regula-

tion of TACC1 may alter the control of mRNA homeostasis in polarized cells and subse- quently participate in the oncogenic processes [16].

To delineate the region of TACC1 that interacts with SNRPG, TACC1 fragment encoded proteins fused to the LEX binding domain were used as baits in two-hybrid assays against the SNRPG prey [80]. As shown in Figure 8, the results observed that the binding region of SNRPG on TACC1 is thus restricted to the N-terminus region of TACC1. GST pull-down and co-immu- noprecipitation experiments confirmed that the N-terminus region of TACC1 is indeed the site of binding for the Sm protein [80]. Inhibiting the interaction between TACC1 and SNRPG, using small molecules or peptides may modulate cancer-cell networks and have therapeutic sig- nificance. Hence investigating the binding ev- ents in the interactions between SNRPG and TACC1, and their relations in regulating splicing, tumorigenesis and tumor development may help identify new avenues to design and devel- op PPI-focused therapeutic drugs. Currently, the physiological relevance of the interactions and the functional basis for their association remain elusive and uncharacterized.

SNRPG and DEAD-box helicase 20 (DDX20)

DEAD-box helicase 20 (DDX20), commonly known as gem-associated protein 3 (Gemin3), is an ATP-dependent enzyme in humans that is encoded by the DDX20 gene [81, 82]. It is a component of the SMN complex that is tremen-dously important in the assembly and recon-struction of different Sm protein complexes [83]. Cleavage of the DDX20 by the poliovirus-encoded proteinase 2Apro has been shown to result in DDX20 inactivation and reducedsnRNP assembly [84]. DDX20 may act as a tumor suppressor in hepatocellular carcinoma and as a tumor promoter in breast cancer [85].According to Chen and co-workers, DDx20 defi- ciency enhances NF-κB activity by impairing theNF-κB-suppressive action of microRNAs. The findings suggest that dysregulation of the microRNA machinery components may also be involved in pathogenesis in various human dis- eases such as cancer [85].

One good example is miRNA-140, which acts asa liver tumor suppressor. Deficiency of DDx20 leads to the impairment of miRNA-140 func-tion. Functional impairment of miRNAs has been linked to hepatocarcinogenesis [85]. Similarly, DDX20 may promote the progression of prostate cancer through the NF-κB pathway [85]. Clinical investigations by Shin and co- workers found that a positive DP103/NF-κB feedback loop promotes constitutive NF-κB activation in invasive breast cancers [86]. The activation of this pathway is linked to cancer progression and the acquisition of chemothera- py resistance. It implies that DP103 has poten- tial as a therapeutic target for breast cancer treatment [86].

DEAD box proteins have been found to be involved in many aspects of RNA metabolism, including Sm-Sm protein interactions, pre-mRNA splicing, mRNA transport, mRNA degra-dation and translation in eukaryotes and pro-karyotes [87-92]. Using a biochemical appro- ach, Charroux and co-workers observed that anti-DDX20 mAbs immunoprecipitated the spli-ceosomal RNPG protein, as well as several other unidentified Sm proteins. Gemin3 inter- acts directly with Sm core proteins, including B/B', D2, and D3 [93]. In addition, DDX20 is uniformly distributed in the cytoplasm, where U snRNP assembly takes place, and it can bespecifically co-immunoprecipitated with the cytoplasmic pool of Sm proteins [93]. Taken together, these findings suggest that DDx20 and SNRPG may play an important role linking the spliceosomal snRNP biogenesis to tumori- genesis and tumor development. Finding small- molecule or peptide inhibitors for the interac- tion between DDX20 and SNRPG may help modulate cancer-cell networks and open up other avenues for the designing and develop- ment of new PPI-focused smart drugs.

Conclusions

Despite the strong and interesting implications associated with SNRPG and its significant prowess as a potential smart-drug discoverytarget in PPI-focused diagnostics and thera- peutic studies [14], the oncogenic potential of SNRPG remains to be proven. The mechanisms and functional basis of its operations in link-ing the splicing machinery to tumorigenesis and tumor development remain elusive and yet to be fully investigated. The findings pre- sented in this study prompt further investiga- tions. However, it is noteworthy that the foun-

dational basis of the views here presented in this article is based solely on the questiona-ble Y2H technique. Despite the popularity, rela- tive methodical simplicity, diversity and high-throughput capacity, as well as screening meth-od for interactomics, Y2H techniques face the problem of false positives [94]. False positiv-es in Y2H are physical interactions detected in the screening in yeast that are not reproducible in an independent system. A list of recurrent false positives exists and often depends on the Y2H system used [94]. Nonetheless, in this study we confirm that there is no data so far reported to prove and support that the Y2H results presented in line with SNRPG are false positives.

Acknowledgements

Research reported in this article was sup- ported by the South African National Research Foundation (NRF) through funding received via a Thuthuka Grant awarded to Abidemi Paul Kappo (Grant No: 107262) and a Doctoral Bursary received by Lloyd Mabonga from the South African Department of Science and Technology (DST). Its contents are solely the responsibility of the authors and do not neces- sarily represent the official views of the South African DST or the NRF.

Disclosure of conflict of interest

None.

Address correspondence to: Abidemi Paul Kappo, Molecular Biophysics and Structural Biology (MBSB) Group, Department of Biochemistry and Microbio-logy, University of Zululand, KwaDlangezwa 3886, South Africa. Tel: +27-35-902-6780; +27-72-084-3110; Fax: +27-35-902-6567; E-mail: KappoA@uni-zulu.ac.za

References

[1] Du L, Grigsby SM, Yao A, Chang Y, Johnson G, Sun H and Nikolovska-Coleska Z. Peptidomi-metics for targeting protein-protein interac- tions between DOT1L and MLL oncofusion pro- teins AF9 and ENL. ACS Med Chem Lett 2018; 9: 895-900.

[2] Robertson NS and Spring DR. Using peptidomi-metics and constrained peptides as valuable tools for inhibiting protein-protein interactions. Molecules 2018; 23.

[3] Zhang G, Andersen J and Gerona-Navarro G. Peptidomimetics targeting protein-protein in-teractions for therapeutic development. Pro- tein Pept Lett 2018; 25: 1076-1089.

[4] Gonzalez MW and Kann MG. Chapter 4: pro-tein interactions and disease. PLoS Comput Biol 2012; 8: e1002819.

[5] Díaz-Eufracio BI, JesúsNaveja J and Medina-Franco JL. Protein-protein interaction modula-tors for epigenetic therapies. Adv Protein Chem Struct Biol 2018; 110: 65-84.

[6] Adams JM and Cory S. The Bcl-2 apoptotic switch in cancer development and therapy. On-cogene 2007; 26: 1324-1337.

[7] Kale J, Osterlund EJ and Andrews DW. BCL-2 family proteins: changing partners in the dance towards death. Cell Death Differ 2018; 25: 65-80.

[8] Taylor lR, Dunyak BM, Komiyama T, Shao H, Ran X, Assimon VA, Kalyanaraman C, Rauch JN, Jacobson MP, Zuiderweg ERP and Gest-wicki JE. High throughput screen for inhibitors of protein-protein interactions in a reconstitut- ed heat shock protein 70 (Hsp70) complex. J Biol Chem 2018; 293: 4014-4025.

[9] Pelay-Gimeno M, Glas A, Koch O and Gross-mann TN. Structure-based design of inhibitors of protein-protein interactions: mimicking pep- tide binding epitopes. Angew Chem Int Ed Engl2015; 54: 8896-8927.

[10] Keskin 0, Gursoy A and Ma B. Principles of protein-protein interactions: what are the pre-ferred ways for proteins to interact? Chem Rev 2008; 108: 1225-44.

[11] Wishart DS, Knox C, Guo AC, Cheng D, Shrivas-tava S, Tzur D, Gautam B and Hassanali M. A knowledgebase for drugs, drug actions and drug targets. Nucleic Acids Res 2008; 36: D901-D906.

[12] Fuller JC, Burgoyne NJ and Jackson RM. Pre-dicting druggable binding sites at the protein-protein interface. Drug Discov Today 2009; 14: 155-161.

[13] xu GG, Guo J and Wu Y. Chemokine receptor ccr5 antagonist maraviroc: medicinal chemis- try and clinical applications. Curr Top Med Chem 2014; 14: 1504-1514.

[14] Schwer B, Kruchten J and Shuman S. Struc- ture-function analysis and genetic interactions of the SmG, SmE, and SmF subunits of the yeast Sm protein ring. RNA 2016; 22: 1320-8.

[15] Simons A, Melamed-Bessudo C, Wolkowicz R, Sperling J, Sperling R, Eisenbach L and Rotter V. PACT: cloning and characterization of a cel-lular p53-binding protein that interacts with Rb. Oncogene 1997: 14: 145-155.

[16] Conte N, Charafe-Jauffret E, Delaval B, Adéla-ïde J, Ginestier C, Geneix J, lsnardon D, Jacque-mier J and Birnbaum D. Carcinogenesis and translational controls: TACC1 is down-regulat-ed in human cancers and associates with

mRNA regulators. Oncogene 2002; 21: 5619-5630.

[17] Li L, Deng B and Xing G. PACT is a negative regulator of p53 and essential for cell growth and embryonic development. Proc Natl Acad Sci U S A 2007; 104: 7951-7956.

[18] Ezkurdia l, Juan D and Rodriguez JM. Multiple evidence strands suggest that there may be asfew as 19 000 human protein-coding genes. Hum Mol Genet 2014; 23: 5866-5878.

[19] Khan F, Allam M, Tincho MB and Pretorius A. Implications of RBBP6 in various types of cancer. Proceedings IWBBIO 2014. Conference Paper Granada; 2014. pp. 7-9.

[20] Hull R, Oosthuysen B and Cajee U. The drosophila retinoblastoma binding protein 6 fami- ly member has two isoforms and is potentially involved in embryonic patterning. Int J Mol Sci 2015: 16: 10242-10266.

[21] Stark C, Breitkreutz BJ, Reguly T, Boucher L, Breitkreutz A and Tyers M. BioGRlD: a general repository for interaction datasets. Nucleic Acids Res 2006; 34: D535-9.

[22] Ohtani M. Plant snRNP biogenesis: a perspective from the nucleolus and cajal bodies. Front Plant Sci 2018; 8: 2184.

[23] Jones S and Thornton JM. Principles of protein-protein interactions. Proc Natl Acad Sci U S A 1996; 93: 13-20.

[24] Conte LL, Chothia C and Janin J. The atomic structure of protein-protein recognition sites. J Mol Biol 1999; 285: 2177-2198.

[25] Clackson T and Wells JA. A hot spot of binding energy in a hormone-receptor interface. Sci- ence 1995; 267: 383-386.

[26] Jochim AL and Arora PS. Systematic analysis of helical protein interfaces reveals targets for synthetic inhibitors. ACS Chem Biol 2010; 5: 919-923.

[27] Carbonell P, Nussinov R and Del Sol A. Energetic determinants of protein binding specifici-ty: insights into protein interaction networks. Proteomics 2009; 9: 1744-1753.

[28] Keskin 0, Ma B and Nussinov R. Hot regions in protein-protein interactions: the organization and contribution of structurally conserved hot spot residues. J Mol Biol 2005; 345: 1281- 1294.

[29] Reichmann D, Rahat 0, Albeck S, Meged R, Dym O and Schreiber G. The modular architec- ture of protein-protein binding interfaces. Proc Natl Acad Sci U S A 2005; 102: 57-62.

[30] Moza B, Buonpane RA, Zhu P, Herfst CA, Rahman AK, McCormick JK, Kranz DM and Sundberg EJ. Long-range cooperative binding ef- fects in a T cell receptor variable domain. Proc Natl Acad Sci U S A 2006; 103: 9867-9872.

[31] Guo W, Wisniewski JA and Ji H. Hot spot-based design of small-molecule inhibitors for protein-protein interactions. Bioorg Med Chem Lett 2014; 24: 2546-2554.

[32] Bogan AA and Thorn KS. Anatomy of hot spots in protein interfaces. J Mol Biol 1998; 280: 1-9.

[33] Chakrabarti P and Janin J. Dissecting protein-protein recognition sites. Proteins 2002; 47: 334-343.

[34] Chene P. Drugs targeting protein-protein interactions. Chem Med Chem 2006; 1: 400-411.

[35] Raj M, Bullock BN and Arora PS. Plucking the high hanging fruit: a systematic approach for targeting protein-protein interactions. Bioorg Med Chem 2013; 21: 4051-4057.

[36] Vo LT, Minet M, Schmitter JM, Lacroute F and Wyers F. Mpe1, a zinc knuckle protein, is an essential component of yeast cleavage and polyadenylation factor required for the cleav- age and polyadenylation of mRNA. Mol Cell Biol 2001; 21: 8346-8356.

[37] Shi Y, Di Giammartino DC, Taylor D, Sarkeshik A, Rice WJ, Yates JR 3rd, Frank J and Manley JL. Molecular architecture of the human pre-mRNA 3'-processing complex. Mol Cell 2009; 33: 365-376.

[38] Gasteiger E, Hoogland C, Gattiker A, Duvaud S, Wilkins MR, Appel RD and Bairoch A. Protein identification and analysis tools on the ExPasy Server. ln: John M Walker, editor. The Proteomics Protocols Handbook. Human Press; 2005. pp. 571-607.

[39] Hermann H, Fabrizio P and Raker VA. snRNP Sm proteins share two evolutionarily con- served sequence motifs which are involved in Sm protein-protein interactions. EMBO J 1995; 14: 2076-2088.

[40] Blijlevens M, Meulen-Muileman IH, Menezes RX, Smit EF and van Beusechem VW. High-throughput RNAi screening reveals cancer-selective lethal targets in the RNA spliceosome. Oncogene 2019; 38: 4142-4153.

[41] Prusty AB, Meduri R, Prusty BK, Vanselow J, Schlosser A and Fischer U. Impaired spliceosomal UsnRNP assembly leads to Sm mRNA down-regulation and Sm protein degradation. J Cell Biol 2017; 216: 2391-2407.

[42] Yoshitake Y, Nakatsura T, Monji M, Senju S, Matsuyoshi H, Tsukamoto H, Hosaka S, Komori H, Fukuma D, lkuta Y, Katagiri T, Furukawa Y, lto H, Shinohara M, Nakamura Y and Nishimura Y. Proliferation potential-related protein, an ideal esophageal cancer antigen for immuno- therapy, identified using complementary DNA microarray analysis. Clin Cancer Res 2004; 10:6437-6448.

[43] Chibi M, Meyer M, Skepu A, G Rees DJ, Moolman-Smook JC and Pugh DJ. RBBP6 interacts with multifunctional protein YB-1 through its RlNG finger domain, leading to ubiquitination

and proteosomal degradation of YB-1. J Mol Biol 2008; 384: 908-916.

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

[45] Kappo MA, Ab E, Hassem F, Atkinson RA, Faro A, Muleya V, Mulaudzi T, Poole JO, McKenzie JM, Chibi M, Moolman-Smook JC, Rees DJ and Pugh DJ. Solution structure of RlNG finger-like domain of retinoblastoma-binding protein-6 (RBBP6) suggests it functions as a U-box. J Biol Chem 2012; 287: 7146-7158.

[46] Ye F, Song J and Wang Y. Proliferation potential-related protein promotes the esophageal can- cer cell proliferation, migration and suppress- es apoptosis by mediating the expression of p53 and interleukin-17. Pathobiology 2018; 85: 322-331.

[47] Kambach C, Walke S and Young R. Crystal structures of two Sm protein complexes and their implications for the assembly of the spliceosomal snRNPs. Cell 1999; 96: 375-387.

[48] Pellizzoni L, Yong J and Dreyfuss G. Essential role for the SMN complex in the specificity of snRNP assembly. Science 2002; 298: 1775-9.

[49] Kroiss M, Schultz J and Wiesner J. Evolution of an RNP assembly system: a minimal SMN complex facilitates formation of UsnRNPs in Drosophila melanogaster. Proc Natl Acad Sci US A 2008; 105: 10045-10050.

[50] Zhang R, So BR, Li P, Yong J, Glisovic T, Wan L and Dreyfuss G. Structure of a key intermediate of the SMN complex reveals Gemin2's crucial function in snRNP assembly. Cell 2011; 146: 384-95.

[51] Grimm C, Chari A and Pelz JP. Structural basis of assembly chaperone-mediated snRNP formation. Mol Cell 2013; 49: 692-703.

[52] Neuenkirchen N, Englbrecht C, 0hmer J, Ziegenhals T, Chari A and Fischer U. Reconstitution of the human U snRNP assembly machinery reveals stepwise Sm protein organization. EMBO J 2015; 34: 1925-1941.

[53] Neuenkirchen N, Chari A and Fischer U. Deciphering the assembly pathway of Sm-class U snRNPs. FEBS Lett 2008; 582: 1997-2003.

[54] Chari A, Golas MM, Klingenhäger M, Neuenkirchen N, Sander B, Englbrecht C, Sickmann A, Stark H and Fischer U. An assembly chaperone collaborates with the SMN complex to generate spliceosomal SnRNPs. Cell 2008; 135: 497-509.

[55] Heinrichs V, Hack W and Lührmann R. Direct binding of small nuclear ribonucleoprotein G to the Sm site of small nuclear RNA. Ultraviolet light cross-linking of protein G to the AAU stretch within the Sm site (AAUUUGUGG) of U1 small nuclear ribonucleoprotein reconstituted in vitro. J Mol Biol 1992; 227: 15-28.

[56] Du X, Li Y, Xia YL, Ai SM, Liang J, Sang P, Ji XL and Liu SQ. Insights into protein-ligand interac- tions: mechanisms, models and methods. Int JMol Sci 2016; 17.

[57] Boulisfane N, Choleza M and Rage F. Impaired minor tri-snRNP assembly generates differen- tial splicing defects of U12-type introns in lymphoblasts derived from a type I SMA patient. Hum Mol Genet 2011; 20: 641-648.

[58] Lefebvre S, Bürglen L and Reboullet S. lden- tification and characterization of a spinal muscular atrophy-determining gene. Cell 1995; 80: 155-165.

[59] Marabti EE and Younis I. The cancer spliceome: reprograming of alternative splicing in cancer. Front Mol Biosci 2018; 5: 80.

[60] Jing J, Zhao Y, Wang C, Zhao Q, Liang Q, Wang S and Ma J. Effect of small nuclear ribonucleoprotein-associated polypeptide N on the proliferation of medulloblastoma cells. Mol Med Rep 2015; 11: 3337-3343.

[61] Benetatos L, Hatzimichael E, Dasoula A, Dranitsaris G, Tsiara S, Syrrou M, Georgiou I and Bourantas KL. CpG methylation analysis of the MEG3 and SNRPN imprinted genes in acute myeloid leukemia and myelodysplastic syndromes. Leuk Res 2010; 34: 148-153.

[62] Lee SH, Appleby V, Jeyapalan JN, Palmer RD, Nicholson JC, Sottile V, Gao E, Coleman N and Scotting PJ. Variable methylation of the imprinted gene, SNRPN, supports a relationship between intra-cranial germ cell tumours and neural stem cells. J Neurooncol 2011; 101: 419-428.

[63] Fallini C, Bassell GJ and Rossoll W. Spinal muscular atrophy: the role of SMN in axonal mRNA regulation. Brain Res 2012; 1462: 81-92.

[64] Anchi T, Tamura K, Furihata M, Satake H, Sakoda H, Kawada C, Kamei M, Shimamoto T, Fukuhara H, Fukata S, Ashida S, Karashima T, Yamasaki l, Yasuda M, Kamada M, lnoue K and Shuin T. SNRPE is involved in cell proliferation and progression of high-grade prostate cancer through the regulation of androgen receptor expression. Oncol Lett 2012; 3: 264- 268.

[65] Johnson PF, Cavanna T, Zicha D and Bates PA. Cluster analysis of networks generated through homology: automatic identification of impor- tant protein communities involved in cancer metastasis. BMC Bioinformatics 2006; 7: 2.

[66] Steinbrecher T and Labahn A. Towards accu- rate free energy calculations in ligand protein-binding studies. Curr Med Chem 2010; 17: 767-785.

[67] Heneghan C, Blacklock C, Perera R, Davis R, Banerjee A, Gill P, Liew S, Chamas L, Hernandez J, Mahtani K, Hayward G, Harrison S, Lasserson D, Mickan S, Sellers C, Carnes D, Homer K, Steed L, Ross J, Denny N, Goyder C, Thomp-

son M and Ward A. Evidence for non-communicable diseases: analysis of Cochrane reviews and randomised trials by World Bank classifi- cation. BMJ Open 2013; 3.

[68] Bhandari GP, Angdembe MR and Dhimal M. State of non-communicable diseases in Nepal. BMC Public Health 2014; 14: 23.

[69] Hanlon L, Avila JL, Demarest RM, Troutman S, Allen M, Ratti F, Rustgi AK, Stanger BZ, Radtke F, Adsay V, Long F, Capobianco AJ and Kissil JL. Notch1 functions as a tumor suppressor in a model of K-ras-induced pancreatic ductal adenocarcinoma. Cancer Res 2010; 70: 4280-6.

[70] Witte MM and Scott RE. The proliferation potential protein related (P2P-R) gene with domains encoding heterogeneous nuclear ribonucleoprotein association and Rb1 binding shows repressed expression during terminal differentiation. Proc Natl Acad Sci U S A 1997; 94: 1212-1217.

[71] Pugh DJ, Eiso AB and Faro A. DWNN, a novel ubiquitin-like domain, implicates RBBP6 in mRNA processing and ubiquintin-like pathway. BMC Struct Biol 2006; 6: 1-12.

[72] Gao S and Scott RE. P2P-R protein overexpression restricts mitotic progression at prometaphase and promotes mitotic apoptosis. J Cell Physiol 2002; 193: 199-207.

[73] Gao S, Witte MM and Scott RE. P2P-R protein localizes to the nucleolus of interphase cells and the periphery of chromosomes in mitotic cells that show maximum P2P-R immunoreac-tivity. J Cell Physiol 2002; 191: 145-154.

[74] Chen J, Tang H and Wu Z. Overexpression of RBBP6, alone or combined with mutant TP53, is predictive of poor prognosis in colon cancer. PLoS One 2013; 8: e66524.

[75] Moela P, Choene MM and Motadi LR. Silencing RBBP6 (retinoblastoma binding protein 6) sensitises breast cancer cells MCF7 to staurosporine and camptothecin-induced cell death. Immunobiology 2014; 219: 593-601.

[76] Muleya V. Structural characterisation of the interaction between RBBP6 and the multifunctional protein YB-1. MSc Thesis. Department of Biotechnology, University of Western Cape. South Africa. URI: http://hdl.handle.net/11394/2324 Accessed June 27, 2019.

[77] Spector DL. Nuclear organisation of pre-mRNA processing. Curr Opin Cell Biol 1993; 5: 442-448.

[78] Still IH, Vince P and Cowell JK. The third member of the transforming acidic coiled coil-containing gene family, TACC3, maps in 4p16, close to translocation breakpoints in multiple myeloma, and is upregulated in various cancer cell lines. Genomics 1999; 58: 165-70.

[79] Ugolini F, Charafe-Jauffret E, Bardou VJ, Geneix J, Adélaïde J, Labat-Moleur F, Penault-Llorca F, Longy M, Jacquemier J, Birnbaum D and Pébusque MJ. WNT pathway and mammary carcinogenesis: loss of expression of candidate tu- mor suppressor gene SFRP1 in most invasive carcinomas except of the medullary type. Oncogene 2001; 20: 5810-5817.

[80] Conte N, Delaval B, Ginestier C, Ferrand A, Isnardon D, Larroque C, Prigent C, Séraphin B, Jacquemier J and Birnbaum D. TACC1-chTOG-Aurora A protein complex in breast cancer. Oncogene 2003; 22: 8102-8116.

[81] Takata A, 0tsuka M, Yoshikawa T, Kishikawa T, Kudo Y, Goto T, Yoshida H and Koike K. A miR- NA machinery component DDX20 controls NF- κB via microRNA-140 function. Biochem Bio- phys Res Commun 2012; 420: 564-569.

[82] Takata A, 0tsuka M and Yoshikawa T. MicroRNA-140 acts as a liver tumor suppressor by controlling NF-κB activity by directly targeting DNA methyltransferase 1 (Dnmt1) expression. Hepatology 2013; 57: 162-170.

[83] Shpargel KB and Matera AG. Gemin proteins are required for efficient assembly of Sm-class ribonucleoproteins. Proc Natl Acad Sci U S A 2005; 102: 17372-17377.

[84] Almstead LL and Sarnow P. Inhibition of U snRNP assembly by a virus-encoded proteinase. Genes Dev 2007; 21: 1086-1097.

[85] Chen W, Zhou P and Li X. High expression of DDX20 enhances the proliferation and metastatic potential of prostate cancer cells throughthe NF-κB pathway. lnt J Mol Med 2016; 37: 1551-1557.

[86] Shin EM, Hay HS, Lee MH, Goh JN, Tan TZ, Sen YP, Lim SW, Yousef EM, 0ng HT, Thike AA, Kong X, Wu Z, Mendoz E, Sun W, Salto-Tellez M, Lim CT, Lobie PE, Lim YP, Yap CT, Zeng Q, Sethi G, Lee MB, Tan P, Goh BC, Miller LD, Thiery JP, Zhu T, Gaboury L, Tan PH, Hui KM, Yip GW, Mi- yamoto S, Kumar AP and Tergaonkar V. DEAD- box helicase DP103 defines metastatic poten- tial of human breast cancers. J Clin Invest 2014: 124: 3807-3824.

[87] Company M, Arenas J and Abelson J. Requirement of the RNA helicase-like protein PRP22 for the release of messenger RNA from spliceosomes. Nature 1991; 349: 487-493.

[88] Ohno M and Shimura Y. A human RNA helicase-like protein, HRH1, facilitates nuclear export of spliced mRNA by releasing the RNAfrom the spliceosome. Genes Dev 1996; 10: 997-1007.

[89] Arenas JE and Abelson J. Prp43: an RNA helicase-like factor involved in spliceosome disassembly. Proc Natl Acad Sci U S A 1997; 94: 11798-11802.

[90] Hamm J and Lamond AI. Spliceosome assembly: the unwinding role of DEAD box proteins. Curr Biol 1998; 8: 532-534.

[91] Staley JP and Guthrie C. Mechanical devices of the spliceosomemotors, clocks, springs and things. Cell 1998; 92: 315-326.

[92] De la Cruz J, Kressler D and Linder P. Unwinding RNA *in Saccharomyces cerevisae*: DEAD box proteins and related families. Trends Biochem Sci 1999; 24: 192-198.

[93] Charroux B, Pellizzoni L, Perkinson RA, Shevchenko A, Mann M and Dreyfuss G. Gem-in3: a novel DEAD box protein that interacts with SMN, the spinal muscular atrophy gene product, and is a component of gems. J Cell Biol 1999; 147: 1181-94.

[94] Brückner A, Polge C and Lentze N. Yeast two-hybrid, a powerful tool for systems biology. Int J Mol Sci 2009; 10: 2763-2788.

CHAPTER 5

A paper submitted to the *American Journal of Cancer Research*
Manuscript Number: AJCR0135174

Predictive Inhibitory Studies of the Putative Interaction between Small Nuclear Ribonucleoprotein Polypeptide G and the RING finger domain of RBBP6

Lloyd Mabonga, Priscilla Masamba, Albertus Kotze Basson and Abidemi Paul Kappo

5.1 Preface − About the Manuscript

The manuscript is a bioinformatics frontier revealing the predicted structural interactions and concommittant conformational changes that occur during the binding events between the small nuclear ribonucleoprotein polypeptide G and the RING finger domain of RBBP6.The ability to computationally-ideate and explore the two proteins in PPI-focused drug discovery is indispensable in developing synthetically feasible small molecule smart-drugs against cancer. The computational derived data accumulated from the molecular interactions determined in this manuscript gives a better understanding into the biological insights of the putative interactions between the small nuclear ribonucleoprotein polypeptide G and the RING finger domain of RBBP6. The docking and MD simulation results can be extrapolated into mapping experimental studies between the two oncogenic proteins. The manuscript also reports on the identification of a putative small molecule inhibitor, (2R)-2-[(2-methyl-5-phenyl-pyrazol-3-yl)carbonylamino]-3-naphthalen-2-yl-propanoic acid (4FI). The 4FI inhibitor exhibits enhanced inhibitory activity against the protein-protein interaction between the small nuclear ribonucleoprotein polypeptide G and the RING finger domain of RBBP6,suggesting its potential use as a “lead compound” in the design of new PPI-focused anti-cancer drugs.

Inhibitory potential of a benzoxazole derivative, 4FI against SNRPG~RING finger domain protein complex in anti-cancer drug design: a molecular dynamics simulation approach

[1]Lloyd Mabonga, [2]Priscilla Masamba, [1]Albert Kotze Basson and [2]Abidemi Paul Kappo

[1]Department of Biochemistry and Microbiology, University of Zululand, KwaDlangezwa 3886, South Africa

[2]Molecular Biophysics and Structural Biology (MBSB) Group, Department of Biochemistry, University of Johannesburg, Kingsway Campus, Auckland Park 2006, South Africa

***Correspondence: Prof Abidemi Paul Kappo**

Molecular Biophysics and Structural Biology (MBSB) Group

C2 Lab Room 426

Department of Biochemistry

Faculty of Science

University of Johannesburg

Kingsway Campus

Auckland Park, 2006

South Africa

Email: akappo@uj.ac.za Tel: +27 11 559 3182, Fax: +27 11 559 2605

ABSTRACT

The etiology of developing smart-drugs has been a subject of intense scientific pursuit for centuries. It has spawned fascinating invocation on the biological insights targeting protein-protein interactions (PPI) as potential therapeutic tools in diseases such as cancer. The small nuclear ribonucleoprotein polypeptide G (SNRPG) and retinoblastoma-binding protein 6 (RBBP6) are two good examples of core-splicing proteins that have been in the limelight over the years due to their regulatory roles in mRNA processing and their active involvement in tumorigenesis and tumor development. Varying expression levels of the two proteins have been reported in different types of cancer. Suggestive evidence has postulated possible involvement between SNRPG and RBBP6 through its RING finger domain. However, the precise mechanisms of their interaction in cancer remain elusive. Therefore, the ability to computationally-ideate and explore the two proteins in PPI-focused drug discovery may yield synthetically feasible small molecule smart-drugs against cancer. This study explored this possible interaction by first modeling the proteins and validating the models by generating Ramachandran plots via PROCHECK. Following the removal of hydrogen atoms from the RING finger domain protein in UCSF Chimera, Autodock Vina was used to determine binding poses between RING and SNRPG. Thereafter a small molecule inhibitor known as (2R)-2-[(2- methyl-5-phenyl-pyrazol-3-yl) carbonylamino]-3-naphthalen-2-yl-propanoic acid (4FI) was identified using I-TASSER. The study showed, by Molecular Dynamics Simulation, that the molecule exhibits enhanced inhibitory activity against the PPI, therefore suggesting its potential use as a "lead compound" in the synthetic design of new PPI-focused anti-cancer drugs.

Keywords: Cancer; 4FI; protein-protein interaction; SNRPG; RBBP6; tumorigenesis.

Running Title: Identification of 4FI as a lead compound in anti-cancer drug design.

1. Introduction

Elucidating protein structure and function in biomedical research has become highly automated with the advent of bioinformatics [1]. Interestingly, bioinformatics tools have turned out to be an essential component as well as an integral part of structural biology. Structural biology and bioinformatics are dynamic areas of scientific research attracting a high level of attention in recent years [1, 2]. The impact of bioinformatics in the last two decades and its current contribution to understanding the molecular events or mechanisms of various diseases has positioned it at the forefront in science and technology [3, 4]. In order to develop novel drug candidates in the treatment and management of diseases, the identification and characterization of protein structure and function becomes very crucial. The determination of the active site of the proteins involved in the disease is important to design specific and selective small molecule inhibitors, ligands or peptides that are capable of modulating protein activity [5]. Consequently, focusing on elucidating the exact biological and biochemical functions of proteins and subsequently analyzing their interactome patterns can be complicated and time consuming using current conventional *in vitro* or *in vivo* methods [6]. Until recently, the majority of newly sequenced proteins have unknown structures and functions. Ironically, the lack of information of a protein structure and function restricts its further utilization in downstream drug design and discovery. Therefore, *in silico* methods for predicting protein interaction patterns in various cellular environments, provides an alternative practical solution [7].

The various types of interactions that exist between proteins are important in the discovery of a variety of biological functions and pathways within the living system [8]. Traditionally, proteins carry out their functions by interacting with other proteins. Protein-protein interactions play a

crucial role in a variety of cellular biochemical processes like DNA transcription and replication, cellular defenses, metabolic cycles, enzymatic catalysis and signaling cascades [9]. Therefore, the correct identification and characterization of such proteins, as well as the detailed investigation of protein-protein interaction network partners can greatly increase our understanding towards the elucidation of protein functions at the molecular level within the cell [8, 9].

Small nuclear ribonucleoprotein polypeptide G (SNRPG) is a core-splicing protein essential in the biogenesis of small nuclear ribonucleoproteins (snRNPs), which are precursors of the spliceosome [10-13]. Retinoblastoma-binding protein 6 is a splicing-associated multi-domain and multi-functional nuclear protein known to play a role in mRNA splicing, cell cycle control and apoptosis [14]. RBBP6 interacts with tumour suppressor proteins p53 and pRb in which the RING finger domain plays an essential role [15, 16]. Varying expression levels of the two proteins have been reported in different types of cancer such as breast, lung, prostate and colon cancer [15, 17-19]. However, not much is known about the putative interactions between RBBP6 and SNRPG in different types of cancers. Chibi and co-workers as well as Kappo and colleagues predicted possible interactions between SNRPG and RBBP6 using a yeast 2-hybrid (Y2H) technique [14, 20]. The findings from these studies suggest possible involvement of SNRPG and RBBP6 through its RING finger domain in tumorigenesis and tumour development [14]. However, the precise mechanisms involved remain elusive and yet to be characterized [9]. Therefore, the ability to computationally-ideate and explore the two proteins in protein-protein interaction (PPI)-focused drug discovery may yield synthetically feasible drugs against cancer. This paper reports the computationally derived interaction of SNRPG and the RING finger domain of RBBP6, and the identification of a putative small molecule inhibitor. The results suggest that (2R)-2-[(2-methyl-5-

phenyl-pyrazol-3-yl)carbonylamino]-3-naphthalen-2-yl-propanoic acid (4FI) inhibitor exhibits enhanced inhibitory activity against the SNRPG~RING finger domain complex, thus suggesting its potential use as a "lead compound" in the design of new protein-protein interaction (PPI) focused anti-cancer drugs.

2. Materials and Methods

Homology Modeling of SNRPG and the RING finger domain of RBBP6

Homology modeling of SNRPG and the RING finger domain of RBBP6 was first performed. The FASTA protein sequences of SNRPG (UniProtKB - P62308 (RUXG_HUMAN)) and the RING finger domain of RBBP6 (UniProtKB - Q7Z6E9 (RBBP6_HUMAN)) were retrieved from the National Centre for Biotechnology Information (NCBI) database. Protein sequence similarity was analyzed using BLAST [21] within the NCBI non-redundant sequence database (https://blast.ncbi.nlm.nih.gov/). Based on the protein similarity tests, sequences with homology identity of 40% and above were chosen as a template and used to generate the three-dimensional structures of SNRPG and the RING finger domain of RBBP6 using MODELLER, a computer program built into the visualization UCSF Chimera software, that models proteins based on their spatial constraints [22, 23]. The accuracy and stereochemistry parameters of the modeled structures were validated by the generation of Ramachandran plots, which analyses the rotations of the polypeptide backbone around the torsional phi-psi dihedral angles; this was achieved using PROCHECK [24].

Molecular Docking

The SNRPG and RING finger domain of RBBP6 were prepared for docking using UCSF Chimera software [22] by deleting all hydrogens from the ligand (RING finger domain of RBBP6), adding them to the receptor (SNRPG) and saving the structures in .mol2 and .pdb formats, respectively. Prior to docking, MetaPocket 2.0 was used to identify the binding sites available on SNRPG. Using AutoDock tools, the files were opened and converted into .pdbqt and rec.pdb formats respectively. A grid box was then created to surround the residues present in the SNRPG binding site with specific X, Y and Z dimensions and centers. Raccoon and Autodock Graphical user interface supplied by MGL tools were then used to dock the RING finger domain of RBBP6 onto SNRPG using default docking parameters. The complex with the least energy function (the lowest Z-score) was then selected for further analysis.

I-TASSER ligand prediction yield

Using I-TASSER structure prediction, ligands for SNRPG were successfully identified; (2R)- 2-[(2-methyl-5-phenyl-pyrazol-3-yl)carbonylamino]-3-naphthalen-2-yl-propanoic acid (4FI), (11alpha,16alpha)-9-fluoro-11,17-dihydroxy-16-methyl-3,20-dioxopregna-1,4-dien-21-yldihydrogen phosphate (3T5), 3-fluoro-N-(naphthalen-2-ylcarbonyl)-D-phenylalanine (GIA)and 1-[(2R,3R,4S,5R)-3,4-dihydroxy-5-(hydroxymethyl)oxolan-2-yl]pyrimidine-2,4-dione (URI)[25-27]. The ligands were selected using the I-TASSER structure-based protein function predictions, which provide annotations on ligand binding sites by matching structural models of the target protein to known proteins in protein function databases [25-27]. From the C-score values, the most reliable ligand of interest was selected. C-score is the confidence score of the prediction, and it ranges from 0-1, where a higher score indicates a more reliable prediction [25-

27]. The selected ligand was thereafter docked onto the SNPRG~RING finger domain complex and subjected to molecular dynamics (MD) simulation.

Molecular Dynamics (MD) Simulation

Molecular dynamics (MD) simulation gives clarity to the physical movements of molecules and atoms within biological complexes, and this was performed on selected docked poses using the AMBER18 software suite [28]. The set-up of MD simulation with Free Energy Workflow (FEW) was used in conjunction with the available free energy calculation procedures. Partial atomic charge of the RING finger domain of RBBP6 protein was calculated using the Restrained Electrostatic Potential (RESP) procedure [28]. The ANTECHAMBER tool within the AMBER18 software suite was used to assign RESP partial charge and General Amber Force Field (GAFF) procedures [29]. The LEAP module in the AMBER18 software suite was then used to add hydrogen molecules to the docked complexes, which were parameterized with the AMBER FF14SB force field (ff03), which is specific for bioorganic systems [28]. The system charge was neutralized through the addition of counter ions. An equilibrated TIP3P water box was then added to the system such that the boundaries of the water molecules had an 8 Å distance between the proteins. Cubic periodic boundary conditions were applied, while long-range electrostatic interactions were treated with the particle-mesh Ewald method with a non-bonding cut-off distance of 10 Å.

Initial energy minimization was applied with a restraint potential of 2kcal/mol using the steepest descent method for 1000 iterations, followed by a conjugate gradient protocol for 2000 steps. This was immediately followed with a full minimization of 1000 iterations without any imposed restraints. Harmonic restraints with force constants 5 kcal/mol was applied to all solute atoms during the heating phase. A canonical ensemble (NVT) of molecular dynamics was carried out for

50 ps, during which the system was gradually annealed from 0 - 300 K using a Langevin thermostat with a coupling coefficient of 1/ps. Thereafter, the system was freely equilibrated at 300 K and kept at constant operating temperature estimating 500 ps for the systems. The SHAKE algorithm was then employed on all atoms covalently bonded to a hydrogen atom during equilibration and production runs. Furthermore, with no restraints imposed, a production run was performed for 30ns in an isothermal isobaric (NPT) ensemble using a Berendsen barostat with a target pressure of 1 bar and a pressure coupling constant of 2 ps.

Post-Dynamic Analysis

The coordinate files of the complexes were saved and their trajectories analyzed every 1 ps using the PTRAJ module in the AMBER18 suite [28]. The thermodynamic properties of the systems, which include the Root-Mean-Square-Deviation (RMSD), Root Mean Square Fluctuation (RMSF), Radius of Gyration (RoG), Solvent-Accessible Surface Area (SASA) and hydrogen bonding were investigated.

Binding Free Energy Calculations

The Prime module was used to compute the ligand binding energies using a physics-based molecular mechanics-generalized born surface area (MM/GBSA) approach to understand protein-ligand binding [28, 30]. The MM/GBSA binding free energy profile of the selected ligand was compared to the top two compounds from each library to rank the virtual screening hits more precisely [28, 30]. Single trajectory tactics were applied and the binding free energy calculations were averaged over 1000 snapshots from a 30ns trajectory [28, 30]. MM/GBSA free energy of binding (ΔG_{bind}) is calculated for the docked poses and Desmond trajectories using the following equation:

$$\Delta G_{Bind} = G_{Complex} - G_{Ligand} - G_{Receptor} \ldots\ldots (1)$$

where $G_{Complex}$, G_{Ligand}, and $G_{Receptor}$ are the energy calculations done in Prime MM/GBSA of the optimized complex (complex), optimized free ligand (ligand), and optimized free receptor (receptor). The following set of equations describes the calculations of the binding free energy:

$$\Delta G_{Bind} = E_{gas} + G_{sol} - TS \quad (2)$$

$$E_{ga} = E_{int} + E_{vdw} + E_{ele} \quad (3)$$

$$G_{sol} = G_{GB} + G_{SA} \quad (4)$$

$$G_{SA} = ySASA \quad (5)$$

where the parameters are denoted as van der Waals energies (E_{vdW}), Electrostatic energies (E_{ele}), Gaseous energy from the ff03 force field terms (E_{Gas}), Solvation free energy (G_{solv}), Polar solvation contribution (G_{GB}), and Nonpolar solvation contribution (G_{SA}). The electrostatic and non-electrostatic interactions between the ligand and the receptor are provided by ΔE_{ele} and ΔE_{VdW}, respectively. ΔG_{gb} and $\Delta G_{surf}$are the electrostatic and non-polar contributions to the solvation free energy [31-35].

3. Results and Discussion

Homology modeling of SNRPG and the RING finger domain of RBBP6

Homology modeling of the proteins was performed by MODELLER [37] and visualized by UCSF Chimera [22]. Generation of the predicted 3D structures with MODELLER requires alignment of the amino acid sequences to existing ones of known structures with percentage identities of at least 30% or more being used as templates. The PDB structures used as templates for modeling SNRPG and the RING finger domain of RBBP6 were 1VU2 chain 4 (Small nuclear ribonucleoprotein G) from *Homo sapiens* (100%) and 3ZTG chain A (RING finger-like domain of RBBP6) from *Homo sapiens* (100%), respectively. The modeled structures of SNRPG and the RING finger domain of RBBP6 showing the proportion of secondary elements in the proteins are shown in Figure 1. The structure of the SNRPG comprise of one short a-helix and seven antiparallel β-sheets made up of

three pairs of β-hairpin motifs. On the other hand, the RING finger domain structure consists only of two a-helices connected together by a long loop from the C-terminus of the protein, which is an ever-present structural feature in RING finger domains. The availability of 3D structural models for SNRPG and the RING finger domain of RBBP6 is important in order to identify their interactive partners as well as to understand their biological functions at the molecular and cellular level [38].

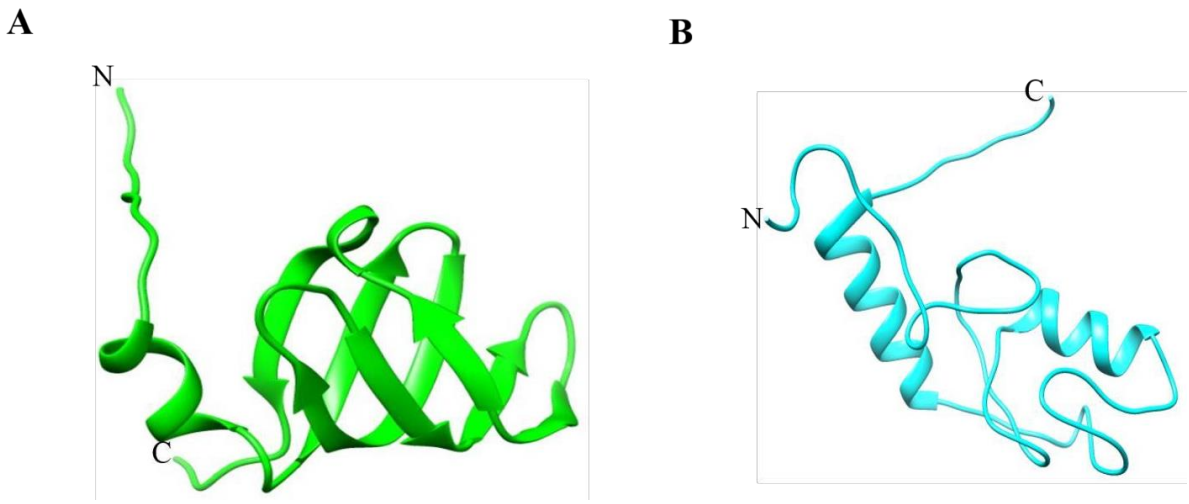


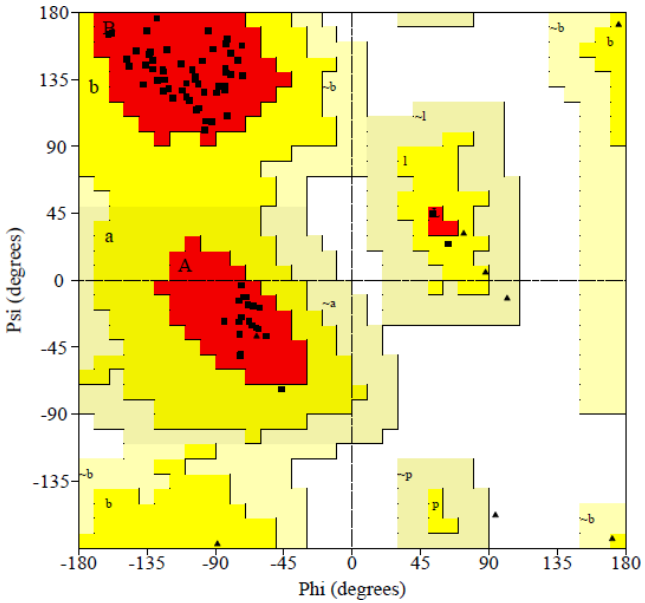
Figure 1. Homology models of **A:** SNRPG comprising of seven anti-parallel β-sheets and a short a-helix and **B:** RING finger domain of RBBP6, which consists of two a-helices connected by a long random coil emerging from the C-terminus of the protein. The figures have been generated using MODELLER (Webb and Sali, 2014).

Generated models often display different accuracies and it remains imperative to verify and validate protein models using sequence similarity, environmental parameters, and the quality of the templates [36]. The structural quality of the protein models were evaluated by Ramachandran plot analysis in RAMPAGE [38, 39]. Validation of the 3D models of SNRPG and the RING finger domain of RBBP6 are shown in Figure 2. The results obtained from the Ramachandran plots reveal that no stereo-chemical parameters were violated during model building; the protein models are stable and of good quality since most of the residues were situated within the favored regions and none in the disallowed regions [24]. The two protein models were therefore used for subsequent analysis.

Molecular docking of SNRPG and the RING finger domain of RBBP6

Understanding the mechanism of PPIs between SNRPG and the RING finger domain of RBBP6 underlies the basis of deciphering the patho-mechanism of diseases such as cancer. Determining the three-dimensional structure of SNRPG~RING finger domain protein complex and identifying "hot spot" regions responsible for the PPI provides a plumb line to determine the binding energy of the protein complex with the view of extrapolating towards targeted therapies using PPI-focused smart-drug modulators [40]. In this context, *in silico* approaches conducted in this study compliment the experimental methods conducted in previous studies [14] by elucidating the 3D structure of the SNRPG~RING finger domain protein complex through protein-protein docking. The results from the docking studies using Racoon and AutodockVina revealed the molecular binding and interaction patterns between SNRPG and the RING finger domain of RBBP6. Four SNRPG~RING finger domain protein complexes were generated having molecular docking scores of -2.86, -2.60, -3.23, and -3.40.

A



B

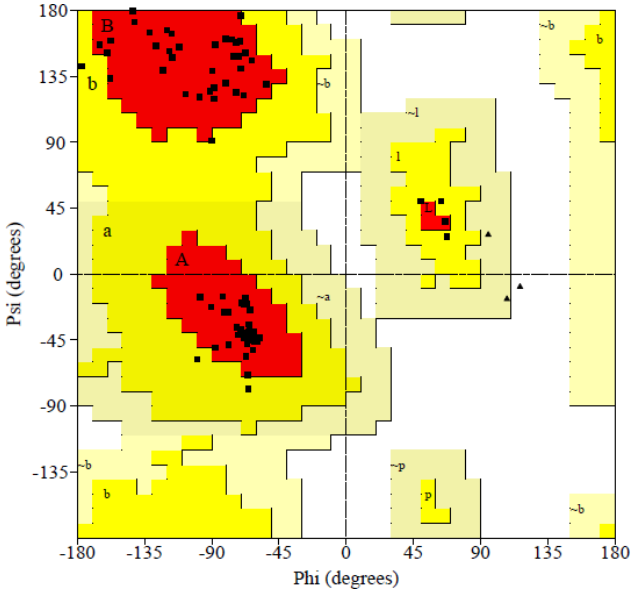


Figure 2. Ramachandran plots of predicted models of **(A)** SNRPG and **(B)** RING finger domain of RBBP6. The overall stereochemical parameters showed that 97.2% of the SNRPG residues and 92.6% RING finger domain residues appeared in the most favored regions, while 2.8% of the SNRPG residues and 7.4% of the RING finger domain residues occurred in the additional allowed regions. No residues from both proteins were present in the outlier regions. Figures were generated using PROCHECK (Laskowski *et al.*, 1993).

The binding scores of the docked complexes determine the binding mode and site, predict the absolute binding affinity and in structure-based drug design, it is a parameter that helps to identify the potential drug hits for a given protein target by searching a large ligand database [41]. Binding scores exhibit different accuracies and computational efficiencies however, the commonly used scoring functions include force-field, empirical, knowledge-based and consensus scoring [41]. Therefore, the results of the docking between SNRPG and the RING finger domain of RBBP6 as depicted by the best binding complex as well as the docking pose with the lowest energy score selected for downstream analysis is shown in Figure 3. Conceptually, protein-protein docking has limitations of not being able to address aspects of protein dynamics, which include the global conformational changes that occur during protein complex formation. However, protein-protein docking compliments experimental assays in PPI-focused drug discovery processes [40, 42].

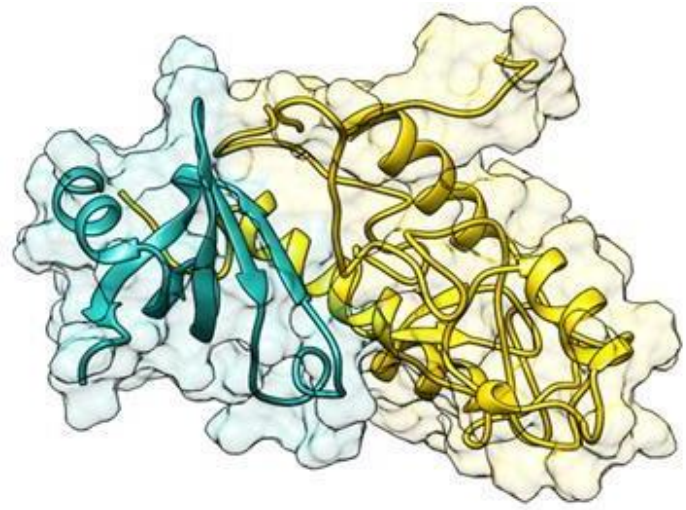
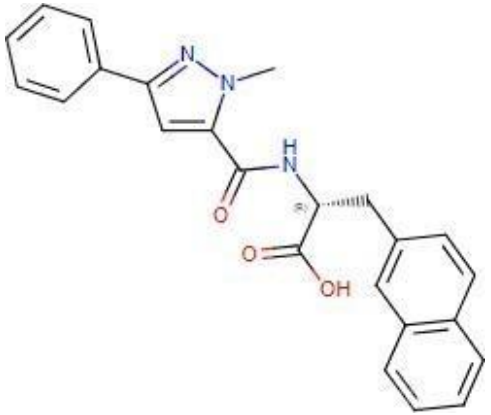
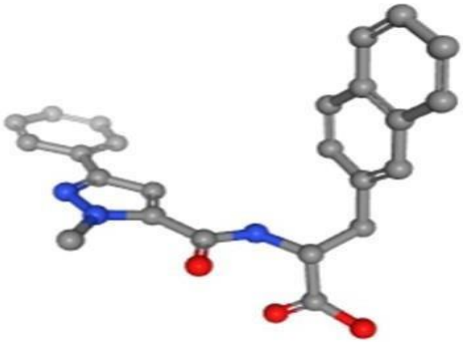
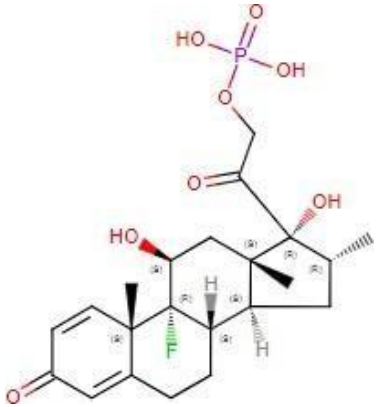
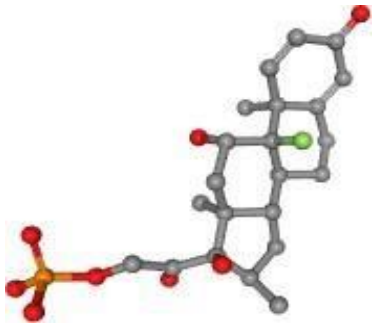


Figure 3. Structure showing the best docking pose of the SNRPG~RING finger domain protein complex. The docking score of this protein complex pose was given as -3.40kcal/mol.

Ligand prediction and molecular docking

Based on the C-score values, the I-TASSER search engine generated four ligands as shown in Table 1. The C-score values are confidence scores representing the assertiveness of the ligand hence, the higher the C-score value, the more trustworthy the estimate [25-27].The ligand with the highest C-score (4FI: 094) was selected and docked onto the SNPRG~RING finger domain protein complex before being immediately subjected to MD simulations. Table 2 depicts the docked poses of the protein complex together with the 4FI ligand and their associated binding scores. The task of predicting ligand binding sites from the sequence of a protein has received increasing attention over the past years. Therefore, computational tools have been developed to accurately predict the precise location of binding sites and the different amino acid residues that directly interact with ligands in a protein. Various approaches for the prediction of ligand binding sites have been proposed [43], both from the structure and sequence based on sequence conservation, geometric criteria of the protein surface or homology transfer from known structures [43, 44]. Predicting ligand binding sites has the potential for yielding high impact on life science research, ranging from functional characterization of novel proteins to applications in drug design, especially if the predictions are specific and accurate enough to help in addressing relevant biological questions.

Table 1. Chemical and cartoon representations of predicted ligands from I-TASSER

Ligand	C-score value	Chemical structure	Cartoon structure
4FI	0.94		
3T5	0.37		

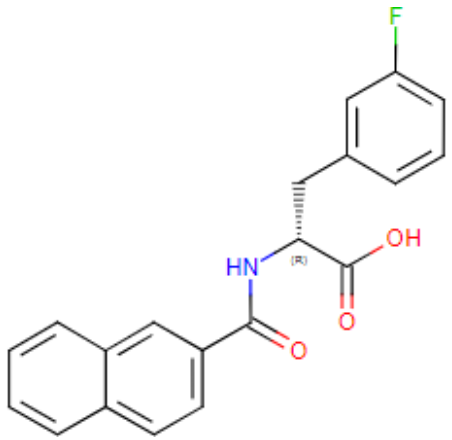
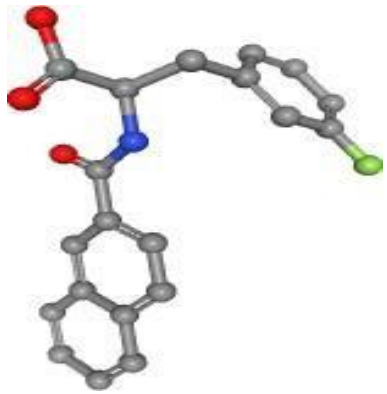
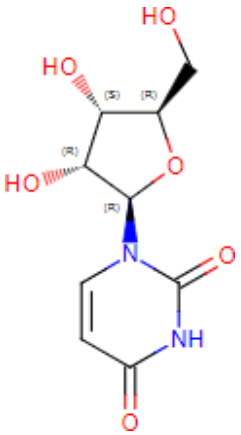
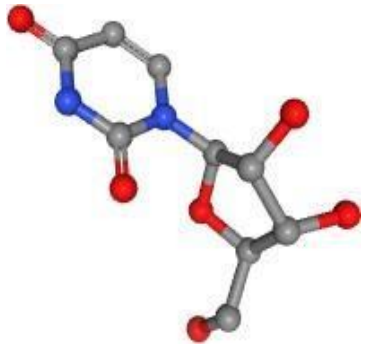
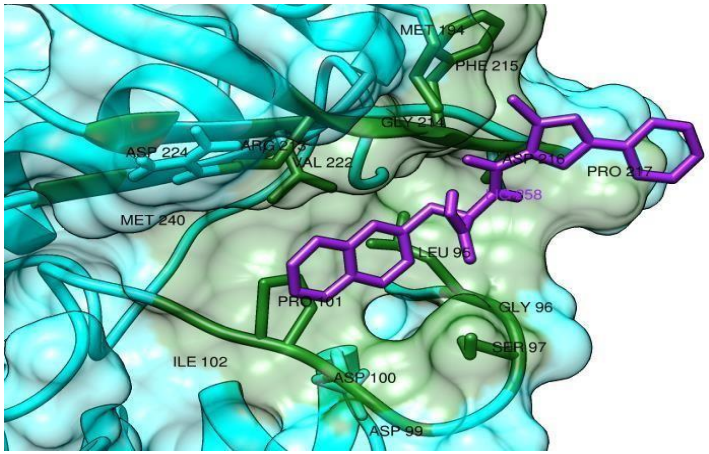
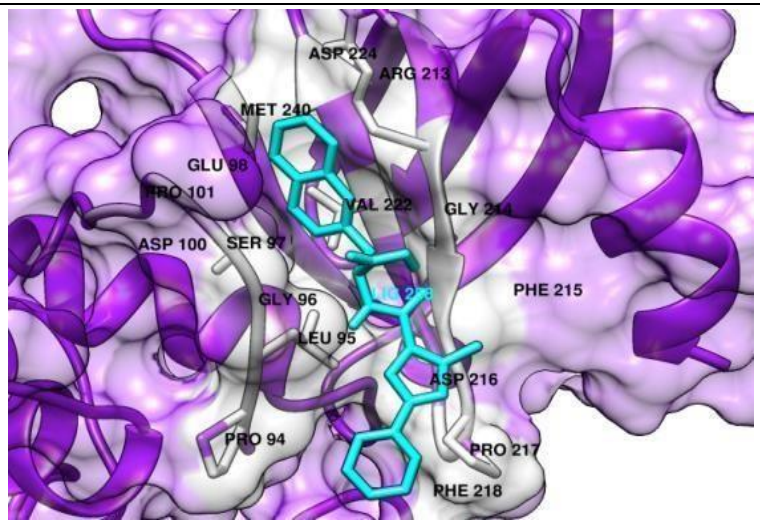
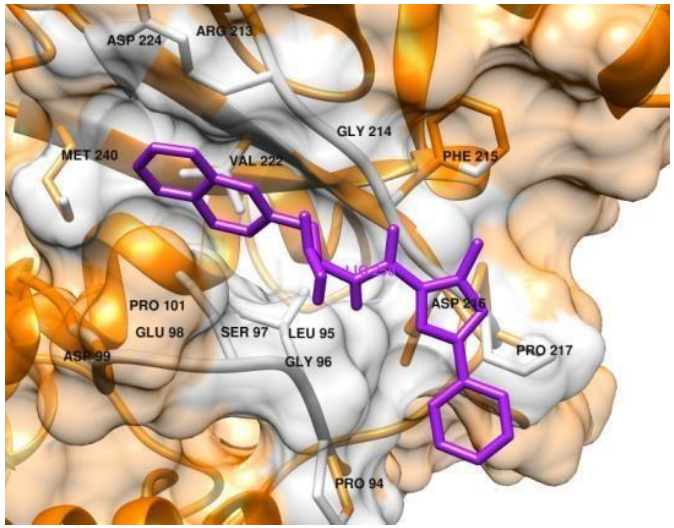
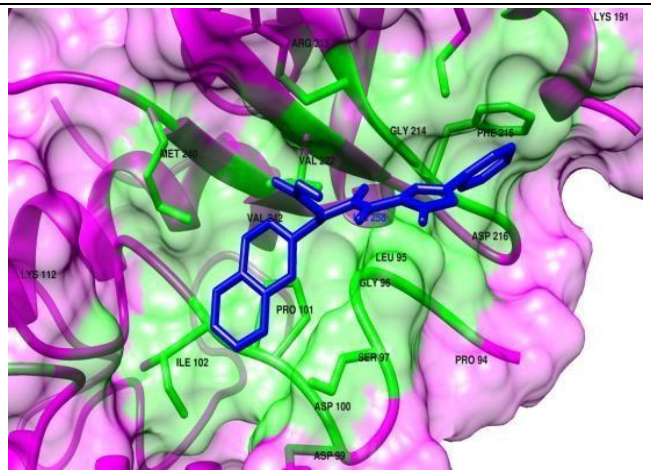
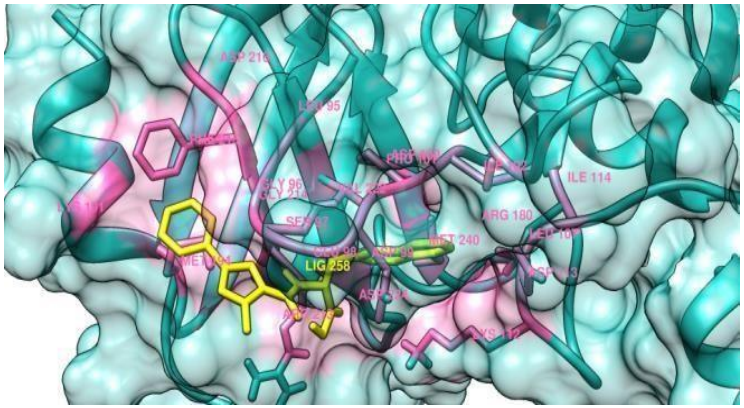
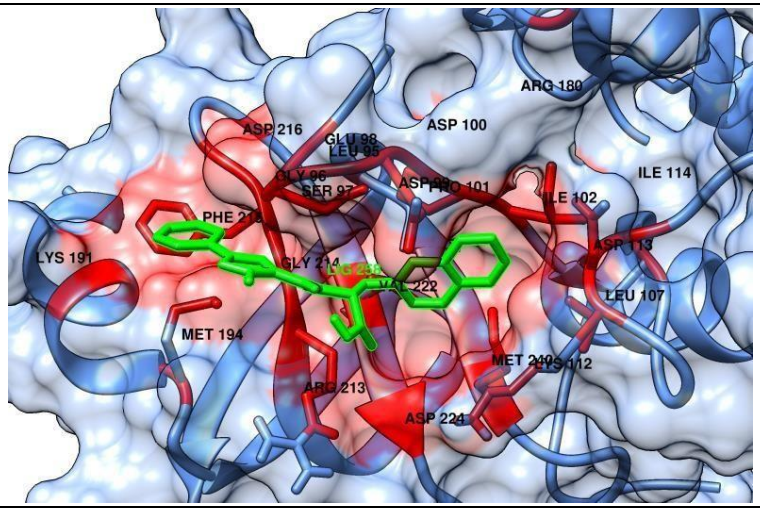
GIA	0.22		
URI	0.05		

Table 2. Molecular docking poses of the 4FI ligand onto the SNRPG~RING finger domain complex. The amino acid residues shown in red are those involved in the interaction between the two proteins irrespective of the docking pose.

No	Docked pose	Interacting Residues	Docking scores
1		**Met194**; **Phe215**; Gly214; Asp216; Pro217; **Arg213**; **Val222**; Asp224; Met240; **Leu95**; **Pro101**; Ile102; Asp100; Asp99; Ser97; Gly96.	-3.2kcal/mol
2		Asp224; **Arg213**; Met240; **Val222**; Gly214; Glu98; **Pro101**; Asp100; **Ser97**; Gly96; **Phe215**; Asp216; Pro217; Pro94.	-2.8kcal/mol

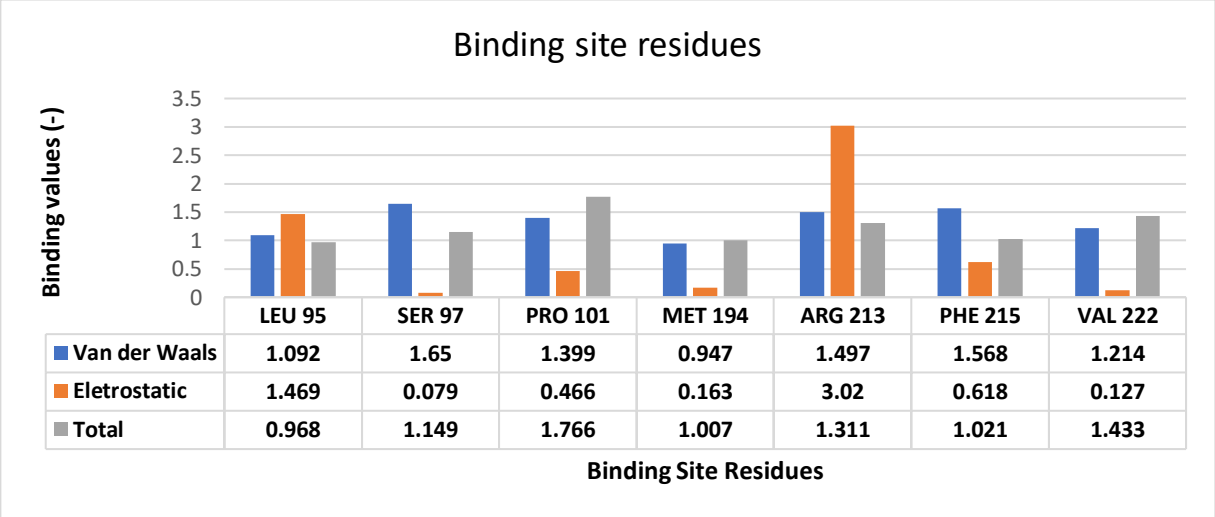
3		Asp224; **Arg213**; Met240; **Val222**; Gly214; **Phe215**; **Pro101**; Glu98; **Ser97**; **Leu95**; Asp99; Gly96; Asp216; Pro217; Pro94	-3.1kcal/mol
4		Lys191; **Arg213**; Met240; **Val222**; Gly214; Phe215; Val242; Asp216; Lys112; Ile102; **Pro101**; **Leu95**; Gly96; Asp216; **Ser97**; Pro94; Asp100; Asp99	-3.1kcal/mol

5		Asp216; **Leu95**; Lys191; **Met194**; Gly96; Gly214; **Ser97**; **Val222**; Asp100; Glu98; Asp99; Met240; **Arg213**; Asp224; Met240; Arg180; Leu107; Asp113; Lys112; Ile114	-3.0kcal/mol
6		Lys191; **Met194**; Phe215; Asp216; Gly96; **Ser97**; Glu98; **Leu95**; Gly214; **Arg213**; Asp100; **Pro101**; Leu107; Ile102; Ile114; Asp113; Met240; Lys112; Asp224	-2.6kcal/mol

The results of the binding site residues in the active site of SNRPG~RING finger domain protein complex when bound to 4FI showed more than 71% participation of hydrophobic amino acid residues with non-polar R-groups, 14% hydrophilic amino acid residues with polar R-groups and 14 % basic and electrically-charged amino acid residues possessing positively charged R-groups. According to Barratt and co-workers, it is generally accepted that the hydrophobic association of non-polar residues results from the tendency of the solvent to form a more ordered structure in the vicinity of hydrophobic groups [45]. Hydrophobic groups that are accessible to solvent prior to the association becomes buried during the complex formation with a consequent increase in the number of unstructured water molecules. Thus, the association of hydrophobic groups or non-polar R-groups clustering is often known to be highly thermodynamically endothermic and entropy driven, due to the disruption of ordered solvation shells around the interacting species [45]. However, in this study, the thermodynamic signature is not evident. Contrastingly, exothermic binding appears to be the rule in the 4FI-SNRPG~RING finger domain ligand-protein complex interactions. From Table 2, it is evident to see that regardless of conformation or docking pose, there are at least seven key interacting resides (highlighted in red) that contributed significantly to the final binding affinity in relation to the van der Waals and electrostatic attractions. Using complex 1 (the docked pose with the best docking score of -3.2kcal/mol), the results are interpreted more clearly on the histogram presented in Figure 4, which elucidates the individual contribution and the total binding input of each binding residue based on the van der Waals and electrostatic attractions involved in the predicted binding event.

In predicting binding sites from the sequence of a protein, it has been suggested that the electrostatic interactions dominate the non-covalent binding in molecular recognition. However, this is not generally true as it is well-known that shape complementarities are also very important

[46]. Over the years, it has been established that predicting binding sites from the sequence of a protein is attributed to contributions from both electrostatic and van der Waals interactions, solvation/desolvation and entropy effects. The amino acid residues Arginine (Arg) and Leucine (Leu) thus contribute the most towards the electrostatic attractions in the predicted binding event hence, suggesting that electrostatic interactions are a key factor for binding affinity. In this study, most of the residues contributed above 1.0 binding values for van der Waals forces. Nevertheless, Proline (Pro) residues contributed the highest total score to the binding and Leucine (Leu) the least.



	LEU 95	SER 97	PRO 101	MET 194	ARG 213	PHE 215	VAL 222
Van der Waals	1.092	1.65	1.399	0.947	1.497	1.568	1.214
Eletrostatic	1.469	0.079	0.466	0.163	3.02	0.618	0.127
Total	0.968	1.149	1.766	1.007	1.311	1.021	1.433

Figure 4. Energy profile of the key binding site residues of SNRPG~RING finger domain protein complex when bound to 4FI.

Generally, it is established that van der Waals energy changes (ΔE_{VDW}) for most ligands upon binding are less than the electrostatic energy changes (ΔE_{ELE}) by orders of magnitude 1 to 2. In this study, the non-polar contribution of salvation-free energy for the SNRPG~RING finger domain protein~4FI complexes are negative thus suggesting that the contribution of the van der Waals energy change is less than that of the electrostatic energy change. Hence, the two major factors to determine the binding affinity are the electrostatic energy change and solvation free

energy change [35, 47]. In drug design, the binding site residues are often used to characterize the binding strength between a receptor and a drug molecule. The fundamental goal of structure-based drug design is to find new lead compounds that bind as tightly as possible to macromolecular receptors [35, 47]. Therefore, the results from this study suggest the potential use of 4FI as a "lead compound" in the design of new PPI-focused anti-cancer drugs against the PPI between SNRPG and RING finger domain.

Analyses of system stability by MD simulations

The 4FI ligand in complex with the SNRPG~RING finger docked pose (with a docking score of -3.23) was prepared and subjected to MD simulations using AMBER18 program. MD simulation was performed for a total of 80 ns and trajectories were analyzed for interaction fingerprints of the compound in the active site over time as shown in Figure 5. The root mean square deviation (RMSD) is the average displacement of the atoms at an instant of the simulation relative to a reference structure, which analyzes the time-dependent motions of the structure [48-51]. It is used to determine the stability of a structure in the timescale of the simulation and to discern if it is diverging from the initial coordinates [48, 50, 51]. A high level of fluctuation was seen in the 4FI~SNRPG~RING finger domain complex at around 15 ns stabilizing at approximately 40 ns till the end of the simulation, suggesting the adjustment of the ligand within the active site of the complex. According to Martínez (2015), high RMSD and RMSF values indicate that the whole structure fluctuates or reflects only large displacements of a small structural subset within an overall rigid structure.

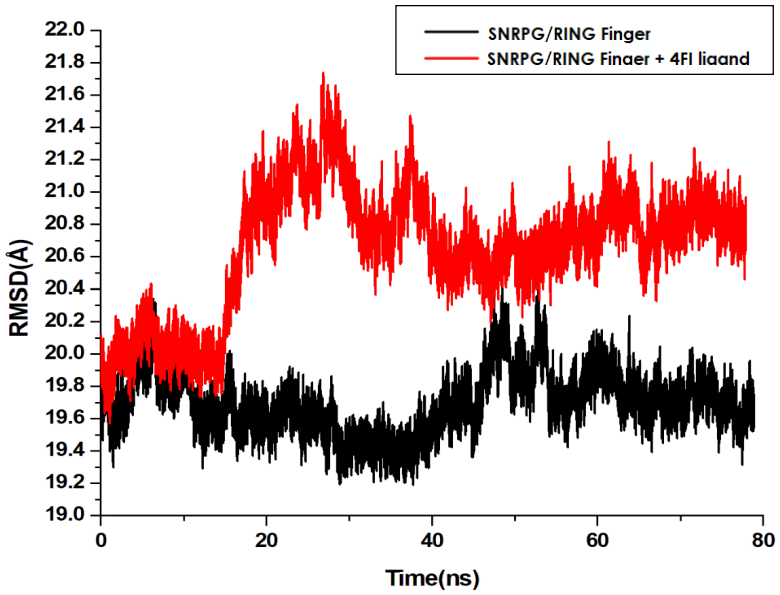


Figure 5. Root Mean Square Deviation (RMSD) of the SNRPG~RING finger domain complex with the 4FI ligand (red) and SNRPG/RING finger domain complex (black).

The root mean square fluctuation (RMSF) determines the average atomic mobility of backbone atoms (N, Ca and C) during MD simulation. It captures, for each atom, the fluctuation of its average position. To understand the conformational changes and explore the effect of ligand binding on protein dynamics, the RMSF of the studied systems is calculated from MD trajectories [50, 51]. This analysis gives an overview of the flexible regions in the protein and corresponds to temperature factors. The RMSF of the instantaneous structures of the two models was plotted as a function of residue (Figure 6). The binding of a ligand to a complex may restrict the active intrinsic movement of residues in the active site, resulting in the loss of flexibility. As shown in the residues in Figure 5, there was a general slight increase in the RMSF profile of the ligand-bound complex in comparison to the apo (SNRPG~RING finger domain complex), where the most significant fluctuation is shown at residues ~120–150.

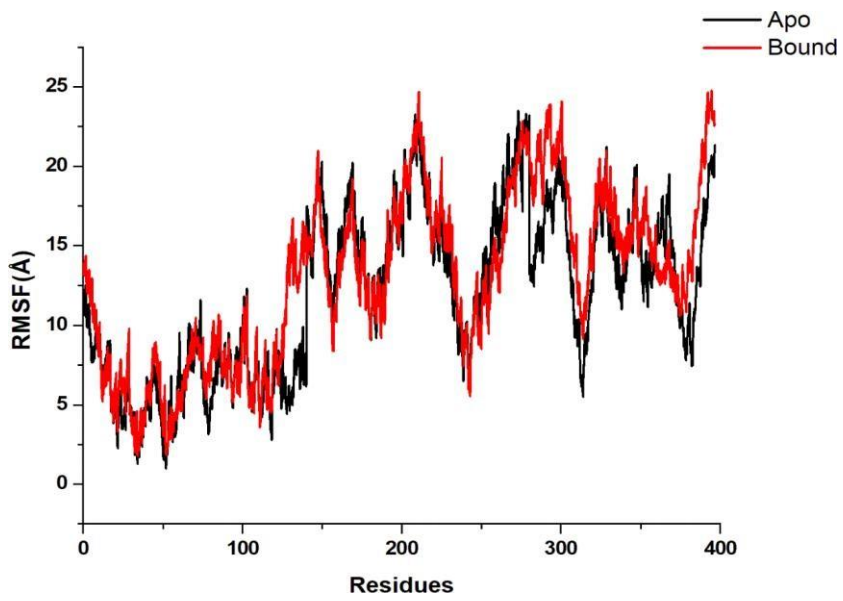


Figure 6. RMSF profile of the SNRPG/RBBP6-RING finger domain complex with the 4FI ligand during the MD simulations.

The radius of gyration (RoG) profile simulated in Figure 7 presents a rigid-body alignment that is very sensitive to the existence of structural subsets with high conformational fluctuations. The core of the protein appears to be more rigid relative to the loops (solvent-exposed) with very high fluctuations. A very significant difference is seen between the profiles of each complex. A considerable rise in gyration occurred just before 20ns up until approximately 40ns, showing very high fluctuations and representing loss of compactness until the profile eventually sees some form of stability. This trajectory correlates to the RMSD profile, which showed high fluctuations at the same time, due to positioning of the ligand within the active site of the complex.

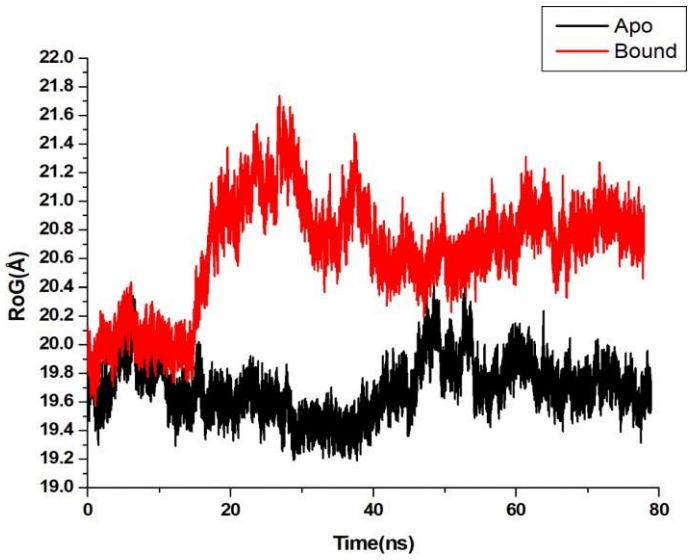


Figure 7. RoG profile of the SNRPG/RING finger domain complex with the 4FI ligand during the MD simulations.

Thermodynamics of the Complexes: Binding Free Energy Calculations

Molecular mechanics generalized Born surface area (MM/GBSA) is arguably one of the most common techniques to predict binding free energy because it is more accurate than most scoring functions of molecular docking [31, 52]. MM/GBSA is less computationally demanding than alternative chemical free energy scoring methods hence, its wide use in biomolecular studies such as protein folding, protein−ligand binding and PPIs amongst others [32, 33]. Free energy drives molecular processes and the accurate determination of the free energy is one of the most important tasks in biomolecular studies [31-33, 35, 52]. In the MM/GBSA approach, the binding free energy is approximated by the interaction energy between the ligand and the receptor complex and the difference in solvation free energy upon binding [32-34]. The correlation between binding free energy (ΔG_{bind}) and its individual components (ΔE_{vdW}, ΔE_{ele}, ΔH, ΔG_{solv}, $-T\Delta S_{tot}$) were analyzed and these are given in Table 3.

Table 3. MM/GBSA binding free energies (kcal/mol) of SNRPG/RING finger in complex with 4FI

System	ΔE_{vdW}	ΔE_{ele}	ΔE_{gas}	ΔG_{solv}	$-T\Delta S_{tot}$	ΔG_{bind}
SNRPG~RING finger domain in complex with 4FI	-36.74	-13.83	-50.57	22.62	- 0.01	-27.96

The inter-relation between ΔG_{bind} and its individual components (ΔE_{vdW}, ΔE_{ele}, ΔH, ΔG_{solv}, $-T\Delta S_{tot}$) depicts positive correlation between the buried nonpolar surface area and the binding free energy for several peptidic ligands [53]. The remarkably strong correlations between ΔG_{bind} (-27.96 kcal/mol) and ΔE_{VdW} (-36.74 kcal/mol) indicates that the binding of the sorting signal (4FI ligand) is mainly driven by hydrophobic interactions. From the results shown in Table 3, the Gibbs energy difference ($-T\Delta S_{tot}$) was calculated using the following equation:

$$\Delta G_{bind} = \Delta H - T\Delta S \approx \Delta E_{ELE} + \Delta E_{VDW} + \Delta G_{sol} - T\Delta S_{tot}, \text{ where}$$

The results revealed a -0.01 kcal/mol value of the Gibbs energy difference ($-T\Delta S_{tot}$). The Gibbs energy represents the energy available for the reaction and this includes enthalpy and entropy contributions. According to Ylilauri and Pentikäinen, if the Gibbs energy difference is negative the reaction is spontaneous since temperature is always a positive number. If the process is in equilibrium, the Gibbs energy difference is equal to zero and if its positive then the reaction is unfavourable and the initial state with the unbound ligand is favoured [54].

From this study, the negative Gibbs energy difference suggests that the interaction between the 4FI ligand and SNRPG~RING finger domain complex does not favor the initial state and is not at

equilibrium. It is actually spontaneous with high chances of favoring the 4FI ligand bound to the SNRPG~RING finger domain complex rather than the apo state. Hence, the binding of the 4FI ligand is expected to reshape the sorting signal binding groove in the SNRPG~RING finger domain complex in a way that optimal van der Waals contacts between the complex and the ligand are possible and non-polar solvent accessible surface area is diminished upon binding. Molecular recognition in drug-target interactions is known to be attributable to the contributions from both electrostatic and van der Waals interactions, solvation/desolvation and entropy effects [35, 55, 56].

The MM/GBSA calculations have the advantage of prediction accuracy at a much lower computational costs in comparison to the much more time-consuming methods widely used in the field of small-molecule drug design, which include post-processing of structure-based virtual screening [57, 58]. MM/GBSA is a very useful tool in analyzing the binding details of drug-target interactions since they can capture very important and vital regions within receptor-ligand systems for rational drug design [57-59]. The binding affinity of the 4FI ligand bound to the SNRPG~RING finger domain complex structure is -27.96 kcal/mol. This negative binding affinity value indicates good binding between the ligand and the two proteins. The results obtained in this study present a channel towards the implementation of MM/GBSA in small molecule drug design to quantitatively-characterize the binding of 4FI to the SNRPG~RING finger domain system. The combination of molecular docking and MM/GBSA rescoring is a promising strategy in the identification of the correct binding poses and the correct ranking of the binding affinities of ligands [60, 61].

4. Conclusion

Cancer remains one of the major public health concerns worldwide. Increasing morbidity and mortality rates, the lack of specificity of most cancer drugs, and the side effects of cancer therapy remain a major challenge. Over the years, PPI-focused drug technology has emerged as a promising arsenal for cancer therapeutic drug discovery. The technology fosters a paradigm proposition that insinuates targeting regulatory core splicing proteins as potential PPI drug candidates for cancer. The SNRPG and RING finger domain of RBBP6 are two good examples of regulatory proteins that have drawn attention over the years due to their regulatory roles in mRNA processing and their active involvement in tumorigenesis and tumor development. Varying expression levels of SNPRG and RBBP6 have been reported in different types of cancer with suggestive evidence postulating possible involvement between SNRPG and the RING finger domain of RBBP6. However, the precise mechanisms of the interaction remains elusive and yet to be characterized. Therefore, this study aimed to computationally-ideate and explore the PPI between the two proteins and further identify possible small molecule inhibitors towards PPI-focused drug discovery. The results predicted possible binding between the proteins with interaction patterns observing relatively high binding free energies as well as high levels of compactness and stability, thereby supporting previous studies suggesting high probability of possible involvement between the two proteins in cancer cell networks. Key conserved interacting residues within the SNPRG~RING finger complex are proposed for the design of small molecule inhibitors as adjuvants with anti-cancer potential. Furthermore, the 4FI ligand exhibited enhanced inhibitory activity against the SNRPG~RING finger complex, favoring its binding to the SNRPG/RING finger complex. The data accumulated from this *in silico* molecular interactions suggest the potential use of 4FI as a “lead compound” in the design of new PPI-focused anti-cancer

drugs against the SNPRG~RING finger domain protein-protein interactions. However, further exploration into the molecular and structural properties of the mechanism of action and strength of this complex through experimental evaluation is indispensable in validating the efficacy of 4FI as a potential anti-cancer agent in order to substantiate the results from this study.

Acknowledgements

Research reported in this manuscript was supported South African Department of Science and Innovation (DSI), the National Research Foundation (NRF) and the University of Zululand. Lloyd Mabonga is grateful to the NRF for a doctoral research support and Dr. Priscilla Masamba for a Scarce-skill Postdoctoral Fellowship at the University of Johannesburg. The content of this paper is the sole responsibility of the authors and do not necessarily represent the official views of the supporting institutions.

Competing Interest

The authors declare no conflict of interest.

References

[1] Liang X, Zhu W, Lv Z, Zou Q. Molecular Computing and Bioinformatics. Molecules 2019; 24: 2358.

[2] Elumalai A and Eswaraiah MC. Review on application of bioinformatics. J Sci Bioinform 2013; 3: 21-27.

[3] Tsuji S and Aburatani H. Machine Learning Applications in Cancer Genome Medicine. Gan to kagaku ryoho. Cancer Chemother 2019; 46: 423-426.

[4] Wang Y, Deng G, Zeng N, Song X and Zhuang Y. Drug-disease association prediction based on Neighborhood information aggregation in neural networks. IEEE Access 2019; 7: 50581-50587.

[5] Stephenson N, Shane E, Chase J, Rowland J, Ries D, Justice N, Zhang J, Chan L and Cao

R. Survey of machine learning techniques in drug discovery. Curr Drug Metab 2019; 20: 185-193.

[6] Chen R, Liu X, Jin S, Lin J and Liu J. Machine learning for drug-target interaction prediction. Molecules 2018; 23: 2208.

[7] Agamah FE, Mazandu GK, Hassan R, Bope CD, Thomford NE, Ghansah A and Chimusa ER. Computational/in silico methods in drug target and lead prediction. Brief Bioinform 2020; 21: 1663-1675.

[8] Kovács IA, Luck K, Spirohn K, Wang Y, Pollis C, Schlabach S, Bian W, Kim D-K, Kishore N and Hao T. Network-based prediction of protein interactions. Nat Commun 2019; 10: 1-8.

[9] Mabonga L and Kappo AP. The oncogenic potential of small nuclear ribonucleoprotein polypeptide G: a comprehensive and perspective view. Am J Transl Res 2019; 11: 6702.

[10] Lührmann R, Kastner B and Bach M. Structure of spliceosomal snRNPs and their role in pre-mRNA splicing. Biochim Biophys Acta 1990; 1087: 265-292.

[11] Palfi Z, Lücke S, Lahm H-W, Lane WS, Kruft V, Bragado-Nilsson E, Séraphin B and Bindereif A. The spliceosomal snRNP core complex of Trypanosoma brucei: cloning and functional analysis reveals seven Sm protein constituents. Proc Natl Acad Sci 2000; 97: 8967-8972.

[12] Stevens SW and Abelson J. Purification of the yeast U4/U6· U5 small nuclear ribonucleoprotein particle and identification of its proteins. Proc Natl Acad Sci 1999; 96: 7226-7231.

[13] Will CL and Lührmann R. Spliceosome structure and function. Cold Spring Harb Perspect Biol 2011; 3: a003707.

[14] Kappo MA, Eiso A, Hassem F, Atkinson RA, Faro A, Muleya V, Mulaudzi T, Poole JO, McKenzie JM and Chibi M. Solution structure of RING finger-like domain of retinoblastoma-binding protein-6 (RBBP6) suggests it functions as a U-box. J Biol Chem 2012; 287: 7146-7158.

[15] Hull R, Oosthuysen B, Cajee U-F, Mokgohloa L, Nweke E, Antunes RJ, Coetzer TH and Ntwasa M. The drosophila retinoblastoma binding protein 6 family member has two isoforms and is potentially involved in embryonic patterning. Int J Mol Sci 2015; 16: 10242-10266.

[16] Pugh DJ, Eiso A, Faro A, Lutya PT, Hoffmann E and Rees DJG. DWNN, a novel ubiquitin-like domain, implicates RBBP6 in mRNA processing and ubiquitin-like pathways. BMC Struct Biol 2006; 6: 1-12.

[17] Khan F, Allan M, Tincho M and Pretorius A. Implications of RBBP6 in various types of Cancer. Proceedings IWBBIO 2014.

[18] Li L, Deng B, Xing G, Teng Y, Tian C, Cheng X, Yin X, Yang J, Gao X and Zhu Y. PACT is a negative regulator of p53 and essential for cell growth and embryonic development. Proc Natl Acad Sci 2007; 104: 7951-7956.

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

[20] Chibi M, Meyer M, Skepu A, Rees DJG, Moolman-Smook JC and Pugh DJ. RBBP6 interacts with multifunctional protein YB-1 through its RING finger domain, leading to ubiquitination and proteosomal degradation of YB-1. J Mol Biol 2008; 384: 908-916.

[21] Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W and Lipman DJ. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 1997; 25: 3389-3402.

[22] Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC and Ferrin TE. UCSF Chimera—a visualization system for exploratory research and analysis. J Comput Chem 2004; 25: 1605-1612.

[23] Yang J, Roy A and Zhang Y. BioLiP: a semi-manually curated database for biologically relevant ligand–protein interactions. Nucleic Acids Res 2012; 41: D1096-D1103.

[24] Lovell SC, Davis IW, Arendall III WB, De Bakker PI, Word JM, Prisant MG, Richardson JS and Richardson DC. Structure validation by Ca geometry: ϕ, ψ and Cβ deviation.

Proteins Struct Function Bioinform 2003; 50: 437-450.

[25] Roy A, Yang J and Zhang Y. COFACTOR: an accurate comparative algorithm for structure-based protein function annotation. Nucleic Acids Res 2012; 40: W471-W477.

[26] Yang J and Zhang Y. I-TASSER server: new development for protein structure and function predictions. Nucleic Acids Res 2015; 43: W174-W181.

[27] Zhang Y. I-TASSER: Fully automated protein structure prediction in CASP8. Proteins Struct Function Bioinform 2009; 77: 100-113.

[28] Case D, Darden T and Kollman P. University of California; San Francisco: 2008. Google Scholar There is no corresponding record for this reference 2018.

[29] Wang J, Wolf RM, Caldwell JW, Kollman PA and Case DA. Development and testing of a general amber force field. J Comput Chem 2004; 25: 1157-1174.

[30] Peddi SR, Sivan SK and Manga V. Molecular dynamics and MM/GBSA-integrated protocol probing the correlation between biological activities and binding free energies of HIV-1 TAR RNA inhibitors. J Biomol Struct Dyn 2018; 36: 486-503.

[31] He X, Liu S, Lee T-S, Ji B, Man VH, York DM and Wang J. Fast, Accurate, and Reliable Protocols for Routine Calculations of Protein–Ligand Binding Affinities in Drug Design Projects Using AMBER GPU-TI with ff14SB/GAFF. ACS Omega 2020; 5: 4611-4619.

[32] Ma G, Yu H, Xu X, Geng L, Wei X, Wen J and Wang Z. Molecular basis for metabolic regioselectivity and mechanism of cytochrome P450s toward carcinogenic 4-(methylnitrosamino)-(3-pyridyl)-1-butanone. Chem Res Toxicol 2019; 33: 436-447.

[33] Perthold JW and Oostenbrink C. GroScore: Accurate Scoring of Protein–Protein Binding Poses Using Explicit-Solvent Free-Energy Calculations. J Chem Inf Model 2019; 59: 5074-5085.

[34] Ugur I, Schatte M, Marion A, Glaser M, Boenitz-Dulat M and Antes I. Ca2+ binding induced sequential allosteric activation of sortase A: An example for ion-triggered conformational selection. PloS One 2018; 13: e0205057.

[35] Wang E, Sun H, Wang J, Wang Z, Liu H, Zhang JZ and Hou T. End-point binding free energy calculation with MM/PBSA and MM/GBSA: strategies and applications in drug design. Chem Rev 2019; 119: 9478-9508.

[36] Muhammed MT and Aki-Yalcin E. Homology modeling in drug discovery: Overview, current applications, and future perspectives. Chem Biol Drug Des 2019; 93: 12-20.

[37] Webb B and Sali A. Comparative protein structure modeling using MODELLER. Curr Protoc Bioinform 2014; 47: 5-6.

[38] Zhu J, Wang S, Bu D and Xu J. Protein threading using residue co-variation and deep learning. Bioinform 2018; 34: i263-i273.

[39] Schwede T. Protein modeling: what happened to the "protein structure gap"? Structure 2013; 21: 1531-1540.

[40] Salmaso V and Moro S. Bridging molecular docking to molecular dynamics in exploring ligand-protein recognition process: An overview. Front Pharmacol 2018; 9: 923.

[41] Hernández-Santoyo A, Tenorio-Barajas AY, Altuzar V, Vivanco-Cid H and Mendoza-Barrera C. Protein-protein and protein-ligand docking. Protein Eng Tech Appl 2013; 63-81.

[42] Kaczor AA, Bartuzi D, Stępniewski TM, Matosiuk D and Selent J. Protein–protein docking in drug design and discovery. Computational Drug Discovery and Design. Edited by Gore M. and Jagtap U. Humana Press, NY, 2018, pp. 285-305.

[43] Gherardini PF and Helmer-Citterich M. Structure-based function prediction: approaches and applications. Brief Funct Genomic Proteomic 2008; 7: 291-302.

[44] Lopez G, Ezkurdia I and Tress ML. Assessment of ligand binding residue predictions in CASP8. Proteins Struct Funct Bioinform 2009; 77: 138-146.

[45] Barratt E, Bingham RJ, Warner DJ, Laughton CA, Phillips SE and Homans SW. Van der Waals interactions dominate ligand− protein association in a protein binding site occluded from solvent water. J Am Chem Soc 2005; 127: 11827-11834.

[46] Gohlke H and Klebe G. Approaches to the description and prediction of the binding affinity

of small-molecule ligands to macromolecular receptors. Angew Chem Int Ed 2002; 41: 2644-2676.

[47] Zhang X, Perez-Sanchez H and C Lightstone F. A comprehensive docking and MM/GBSA rescoring study of ligand recognition upon binding antithrombin. Curr Top Med Chem 2017; 17: 1631-1639.

[48] Martínez L. Automatic identification of mobile and rigid substructures in molecular dynamics simulations and fractional structural fluctuation analysis. PloS One 2015; 10: e0119264.

[49] Sargsyan K, Grauffel Cd and Lim C. How molecular size impacts RMSD applications in molecular dynamics simulations. J Chem Theory Comput 2017; 13: 1518-1524.

[50] Warnau J, Wöhlert D, Okazaki K-i, Yildiz Oz, Gamiz-Hernandez AP, Kaila VR, Kühlbrandt W and Hummer G. Ion binding and selectivity of the Na+/H+ antiporter MjNhaP1 from experiment and simulation. J Phys Chem B 2019; 124: 336-344.

[51] Xiong Y, Karuppanan K, Bernardi A, Li Q, Kommineni V, Dandekar AM, Lebrilla CB, Faller R, McDonald KA and Nandi S. Effects of N-glycosylation on the structure, function, and stability of a plant-made Fc-fusion anthrax decoy protein. Front P Sci 2019; 10: 768.

[52] Li W. Residue–Residue Mutual Work Analysis of Retinal–Opsin Interaction in Rhodopsin: Implications for Protein–Ligand Binding. J Chem Theory Comput 2020; 16: 1834-1842.

[53] Myslinski JM, DeLorbe JE, Clements JH and Martin SF. Protein–ligand interactions: thermodynamic effects associated with increasing nonpolar surface area. J Am Chem Soc 2011; 133: 18518-18521.

[54] Ylilauri M and Pentikäinen OT. MMGBSA as a tool to understand the binding affinities of filamin–peptide interactions. J Chem Inf Model 2013; 53: 2626-2633.

[55] Chen BC, Tu SL, Zheng BA, Dong QJ, Wan ZA and Dai QQ. Schizandrin A exhibits potent anticancer activity in colorectal cancer cells by inhibiting heat shock factor 1. Biosci Rep 2020; 40:

[56] Tang Q, Fu W, Zhang M, Wang E, Shan L, Chai X, Pang J, Wang X, Xu X and Xu L.

Novel androgen receptor antagonist identified by structure-based virtual screening, structural optimization, and biological evaluation. Eur J Med Chem 2020; 192: 112156.

[57] Roca C, Martinez-González L, Daniel-Mozo M, Sastre J, Infantes L, Mansilla A, Chaves-Sanjuan A, González-Rubio JM, Gil C and Cañada FJ. Deciphering the inhibition of the neuronal calcium sensor 1 and the guanine exchange factor ric8a with a small phenothiazine molecule for the rational generation of therapeutic synapse function regulators. J Med Chem 2018; 61: 5910-5921.

[58] Xu W, Lau YH, Fischer G, Tan YS, Chattopadhyay A, de la Roche M, Hyvönen M, Verma C, Spring DR and Itzhaki LS. Macrocyclized extended peptides: inhibiting the substrate-recognition domain of tankyrase. J Am Chem Soc 2017; 139: 2245-2256.

[59] Zhang X, Wong S and Lightstone F. Toward Fully Automated High Performance Computing Drug Discovery: A Massively Parallel Virtual Screening Pipeline for Docking and Molecular Mechanics/Generalized Born Surface Area Rescoring to Improve Enrichment. J Chem Inf Model 2014; 54: 324-337.

[60] Wang W, Liu C, Li H, Tian S, Liu Y, Wang N, Yan D, Li H and Hu Q. Discovery of novel and potent P2Y14R antagonists via structure-based virtual screening for the treatment of acute gouty arthritis. J Adv Res 2020; 23: 133-142.

[61] Wäschenbach L, Gertzen CG, Keitel V and Gohlke H. Dimerization energetics of the G-protein coupled bile acid receptor TGR5 from all-atom simulations. J Comput Chem 2020; 41: 874-884.

CHAPTER SIX

A paper already accepted by the *American Journal of Translational Research*
Manuscript Number: AJTR0135302

Microscale Thermophoresis Analysis of the Molecular Interaction between Small Nuclear Ribonucleoprotein Polypeptide G and the RING finger domain of RBBP6 towards Anti-Cancer Drug Discovery

Lloyd Mabonga, Priscilla Masamba, Albertus Kotze Basson and Abidemi Paul Kappo

Preface − About the Manuscript

Bioinformatics evidence has been provided for SNRPG and the RING finger of RBBP6 domain as interactomes in cancer protein network. However, the binding strength has not been validated. This study determined the binding affinity and binding strength between SNRPG and the RING finger domain of RBBP6 and provides the first mechanistic insight into the biophysical interaction between the two proteins using the Microscale Thermophoresis (MST) technique. An accurate deciphering of the binding affinity between the two proteins is essential for therapeutic optimization and modifications by small molecule inhibitors towards PPI-focused anti-cancer drug discovery. Identifying small molecule adjuvants to modulate the interaction could therefore lead to a possible breakthrough in the treatment of cancer.

Am J Transl Res 2021;13(10):XXX-XXX
www.ajtr.org /ISSN:1943-8141/AJTR0135302

Original Article

Microscale thermophoresis analysis of the molecular interaction between small nuclear ribonucleoprotein polypeptide G and the RING finger domain of RBBP6 towards anti-cancer drug discovery

Lloyd Mabonga[1], Priscilla Masamba[2], Albertus Kotze Basson[2], Abidemi Paul Kappo[2]

[1]*Department of Biochemistry and Microbiology, University of Zululand, KwaDlangezwa 3886, South Africa;*
[2]*Molecular Biophysics and Structural Biology (MBSB) Group, Department of Biochemistry, Faculty of Science, University of Johannesburg, Kingsway Campus, Auckland Park 2006, South Africa*

Received May 16, 2021; Accepted August 16, 2021; Epub XXX, 2021; Published XXX, 2021

Abstract: Regulatory core-splicing proteins are now becoming highly promising therapeutic targets for the development of anti-cancer drugs. SNRPG and RBBP6 are two good examples of regulatory core-splicing proteins involved in tumorigenesis and tumor development whose multi-functional role is primarily mediated by protein-protein interactions. Over the years, skepticism abutting from the two onco-proteins has been mounting. Suggestive evidence using yeast 2-hybrid technique observed possible involvement between SNRPG and the RING finger domain of RBBP6. However, the putative interaction remains elusive and yet to be characterized. In this study, we developed the first MST-based assay to confirm the interaction between SNRPG and the RING finger domain of RBBP6. The results demonstrated a strong binding affinity between SNRPG and the RING finger domain of RBBP6 with a K_D in the low nanomolar concentration range of 3.1596 nM. The results are congruent with previous findings suggesting possible involvement between the two proteins in cancer-cell networks, thereby providing a new mechanistic insight into the interaction between SNRPG and the RING finger domain of RBBP6. The interaction is therapeutically relevant and represents a great milestone in the anti-cancer drug discovery space. Identification of small molecule inhibitors to modulate the binding affinity between the two proteins would therefore be a major breakthrough in the development of new PPI-focused anti-cancer drugs.

Keywords: Anti-cancer drug discovery, cancer, microscale thermophoresis, RBBP6, RING finger, SNPRG

Introduction

The rational optimization of protein-protein interactions (PPIs) is becoming increasingly important in modern drug discovery processes [1-3]. This is currently driven by targeting the undruggable molecular space with the aim of designing new therapeutic agents that can selectively target intractable disease-specific molecular mechanisms or pathways. PPIs are an attractive class of molecular targets in the drug discovery parlance. Drugging the undruggable proteome space with the aim of designing new therapeutic agents is an indispensable arsenal in curbing pathophysiological cues and disease progression [4]. In this context, PPIs of cancer-implicated proteins are considered high-value targets in drug development programmes [1]. Most cancer-implicated proteins possess structural domains that have a higher ratio of infidelity as compared to their non-cancer implicated counterparts, making them more prone to interaction with a wide diversity of proteins [4-6]. Cancer-implicated proteins have many interacting partners and occupy a central position in cancer-cell protein networks [6-8]. Thus, protein interactions between these macromolecules have a higher probability of being related to cancer processes than non-interacting proteins, making them therapeutically vulnerable for anti-cancer drug discovery [6-8].

Targeting cancer-implicated PPIs is a powerful arsenal to address mechanistic cues in tumori-

genesis and tumour development. Small Nuclear Ribonucleoprotein Polypeptide G (SNR- PG) and Retinoblastoma Binding Protein 6 (RBBP6) are two good examples of cancer- implicated proteins whose functions are pre- dominantly mediated by PPIs [9, 10]. The two proteins are connected by a wide array of bio- logical processes and play critical roles in pathophysiological cues. SNRPG is a core-spli- cing protein that is essential in the biogenesis of small nuclear ribonucleoproteins (snRNPs), which are precursors of the spliceosome [11-14]. RBBP6 is a splicing-associated multi-domain and multi-functional nuclear protein known to play a role in mRNA splicing, cell cycle control and apoptosis. RBBP6 interacts with tumour suppressor proteins p53 and pRb in which the Really Interesting New Gene (RING) finger domain plays an essential role [15-17].

Varying expression levels of the two proteins have been reported in different types of cancer such as breast, lung, prostate and colon cancer, but very little is known about the putative interactions between RBBP6 and SNRPG in these different types of cancers [15, 18-20]. Chibi and co-workers [21] as well as Kappo and co-workers [16] predicted possible interactions between SNRPG and RBBP6 using the yeast 2-hybrid (Y2H) technique. The findings suggest possible involvement of SNRPG and RBBP6 through its RING finger domain in tumorigenesis and tumour development. However, the precise mechanisms involved remain elusive and yet to be characterised [22].

Robust and reliable determination of the binding affinity between SNRPG and its putative interactive partner, the RBBP6 RING finger domain, is a critical step in understanding the relationship between the splicing machinery, tumorigenesis and tumour development. More so, quantitative characterization of intermolecular interaction affinity between oncogenic core splicing proteins is highly necessary to develop novel and effective drugs for therapeutic interventions in cancer [23]. Most analytical techniques for PPIs are expensive (e.g., mass spectrometry), time-consuming [e.g., Surface Plasmon Resonance (SPR), Isothermal Titration Calorimetry (ITC)] and require high amounts of sample (e.g., Size-exclusion Chromatography, Isothermal Titration Calorimetry). However, MicroScale Thermophoresis (MST) is an attractive alternative technique with advantages of speed, ultra-low sample consumption and high-throughput/cost-efficiency. Just like ITC and SPR, MST can be used to determine the equilibrium dissociation constant (K_D) and other thermodynamic parameters [2, 23-25].

MicroScale Thermophoresis (MST) is a versa- tile optical fluorescent technique used to quan- tify binding affinities in solution between a tar- get molecule and its interactive partner [25]. This biophysical technique (Figure 1) detects variations in fluorescence signals resultingfrom infrared laser-induced temperature gradi- ents. The variation in the fluorescence signal correlates with the binding of a ligand to the fluorescent target. This effect is known as TRIC (temperature-related intensity change) [2, 24]. The TRIC signals are additory and contribute to the high sensitivity and robustness of MST measurements in molecular binding events [2, 24].

However, the outstanding merit of utilising MST over other routinely used PPI methods is its ability to determine K_D values in complex sample matrices [2, 23, 24]. Although MST measurements is performed using intrinsic fluorescence of proteins, labelling of the target proteins with a suitable fluorophore is required [2, 24]. Different site-specific labelling strategies have been proposed and applied. The His-tag is the most popular and widely used affinity tag for purification, immobilization or detection of proteins. The application of tris-NTA-based labelling of His-tagged proteins is commonly used for MST measurements [24]. Thus, MST can be used to determine the binding affinity and binding strength between protein-protein biophysical interaction with very low sample consumption and high sensitivity. In this study, we used the innovative MST to establish an experimental assay for fast, precise, cost-efficient and quality-controlled characterization of the binding affinity, binding stoichiometry and interaction thermodynamics between SNRPG and the RING finger domain of RBBP6. Thestudy provides novel insights into the molecular mechanisms between the two proteins towards PPI-focused anticancer drug discovery.

Experimental procedures

Bacterial expression and purification

Codon optimized DNA sequences incorporating *BamHI* and *XhoI* restriction sites were amplified

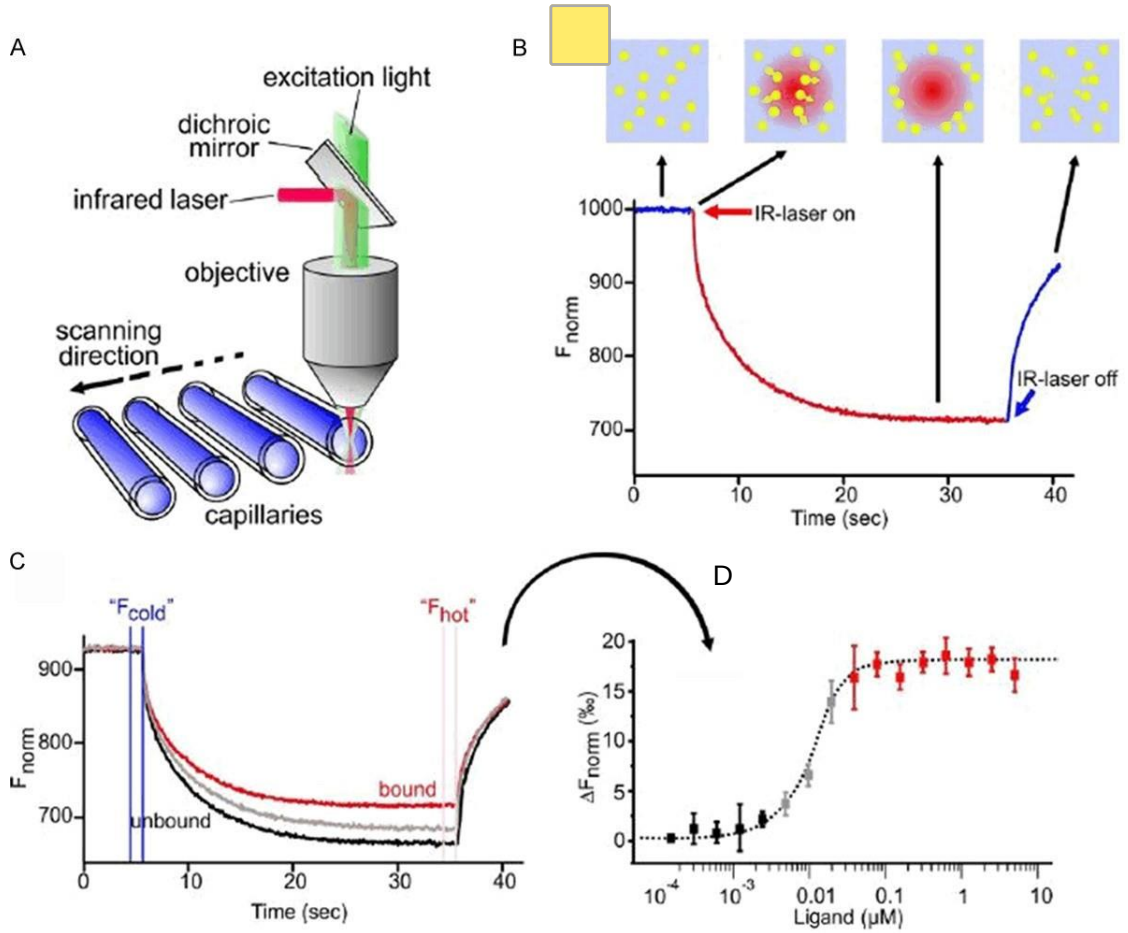


Figure 1. MicroScale Thermophoresis (MST) set-up. A. MST measurements are conducted in small glass capillaries. Infrared and fluorescence lasers trigger the MST effect and generate sample tracking. B. Temperature-related intensity change (TRIC) and thermophoresis account for the time-dependent change in fluorescence upon infrared-heating of the sample capillaries. C. Multiple MST traces are recorded for different mixture ratios of target and ligand molecules. D. Dose-response analysis of the MST traces allows for determination of the steady-state affinity of the target-ligand interaction (Figure extracted from Schubert and Langst [26]).

from the full-length cDNA sequences of SNRPG (UniProtKB-P62308 (RUXG_HUMAN)) and the RING finger domain (pdb: 3tzg) of RBBP6. The genes were cloned into the pQE30 and pGEX-6P-2 protein expression vectors and were purchased from GenScript (New Jersey, USA) to be used for protein expression. Expression of both proteins was induced with 0.5 mM isopro- pyl 1-thio-p-D-galactopyranoside (IPTG) con-centration at 25°C. The expressed SNRPG pro- tein was purified using a Nickel-NTA column recharged with cobalt, whereas the RING finger domain of RBBP6 was purified by the use of a glutathione-agarose (SIGMA® Aldrich) column and Econo® Chromatography Column (Amersham Pharmacia). The concentrations of the eluted proteins were determined using a NanoDrop® ND2000 spectrophotometer (Thermo Fisher Scientific).

MicroScale thermophoresis measurement

A 100 µl of 100 nM 6X His-tag-SNRPG protein labelled in 1X phosphate-buffered saline (pH 7.4) supplemented with 0.05% Tween-20 (PBS-T) was mixed with 100 µl of 100 nM of Monolith His-Tag Labeling Kit RED-tris-NTA 2nd Generation (MO-L018) (NanoTemper Techno- logies, Munich, Germany) diluted in 1X PBS-T buffer to a final concentration where the fluo- rescent signals of the SNRPG proteins were similar and above the typical detection limit of the Monolith NT.115 instrument (NanoTemper Technologies, Munich, Germany). The mixture

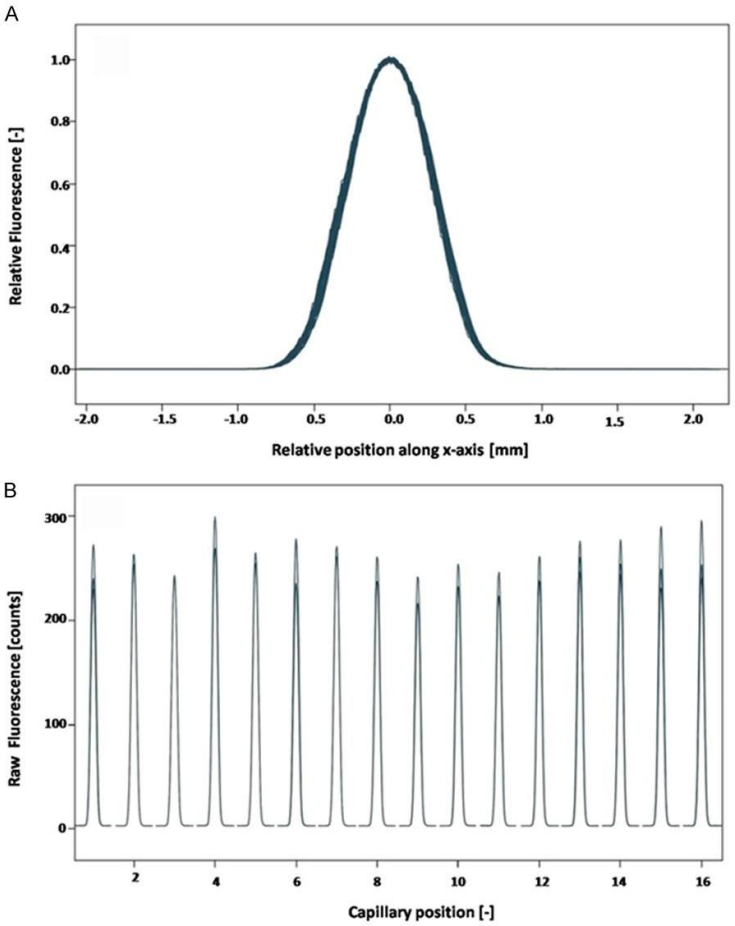


Figure 2. MST capillary scans for the 6X His-tag-SNRPG protein. A. The capillary scan graph overlayed perfectly suggesting no SNRPG protein adsorption onto the capillaries. B. The capillary scans output observed no fluctuations in protein fluorescence suggesting successful labelling of the SNRPG protein.

was incubated for 30 minutes at room temperature in the dark prior to the MST experiment. The final concentration of the fluorescently labelled 6X His-tag-SNRPG protein was 50 nM.

MicroScale thermophoresis experiments were performed on a NanoTemper® Monolith NT.115 (NanoTemper Technologies GmbH, Munich, Germany) as shown in Figure 1. Samples were prepared and loaded into premiun treated capillaries. A 16-tube serial dilution of the non-fluorescent GST-RING finger domain of RBBP6 ranging from 1.11 µM to 3.39×10^{-5} µM was titrated against a fixed concentration (50 nM) of the fluorescent 6x His-tag-SNRPG. The sample was mixed and added to each tube to a final volume of 20 µl using low-bind pipette tips. Having prepared the serial dilutions, 4 µl samples were filled into the capillaries through capillary action, resulting in low sample consumption. Triplicate MST measurements were performed on the NanoTemper® Monolith NT.115 (NanoTemper Technologies GmbH, Munich, Germany) in premium capillaries (MO-K025) at 25°C using 40% MST power and 40% light-emitting diode (LED) with laser off/on times of 5 sec-onds and 30 seconds, respec- tively. The system was operat- ed with the latest version of the MO control software (v1.6) with data analysis performed using the NanoTemper® an- alysis software settings opti-mised for TRIC-sensitive dyes.

Results

The fluorescence labelling of the 6X His-tag SNRPG protein displayed a high affinity of 3.8±0.5 nM using the RED- tris-NTA 2nd Generation dye. The novel Monolith His- Tag labelling Kit RED-tris-NTA 2nd Generation kit (MO-L018) comparatively yields higher binding amplitudes and signal-to-noise ratios [2, 27]. As shown in Figure 2A, the MST capillary scans for the 6X His-tag SNRPG protein depicted a perfectly overlayed graph, suggesting no protein adsorption of the fluorescently labelled His-tag SNRPG protein onto the capillaries. There were no fluctuations in protein fluorescence in the capillary scans output, suggesting successful labelling of the 6X His- tag SNRPG protein. It is always vital to check and avoid adsorption of proteins and reference ligands on the capillary walls as this apparently decreases or ablates ligand binding due to loss of material adversely, subsequently affectingthe MST signal and results [2, 23, 27].

Following the successful fluoresecence label-ling of the 6X His-tag SNRPG, pre-test binding checks were conducted to confirm whether there was detectable binding between the 6X

Table 1. Overview of the MST dataset between the SNRPG and RING finger domain of RBBP6

Description	Experiment
Target Name	SNRPG
Target Concentration	50 nM
Ligand Name	Ring finger domain of RBBP6
Ligand Concentration	1.11 µM to 3.39×10^{-5} µM
n	3
Excitation Power	40%
MST Power	40%
Temperature	25.0°C
K_D	3.1596×10^{-9}
K_D Confidence	$\pm7.627\times10^{-9}$
Response Amplitude	5.2101186
Target Concentration	5×10^{-8} [Fixed]
Unbound	861.21
Bound	866.42
Std. Error of Regression	1.0692345
Reduced X^2	0.96629751
Signal to Noise	5.2341793

His-tag SNRPG and the GST-RING finger domain of RBBP6. The pre-test binding checks observed positive results suggesting detect- able binding between the two proteins. Pre-tests are highly commended to facilitate ad- justments to the labelling or concentrations, thus minimizing wastage of materials from failed or indeterminant binding affinity experi- ments due to insufficient fluorescence [2, 27, 28].

The subsequent characterization and binding affinity measurements of the binding event between the SNRPG and RING finger domain of RBBP6 were conducted and analysed by MO Affinity Analysis software v2.3. An MST on- time of 1.5 seconds was used for analysis and calculation of the K_D value (n=3 independent measurements, error bars represent the standard deviation). The MST raw dataset of the merged PPI dose-response between SNRPG and RBBP6 RING finger domain is summarised in Table 1. Additional MST raw data of merged dose-response of the PPIs are embedded as Table S1, while the resulting dose-response curves were fitted to a one-site binding model to extract K_D values from a K_D-binding model assuming a 1:1 binding stoichiometry.

As indicated on the raw MST trace plot, the relative change in fluorescence was observed as the titration concentration of the RING finger domain of RBBP6 increased. The thermograph showed no signs of adhesion or aggregation (Figure 3). The result is a clear indication of the binding event between the two proteins, transiting from the unbound state to the boundstate.

According to Mrozowich and co-workers [23], an accurate binding curve should observe three significant points denoting no binding, slope of the binding and complete binding. The titration curve as shown in Figure 4 displays a typical sigmoidal shape with an atypical peak in the thermophoresis signal close to the apparent point of saturation. The observation suggests that the changes that affect fluoresence upon binding are identical and that different thermophoretic properties are formed during the binding event. The binding affinity measurement between SNRPG and the RING finger domain of RBBP6 observed a K_D of3.1596 nM under aqueous buffer conditions.

Discussion

The identification of small molecule inhibitors that may be able to modulate the binding affinity between regulatory core-splicing proteins SNRPG and the RING finger domain of RBBP6 represents a major strategy in the development of new PPI-focused anti-cancer drugs. The two oncogenic and regulatory core-splicing proteins are varyingly expressed in different types of cancers and have been overlooked as potential therapeutic arsenals for many years. The molecular mechanisms by which SNRPG and the RING finger domain of RBBP6 mediate their oncogenic networks still remain unknown and uncharacterized [15, 18-20]. The understanding of PPIs of regulatory core-splicing proteins is biologically interesting and is essential for therapeutic optimization and modifications by small molecule inhibitors towards PPI-focused anticancer drug design [29, 30]. The inhibition of the oncogenic activity of the two splicing-associated proteins by developing PPI-focused anticancer modulators appears to be a promising therapeutic alternative in cancer.

The rational optimization of molecular interactions is becoming increasingly important in PPI-focused anti-cancer drug discovery pro- cesses. The putative interaction between SNRPG and the RING finger domain of RBBP6

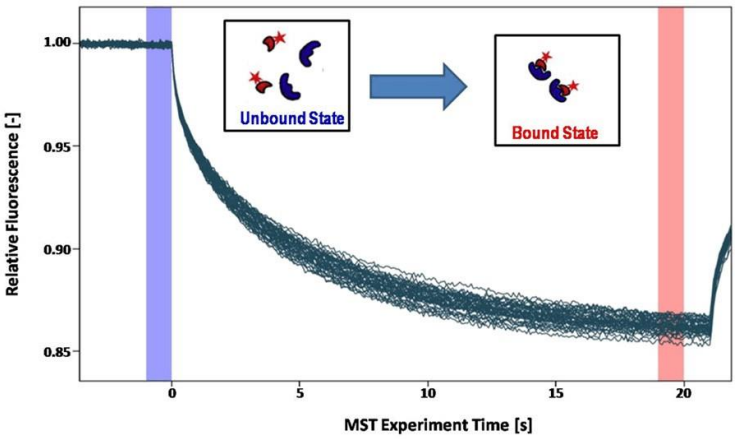


Figure 3. Thermograph of SNRPG binding to the RING finger domain of RBBP6 at 25°C. Multiple MST traces were recorded for different mixture ratios of the SNRPG and RING finger domain of RBBP6. The cold region is set to 0 seconds (blue) and the hot region to 20 seconds (red) to determine the K_D of the interaction and to avoid potential convection phenomena.

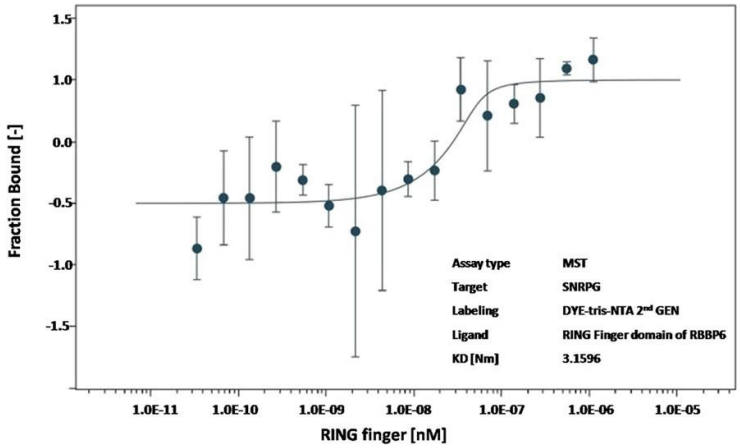


Figure 4. Dose-response curve for the binding interaction between SNRPG and the RING finger domain of RBBP6. The concentration of SNRPG protein was kept constant at 50 nM, while the ligand concentration varied from 1.11µM to 0.03 nM. The binding affinity measurement for the interaction ob- served a K_D of 3.1596 nM. The experiments were performed at 25°C with a 30 min incubation at medium MST and 40% LED power.

has been suggested over the years using the yeast 2-hybrid (Y2H) technique. Two copies of SNRPG (conformational isomers of the same protein) were identified as part of the five substrates that bind to the RING finger domain of RBBP6 [16]. In another study, Chibi and coworkers [21] also suggested that SNRPG interacts with the N-terminal domain of RBBP6, which is a crucial component of the RNA processing machinery in the cell. The findings from the two studies are biologically interesting considering that the two proteins are highly active in cancer-cell networks and vary in their expression profiles within different types of cancers. More so, the two proteins play active roles as core-splicing regulatory proteins in mRNA metabolism, which is a crucial process in tumor development and tumorigenesis. The findings suggest a high likelihood of the possible involvement between SNRPG and the RING finger domain of RBBP6 in tumorigenesis and tumor development.

Accordingly, PPIs identified using the Y2H assay are subject to verification by a series of other biochemical assays such as co-immunoprecipitation (co-IP), pull-down and co-localization experiments. This path works well only for strong PPIs. Weak binary PPIs might not be readily detectable by less-sensitive assays [31, 32]. Most analytical techniques for PPIs are expensive, time-consuming and require high amounts of sample. For this reason, a microscale thermophoresis analysis for the characterisation of the bind-ing affinity between SNRPG and the RING finger domain of RBBP6 was developed.

This study provides the first evidence of quantitative interaction affinity measurement involving the two oncogenic proteins using the MST assay. It provides an optimized methodology suitable for studying other potential oncogenic regulatory core-splicing partners. The results from the MST experiment observed high binding affinity between SNRPG and the RING finger domain of RBBP6 with a K_D value in the low nanomolar concentration range of 3.1596 nM. The experiment was conducted at 25°C with 30 minutes incubation time at medium MST and 40% LED power. The results confirm the findings by Kappo and co-

workers [16] that suggested possible involvement between the SNRPG and the RING finger domain of RBBP6. Indeed, there is strong binding between the two proteins as depicted from the low nanomolar range K_D value.

Generally, weak PPIs in the low nanomolar range are poorly understood even though they are crucial in mediating therapeutically rele-vant biological processes in health and dis- ease states. According to Salahudeen and Nishtala [33], the lower the K_D value in weak PPIs, the stronger the binding and the higher the binding affinity. Conversely, the higher the K_D value is, the weaker the binding and the lower the binding affinity. From a drug discove- ry point of view, the aim is to identify disease- relevant PPIs with lower K_D values (i.e., high binding affinity). Such PPIs have strong binding strength and can be used as targets for inhibi- tion using small molecule inhibitors for PPI- focused anticancer drug discovery [33, 34]. Therefore, this study could be used as a start- ing point to perform high-throughput screening with SNRPG~RING finger domain of RBBP6 inhibitors to evaluate the ability of small mole- cules to modulate the affinity of the two pro- teins. The modulation of this interaction would represent a major breakthrough in the deve- lopment of new strategies targeting immune escape in oncology.

The successful characterization of PPIs between oncogenic proteins and the determi- nation of their binding affinity measurements using the MST technique has been reported in various cancer-related studies. Magnez and co- workers [35] have used the MST assay to deter-mine the binding affinity between the trans- membrane glycoprotein PD-1 and a type I trans- membrane protein PDL1, which are involved in tumour escape processes towards designing small molecule inhibitors as anti-cancer drug agents. In another study, Liberelle and co-work-ers [36] provided the first MST interaction affinity measurement involving the oncogenic ErbB2 protein tyrosine kinase receptor and its membrane partner, the MUC4 mucin, to vali-date finding small molecule binding affinities for targeting the MUC4-ErbB2 protein complex for drug discovery in cancer. These studies are useful indices to justify the potential of using the results from this MST assay to further explore the interaction between SNRPG and the RING finger domain of RBBP6 as potential anti-cancer drug targets in PPI-focused drug discovery.

The physical interaction between SNRPG and the RING finger domain of RBBP6 provides the first mechanistic insight of their structure-func-tion relationship. The observation is of particu-lar interest in pharmacological research sinceit provides a platform to study their possible involvement in cancer-cell networks towards anti-cancer drug discovery [37, 38]. An accu-rate deciphering of the binding affinity between the two proteins is essential for therapeutic optimization and modifications by small mole-cule inhibitors [29, 30]. Currently, there are no developed therapeutic approaches that target blockage of the SNRPG-RING finger domain of RBBP6 protein complex. Over the years, many similar studies in cancer reporting the suc-cessful modulation of therapeutically relevant onco-proteins have been conducted yielding promising results. Some drugs have already been approved, while others have entered clini-cal trials (summarised in Table 2).

PPI-focused anticancer drug strategies target-ing interactions such as the MDM2/p53, Bcl-2/Bax, xIAP/caspase-9, Hsp90/Cdc37, c-Myc/Max, KRAS/PDEδ, CD40/CD40L, Skp2/Skp, Keap1/Nrf2 and PD-1/PD-L1 have shed light on the role of protein complexes in the quest to drug the once undruggable proteome space.The studies indicate that PPIs have great po- tential as intervention targets for novel treat- ments of refractory types of cancers and their regulation is an indispensably promising strat- egy in drug discovery. Blocking the two regula- tory splicing proteins can help generate new anti-cancer 'lead' compounds and thus pro- duce new treatment drugs [39, 40].

Targeting core-splicing regulatory proteins for anti-cancer drugs remains the epitome of fu-ture prospectives in PPI-focused drug discov-ery. Given the complexity of splicing regulation and its centrality in driving biological process- es in pathological states, targeting the interac- tion between SNRPG and the RING finger domain of RBBP6 for drug discovery in cancer may provide a better understanding into the future of PPI-focused drug discovery from a dif- ferent perspective. The biological connection between the splicing machinery and apopto- sis, a phenomenon that allows the regulated

Table 2. Examples of PPIs that have yielded modulators that are either approved or in clinical trials

PPI	Related disease	Drug	Status	References
PD-1/PD-L1	Non-small lung cancer	Keytruda	Approved-2014	[41]
PD-1/PD-L1	Non-small lung cancer	Opdivo	Approved-2014	[42]
PD-1/PD-L1	Non-small lung cancer	Tecentriq	Approved-2016	[43]
Bcl-2/Bax	Chronic lymphocytic leukemia	ABT-199	Approved-2016	[44]
PD-1/PD-L1	Merkel cell carcinoma	Bavencio	Approved-2017	[45]
PD-1/PD-L1	Non-small lung cancer	Imfinzi	Approved-2017	[46]
MDM2/p53	Acute myeloid leukemia	Idasanutlin	Phase III	[47]
MDM2/p53	Metastatic melanoma	AMG232	Phase I/II	[48]
MDM2/p53	Solid tumor with p53 wild type status	CGM097	Phase I	[49]
MDM2/p53	Advanced solid tumor, lymphoma	DS-3032b	Phase I	[50]
MDM2/p53	Neoplasm malignant	SAR405838	Phase I	[51]
MDM2/p53	Advanced solid tumors, lymphomas	ALRN-6924	Phase I/II	[52]
XIAP/caspase-9	Relapsed or refractory multiple myeloma	LCL-161	Phase II	[53]
XIAP/caspase-9	Recurrent head and neck squamous cell carcinoma	TL32711	Phase I	[54]
XIAP/caspase-9	Solid tumors, lymphoma	ASTX-660	Phase I/II	[55]
XIAP/caspase-9	Solid cancers	GDC-0917	Phase I	[56]
B-catenin/CBP	Liver cirrhosis	RPI-724	Phase I/II	[57]
PD-1/PD-L1	Prostatic neoplasms	CA-170	Phase II	[58]
CD40/CD40L	Advanced solid tumors	ABBV-428	Phase I	[59]

destruction and disposal of damaged or unwanted cells, remains an overlooked arsenal in designing anti-cancer therapies. Defects in the regulation of apoptosis have been associated with dysfunctional splicing patterns of a large number of apoptotic factors in tumorigenesis. Therefore, the modulation of anti-apoptotic and pro-apoptotic proteins via pharmaceutical manipulation of regulatory core-splicing proteins may open up new therapeutic avenues for the treatment of cancer. Thus, targeting the interaction between SNRPG and the RING finger domain of RBBP6 for drug discovery is a "bottom-up" approach in addressingthe issues surrounding tumorigenesis and tumor development.

Conclusion

The panoply of using MST technology to detect and quantify high-affinity and therapeutically relevant PPIs towards PPI-focused drug discovery remains of particular interest in biomedical research. Deciphering the binding affinity of therapeutic proteins in cancer is essential for their rational optimization towards designing PPI-focused anti-cancer adjuvants. The MST analysis presented in this study provides the first mechanistic *in vitro* insight of the inter- action between SNRPG and the RING finger domain of RBBP6. The obtained results are coherent and in perfect agreement with previous suggestions implicating possible involvement between the two proteins. The study affirms and strongly establishes scientific pursuit insinuating the possible *in vivo* involvement between the two regulatory core-splicing proteins in cancer-cell networks. The study strongly disqualifies the oversight placed on the two onco-proteins in developing PPI-focus- ed smart drugs, thus showing SNRPG and the RING finger domain of RBBP6 as potential anti-cancer drug candidates. Further exploration into their molecular and structural mechanism of action could significantly validate their effi- cacy as potential PPI-focused anti-cancer drug- gable targets. Identifying small molecule 'lead' compounds capable of modulating the interaction between SNRPG and the RING finger domain of RBBP6 could be the 'missing link' in the puzzle of the "quest for the cure".

Acknowledgements

Lloyd Mabonga and Priscilla Masamba are thankful to the Department of Science and Innovation (DSI) and the National Research Foundation (NRF) of South Africa for a Doctor- al and Postdoctoral Fellowship respectively. Abidemi Paul Kappo is grateful for support

provided by the University of Johannesburg University Research Committee.

Disclosure of conflict of interest

None.

Address correspondence to: Abidemi Paul Kappo, Molecular Biophysics and Structural Biology (MB- SB) Group, Department of Biochemistry, C2 Lab Room 426, University of Johannesburg, Kingsway Campus, Auckland Park 2006, South Africa. Tel: +27-11-559-3182; Fax: +27-11-559-2605; E-mail: akappo@uj.ac.za

References

[1] Plach MG, Grasser K and Schubert T. MicroScale Thermophoresis as a tool to study protein-peptide interactions in the context of large eukaryotic protein complexes. Bio Protoc 2017; 7: e2632.

[2] Rainard JM, Pandarakalam GC and McElroy SP. Using microscale thermophoresis to characterize hits from high-throughput screening: a European lead factory perspective. SLAS Discov 2018; 23: 225-241.

[3] Wienken CJ, Baaske P, Rothbauer U, Braun D and Duhr S. Protein-binding assays in biological liquids using microscale thermophoresis. Nat Commun 2010; 1: 100.

[4] Nimmagadda A, Shi Y and Cai J. Y-AApeptides as a new strategy for therapeutic development. Curr Med Chem 2019; 26: 2313-2329.

[5] Jonsson PF, Cavanna T, Zicha D and Bates PA. Cluster analysis of networks generated through homology: automatic identification of important protein communities involved in cancer metastasis. BMC Bioinformatics 2006; 7: 2.

[6] Steinbrecher T and Labahn A. Towards accurate free energy calculations in ligand protein-binding studies. Curr Med Chem 2010; 17: 767-785.

[7] Bhandari GP, Angdembe MR, Dhimal M, Neupane S and Bhusal C. State of non-communicable diseases in Nepal. BMC Public Health 2014; 14: 23.

[8] Heneghan C, Blacklock C, Perera R, Davis R, Banerjee A, Gill P, Liew S, Chamas L, Hernandez J and Mahtani K. Evidence for non-communicable diseases: analysis of Cochrane reviews and randomised trials by world bank classification. BMJ Open 2013; 3: e003298.

[9] Conte N, Charafe-Jauffret E, Delaval B, Adélaïde J, Ginestier C, Geneix J, Isnardon D, Jacquemier J and Birnbaum D. Carcinogenesis and translational controls: TACC1 is down-regulated in human cancers and associates with mRNA regulators. Oncogene 2002; 21: 5619-5630.

[10] Shi Y, Di Giammartino DC, Taylor D, Sarkeshik A, Rice WJ, Yates JR 3rd, Frank J and Manley JL. Molecular architecture of the human pre-mRNA 3' processing complex. Mol Cell 2009; 33: 365-376.

[11] Lührmann R, Kastner B and Bach M. Structure of spliceosomal snRNPs and their role in pre-mRNA splicing. Biochim Biophys Acta 1990; 1087: 265-292.

[12] Palfi Z, Lücke S, Lahm HW, Lane WS, Kruft V, Bragado-Nilsson E, Séraphin B and Bindereif A. The spliceosomal snRNP core complex of Trypanosoma brucei: cloning and functional analysis reveals seven Sm protein constituents. Proc Natl Acad Sci U S A 2000; 97: 8967-8972.

[13] Stevens SW and Abelson J. Purification of the yeast U4/U6.U5 small nuclear ribonucleoprotein particle and identification of its proteins. Proc Natl Acad Sci U S A 1999; 96: 7226-7231.

[14] Will CL and Lührmann R. Spliceosome structure and function. Cold Spring Harb Perspect Biol 2011; 3: a003707.

[15] Hull R, Oosthuysen B, Cajee UF, Mokgohloa L, Nweke E, Antunes RJ, Coetzer TH and Ntwasa M. The drosophila retinoblastoma binding protein 6 family member has two isoforms and is potentially involved in embryonic patterning. Int J Mol Sci 2015; 16: 10242-10266.

[16] Kappo MA, Eiso A, Hassem F, Atkinson RA, Faro A, Muleya V, Mulaudzi T, Poole JO, McKenzie JM, Chibi M, Moolman-Smook JC, Rees DJ and Pugh DJ. Solution structure of RING finger-like domain of retinoblastoma-binding protein-6 (RBBP6) suggests it functions as a U-box. J Biol Chem 2012; 287: 7146-7158.

[17] Pugh DJ, Eiso A, Faro A, Lutya PT, Hoffmann E and Rees DJ. DWNN, a novel ubiquitin-like domain, implicates RBBP6 in mRNA processing and ubiquitin-like pathways. BMC Struct Biol 2006; 6: 1.

[18] Khan F, Allam M, Tincho MB and Pretorius A. Implications of RBBP6 in various types of cancer. Proceedings IWBBIO 2014.

[19] Li L, Deng B, Xing G, Teng Y, Tian C, Cheng X, Yin X, Yang J, Gao X, Zhu Y, Sun Q, Zhang L, Yang X and He F. PACT is a negative regulator of p53 and essential for cell growth and embryonic development. Proc Natl Acad Sci U S A 2007; 104: 7951-7956.

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

[21] Chibi M, Meyer M, Skepu A, G Rees DJ, Moolman-Smook JC and Pugh DJ. RBBP6 interacts with multifunctional protein YB-1 through its RING finger domain, leading to ubiquitination and proteosomal degradation of YB-1. J Mol Biol 2008; 384: 908-916.
[22] Mabonga L and Kappo AP. The oncogenic potential of small nuclear ribonucleoprotein polypeptide G: a comprehensive and perspective view. Am J Transl Res 2019; 11: 6702-6716.
[23] Mrozowich T, Meier-Stephenson V and Patel TR. Microscale thermophoresis: warming up to a new biomolecular interaction technique. Biochemist 2019; 41: 8-12.
[24] Bartoschik T, Galinec S, Kleusch C, Walkiewicz K, Breitsprecher D, Weigert S, Muller YA, You C, Piehler J, Vercruysse T, Daelemans D and Tschammer N. Near-native, site-specific and purification-free protein labeling for quantitative protein interaction analysis by MicroScale Thermophoresis. Sci Rep 2018; 8: 4977.
[25] Berleth M, Berleth N, Minges A, Hänsch S, Burkart RC, Stork B, Stahl Y, Weidtkamp-Peters S, Simon R and Groth G. Molecular analysis of protein-protein interactions in the ethylene pathway in the different ethylene receptor subfamilies. Front Plant Sci 2019; 10: 726.
[26] Schubert T and Längst G. Studying epigenetic interactions using microscale thermophoresis (MST). AIMS Biophys 2015; 2: 370-380.
[27] Bekic I, Molnar M and Tschammer N. Protein labeling-improved quantitation of biomolecular interactions by MST using the His-Tag labeling kit RD-tris-NTA 2nd generation. 2018.
[28] Lata S, Reichel A, Brock R, Tampé R and Piehler J. High-affinity adaptors for switchable recognition of histidine-tagged proteins. J Am Chem Soc 2005; 127: 10205-10215.
[29] Chen DS and Mellman I. Oncology meets immunology: the cancer-immunity cycle. Immunity 2013; 39: 1-10.
[30] Rueckert C and Guzmán CA. Vaccines: from empirical development to rational design. PLoS Pathog 2012; 8: e1003001.
[31] Ito T, Ota K, Kubota H, Yamaguchi Y, Chiba T, Sakuraba K and Yoshida M. Roles for the two-hybrid system in exploration of the yeast protein interactome. Mol Cell Proteomics 2002; 1: 561-566.
[32] Phizicky EM and Fields S. Protein-protein interactions: methods for detection and analysis. Microbiol Rev 1995; 59: 94-123.
[33] Salahudeen MS and Nishtala PS. An overview of pharmacodynamic modelling, ligand-binding approach and its application in clinical practice. Saudi Pharm J 2017; 25: 165-175.
[34] Kenakin T and Primer AP. Techniques for more effective and strategic drug discovery. Amsterdam, The Netherlands: Academic Press; 2014.
[35] Magnez R, Thiroux B, Taront S, Segaoula Z, Quesnel B and Thuru X. PD-1/PD-L1 binding studies using microscale thermophoresis. Sci Rep 2017; 7: 17623.
[36] Liberelle M, Magnez R, Thuru X, Bencheikh Y, Ravez S, Quenon C, Drucbert AS, Foulon C, Melnyk P, Van Seuningen I and Lebègue N. MUC4-ErbB2 oncogenic complex: binding studies using microscale thermophoresis. Sci Rep 2019; 9: 16678.
[37] Mellman I, Coukos G and Dranoff G. Cancer immunotherapy comes of age. Nature 2011; 480: 480-489.
[38] O'Shea JJ, Kontzias A, Yamaoka K, Tanaka Y and Laurence A. Janus kinase inhibitors in autoimmune diseases. Ann Rheum Dis 2013; 72 Suppl 2: ii111-ii115.
[39] Sharma P and Allison JP. The future of immune checkpoint therapy. Science 2015; 348: 56-61.
[40] Sunshine J and Taube JM. Pd-1/pd-l1 inhibitors. Curr Opin Pharmacol 2015; 23: 32-38.
[41] Reck M, Rodríguez-Abreu D, Robinson AG, Hui R, Csőszi T, Fülöp A, Gottfried M, Peled N, Tafreshi A and Cuffe S. Pembrolizumab versus chemotherapy for PD-L1-positive non-small-cell lung cancer. N Engl J Med 2016; 375: 1823-1833.
[42] Borghaei H, Paz-Ares L, Horn L, Spigel DR, Steins M, Ready NE, Chow LQ, Vokes EE, Felip E, Holgado E, Barlesi F, Kohlhäufl M, Arrieta O, Burgio MA, Fayette J, Lena H, Poddubskaya E, Gerber DE, Gettinger SN, Rudin CM, Rizvi N, Crinò L, Blumenschein GR Jr, Antonia SJ, Dorange C, Harbison CT, Graf Finckenstein F and Brahmer JR. Nivolumab versus docetaxel in advanced nonsquamous non-small-cell lung cancer. N Engl J Med 2015; 373: 1627-1639.
[43] Socinski MA, Jotte RM, Cappuzzo F, Orlandi F, Stroyakovskiy D, Nogami N, Rodríguez-Abreu D, Moro-Sibilot D, Thomas CA and Barlesi F. Atezolizumab for first-line treatment of metastatic nonsquamous NSCLC. N Engl J Med 2018; 378: 2288-2301.
[44] Korycka-Wolowiec A, Wolowiec D, Kubiak-Mlonka A and Robak T. Venetoclax in the treatment of chronic lymphocytic leukemia. Expert Opin Drug Metab Toxicol 2019; 15: 353-366.
[45] Boyerinas B, Jochems C, Fantini M, Heery CR, Gulley JL, Tsang KY and Schlom J. Antibody-dependent cellular cytotoxicity activity of a novel anti-PD-L1 antibody avelumab (MSB-0010718C) on human tumor cells. Cancer Immunol Res 2015; 3: 1148-1157.
[46] Antonia SJ, Villegas A, Daniel D, Vicente D, Murakami S, Hui R, Yokoi T, Chiappori A, Lee KH and de Wit M. Durvalumab after chemoradiotherapy in stage III non-small-cell lung cancer. N Engl J Med 2017; 377: 1919-1929.

[47] Lehmann C, Friess T, Birzele F, Kiialainen A and Dangl M. Superior anti-tumor activity of the MDM2 antagonist idasanutlin and the Bcl-2 inhibitor venetoclax in p53 wild-type acute myeloid leukemia models. J Hematol Oncol 2016; 9: 50.

[48] Sun D, Li Z, Rew Y, Gribble M, Bartberger MD, Beck HP, Canon J, Chen A, Chen X, Chow D, Deignan J, Duquette J, Eksterowicz J, Fisher B, Fox BM, Fu J, Gonzalez AZ, Gonzalez-Lopez De Turiso F, Houze JB, Huang X, Jiang M, Jin L, Kayser F, Liu JJ, Lo MC, Long AM, Lucas B, McGee LR, McIntosh J, Mihalic J, Oliner JD, Osgood T, Peterson ML, Roveto P, Saiki AY, Shaffer P, Toteva M, Wang Y, Wang YC, Wortman S, Yakowec P, Yan X, Ye Q, Yu D, Yu M, Zhao X, Zhou J, Zhu J, Olson SH and Medina JC. Discovery of AMG 232, a potent, selective, and orally bioavailable MDM2-p53 inhibitor in clinical development. J Med Chem 2014; 57: 1454-1472.

[49] Holzer P, Masuya K, Furet P, Kallen J, Valat-Stachyra T, Ferretti S, Berghausen J, Bouisset-Leonard M, Buschmann N, Pissot-Soldermann C, Rynn C, Ruetz S, Stutz S, Chène P, Jeay S and Gessier F. Discovery of a dihydroisoquinolinone derivative (NVP-CGM097): a highly potent and selective MDM2 inhibitor undergoing phase 1 clinical trials in p53wt tumors. J Med Chem 2015; 58: 6348-6358.

[50] Arnhold V, Schmelz K, Proba J, Winkler A, Wünschel J, Toedling J, Deubzer HE, Künkele A, Eggert A and Schulte JH. Reactivating TP53 signaling by the novel MDM2 inhibitor DS-3032b as a therapeutic option for high-risk neuroblastoma. Oncotarget 2018; 9: 2304-2319.

[51] De Weger V, Lolkema M, Dickson M, Le Cesne A, Wagner A, Merqui-Roelvink M, Varga A, Tap W, Schwartz G and Demetri G. 378 A first-in-human (FIH) safety and pharmacological study of SAR405838, a novel HDM2 antagonist, in patients with solid malignancies. Eur J Cancer 2014; 50: 121-122.

[52] Carvajal LA, Neriah DB, Senecal A, Benard L, Thiruthuvanathan V, Yatsenko T, Narayanagari SR, Wheat JC, Todorova TI, Mitchell K, Kenworthy C, Guerlavais V, Annis DA, Bartholdy B, Will B, Anampa JD, Mantzaris I, Aivado M, Singer RH, Coleman RA, Verma A and Steidl U. Dual inhibition of MDMX and MDM2 as a therapeutic strategy in leukemia. Sci Transl Med 2018; 10: eaao3003.

[53] West A, Martin BP, Andrews D, Hogg S, Banerjee A, Grigoriadis G, Johnstone R and Shortt J. The SMAC mimetic, LCL-161, reduces survival in aggressive MYC-driven lymphoma while promoting susceptibility to endotoxic shock. Oncogenesis 2016; 5: e216.

[54] Benetatos CA, Mitsuuchi Y, Burns JM, Neiman EM, Condon SM, Yu G, Seipel ME, Kapoor GS, LaPorte MG, Rippin SR, Deng Y, Hendi MS, Tirunahari PK, Lee YH, Haimowitz T, Alexander MD, Graham MA, Weng D, Shi Y, McKinlay MA and Chunduru SK. Birinapant (TL32711), a bivalent SMAC mimetic, targets TRAF2-associated cIAPs, abrogates TNF-induced NF-κB activation, and is active in patient-derived xenograft models. Mol Cancer Ther 2014; 13: 867-879.

[55] Ward GA, Lewis EJ, Ahn JS, Johnson CN, Lyons JF, Martins V, Munck JM, Rich SJ, Smyth T, Thompson NT, Williams PA, Wilsher NE, Wallis NG and Chessari G. ASTX660, a novel non-peptidomimetic antagonist of cIAP1/2 and xIAP, potently induces TNFa-dependent apoptosis in cancer cell lines and inhibits tumor growth. Mol Cancer Ther 2018; 17: 1381-1391.

[56] Wong H, Gould SE, Budha N, Darbonne WC, Kadel EE, La H, Alicke B, Halladay JS, Erickson R and Portera C. Learning and confirming with preclinical studies: modeling and simulation in the discovery of GDC-0917, an inhibitor of apoptosis proteins antagonist. Drug Metab Dispos 2013; 41: 2104-2113.

[57] Kimura K, Ikoma A, Shibakawa M, Shimoda S, Harada K, Saio M, Imamura J, Osawa Y, Kimura M, Nishikawa K, Okusaka T, Morita S, Inoue K, Kanto T, Todaka K, Nakanishi Y, Kohara M and Mizokami M. Safety, tolerability, and preliminary efficacy of the anti-fibrotic small molecule PRI-724, a CBP/p-catenin inhibitor, in patients with hepatitis C virus-related cirrhosis: a single-center, open-label, dose escalation phase 1 trial. EBioMedicine 2017; 23: 79-87.

[58] Musielak B, Kocik J, Skalniak L, Magiera-Mularz K, Sala D, Czub M, Stec M, Siedlar M, Holak TA and Plewka J. CA-170-a potent small-molecule PD-L1 inhibitor or not? Molecules 2019; 24: 2804.

[59] Ye S, Cohen D, Belmar NA, Choi D, Tan SS, Sho M, Akamatsu Y, Kim H, Iyer R and Cabel J. A bispecific molecule targeting CD40 and tumor antigen mesothelin enhances tumor-specific immunity. Cancer Immunol Res 2019; 7: 1864-1875.

Table S1. MST raw data of merged dose-response of the PPIs between SNRPG and RING finger do-main of RBBP6

Dose	Response (Average)	Std. Dev.	N
1.11E-06	867.28155	0.9235	3
5.55E-07	866.9055	0.26835	2
2.775E-07	865.66652	1.66455	2
1.3875E-07	865.42005	0.81742	2
6.9375E-08	864.91562	2.33324	2
3.46875E-08	866.01389	1.33826	3
1.734375E-08	862.59946	1.25891	2
8.671875E-09	862.22974	0.74229	2
4.335938E-09	861.75527	4.2119	2
2.167969E-09	860.03	5.31867	2
1.083984E-09	861.10854	0.89908	2
5.41992E-10	862.19	0.63159	2
2.70996E-10	862.75342	1.93067	3
1.35498E-10	861.43567	2.59197	3
6.7749E-11	861.44066	1.97434	3
3.3875E-11	859.30191	1.3296	3

CHAPTER SEVEN

General Discussion and Conclusion

7.1 General Discussion

For several years, tremendous efforts have been linked to finding solutions to address existing cancer challenges. Modern drug discovery is driven by molecular targets with the aim of identifying new therapeutic agents that can selectively target disease-specific molecular mechanisms or pathways (Díaz-Eufracio *et al.*, 2018a). Novel techniques to design and develop cancer therapy drugs and treatment methods using protein-protein interaction (PPIs) studies have become promising targets for therapeutic discovery. There are more than 645,000 reported disease-relevant PPIs in the human interactome. However, only 2% of these have been targeted with drugs by the year 2011. Most of the remaining disease-relevant PPIs such as transcription factors and many other signaling proteins have been widely considered 'undruggable' and remain elusive, under-explored and yet to be fully understood (Díaz-Eufracio *et al.*, 2018b; González-Abuín *et al.*, 2012; Robertson and Spring, 2018; Zhang *et al.*, 2018).

Inhibiting PPIs using small molecules or peptides has diagnostic and therapeutic significance (Díaz-Eufracio *et al.*, 2018b; Beloglazkina *et al.*, 2020; Wu *et al.*, 2019). In this context, PPIs are an attractive emerging class of molecular targets that are critically important in the progression of many disease states (Robertson and Spring, 2018; Zhang *et al.*, 2018). If well exploited, they can be engineered to provide therapeutically tractable ways of tweaking and manipulating the interplay in order to address the progression of many disease states (Du *et al.*, 2018). Inhibiting

PPIs is a tremendously important diagnostic and therapeutic strategy that may lead to greatly protracted remissions and even curative therapies for a number of diseases (Stevers *et al.*, 2017).

The emergence of new technologies has unveiled the potential of PPIs in drug discovery and has enabled regular discovery of small molecule PPI modulators as significant smart-drug targets (Jana *et al.*, 2017; Pelay-Gimeno *et al.*, 2015). The efficient mimicking of peptides in their bioactive conformation is a long-standing goal in the design of PPI inhibitors as drugs. Advances in PPI-focused technology have facilitated a display of side chain functionalities in analogy to peptide secondary structures, yielding molecules that are generally referred to as peptidomimetics (Pelay-Gimeno *et al.*, 2015). Peptidomimetics are compounds whose essential elements (pharmacophore) mimic a natural peptide or protein in 3D space and which retain the ability to interact with the biological target and produce the same biological effect (Akram *et al.*, 2014).

Peptidomimetics tend to mimic peptide side chains to take advantage of the binding affinity of a number of hot-spot residues. The use of peptidomimetics has recently come of age with new drugs going into clinical trials. Research into these compounds will continue to be an indispensable tool to target PPIs in drug discovery for the foreseeable future (Akram *et al.*, 2014; Mabonga and Kappo, 2019a). In this context, SNRPG (also referred to as SmG in this study) and the RING finger domain of RBBP6 have attracted significant attention because of their implicated roles in tumorigenesis and tumor development. Suggestive evidence of their varying expression levels has been reported in different types of cancers, which include breast, lung, prostate and colon cancer. The accumulating evidence suggests that the two splicing machinery components play significant roles in the initiation and progression of cancers (Mabonga and Kappo, 2019b).

Additionally, SNRPG and the RING finger domain of RBBP6 have a wide interacting network and their functions are predominantly mediated by protein-protein interactions, making them promising anti-cancer therapeutic targets in PPI-focused drug technology. According to Jonsson and co-workers most cancer-implicated proteins possess structural domains that have a higher ratio of infidelity as compared to non-cancer implicated proteins, making them more prone to interact with a wide diversity of proteins (Jonsson *et al.,* 2006). Cancer-implicated proteins have a wide interactome and occupy central positions in cancer-cell protein networks (Bhandari *et al.*, 2014; Heneghan *et al.*, 2013; Steinbrecher and Labahn, 2010). Moreover, protein interactions between cancer-implicated proteins have a higher probability of being related to cancer processes than non-interacting proteins (Du *et al.,* 2018; Jonsson *et al.*, 2006; Du *et al.*, 2016; Hanlon *et al.*, 2010).

In this regard, SNRPG and the RING finger domain of RBBP6 have been identified as potential diagnostic markers for cancer treatment. Understanding their roles in tumorigenesis and tumor development is an indispensable arsenal in the development of molecular-targeted therapies. Several anti-tumor drugs linked to splicing machinery components have been reported in different types of cancers and some have already entered the clinical setting. However, targeting SNRPG and the RING finger domain of RBBP6 as drug development tools has been an overlooked and under-developed strategy in cancer therapy. Suggestive evidence proposed putative interactions between the two regulatory core-splicing oncogenic proteins (Chibi *et al.*, 2008; Kappo *et al.*, 2012). Chibi and co-workers predicted possible interactions between SNRPG and RBBP6 through its N-terminal domain which is a crucial component of the RNA processing machinery in the cell (Chibi *et al.*, 2008). The N-terminal domain of RBBP6 consists of the Domain With No Name (DWNN), the Zinc knuckle and Really Interesting New Gene (RING)

finger domain. However, the precise mechanisms and binding events involved remain elusive and yet to be fully investigated.

In another study, Kappo and co-workers identified two copies of SNRPG, which are conformational isomers of the same protein as part of the five substrates that bind to the N-terminal RING finger domain of RBBP6 (Kappo *et al.*, 2012). The results obtained from the two studies by Chibi and co-workers (Chibi *et al.*, 2008) as well as Kappo and colleagues (Kappo *et al.*, 2012) form part of the background of the study conducted in this work. The findings from the two studies substantiated the possible interaction between SNRPG and RBBP6 through its RING finger domain in tumorigenesis and tumour development. However, the precise mechanisms involved remain elusive, hence the need to conduct further investigations and establish the relationship between the two cancer proteins towards PPI-focused anti-cancer drug discovery.

In this study, *in silico* studies were conducted to predict the structural and biophysical characteristics of SNRPG and the RING finger domain of RBBP6. The putative interaction between the two proteins was predicted and a possible lead compound that can be used as a small molecule inhibitor towards anti-cancer drug discovery was further identified. The *in silico* studies predicted significant interactions between SNRPG and the RING finger domain of RBBP6. Detailed analysis of the predicted three-dimensional structure of the complex between SNRPG and the RING finger domain of RBBP6 identified "hot spot" regions and binding site residues active in the binding. The determination of *in silico* predictions as determined from Racoon and Autodock Vina provided the fundamental motivation to further explore this interaction using the experimental MST assay. The assay provided a plumb line to determine the thermodynamic parameters of the binding between the two proteins with the view of

extrapolating the findings towards designing PPI-focused smart-drug modulators (Salmaso and Moro, 2018).

The study predicted and identified a potential ligand (4FI), which was subsequently docked onto the SNPRG~RING finger domain protein complex. The predicted binding site residues on the active site of the SNPRG~RING finger domain protein complex when bound to the 4FI ligand denoted more than 71 % participation of hydrophobic amino acid residues bearing non-polar R-groups, 14% hydrophilic amino acid residues with non-polar R-groups and 14% basic and electrically charged amino acid residues possessing positively charged R-groups. A high level of fluctuation occurred in the 4FI~SNRPG~RING finger domain complex at around 15 ns, stabilizing at approximately 40 ns till the end of the simulation, thereby suggesting the adjustment of the ligand within the active site of the complex.

The docked pose of SNRPG in complex with the RING finger domain of RBBP6 was prepared and subjected to MD simulations using AMBER18 program. MD simulations were performed for a total of 80 ns and trajectories were analyzed for ligand root-mean-square-deviation (RMSD), root mean square fluctuation (RMSF) and radius of gyration (RoG) values and interaction fingerprints of the compounds in the active site over time. The results suggested that the 4FI ligand reshapes the sorting signal binding groove in a way that optimal van der Waals contacts are possible and non-polar solvent accessible surface area is diminished upon peptide binding. The opening of the dynamic loop is particularly crucial in establishing the optimized 4FI ligand binding groove conformation. The RMSF results showed that the core of the protein is more rigid relative to the loops (solvent-exposed) with very high fluctuations. The simulation denotes the stabilization of the SNRPG~RING finger domain complex throughout the simulation and the overall 4FI~SNRPG~RING finger domain complex does not deviate much from the

initial structure. Superimposition of the 3D structure of the SNRPG~RING finger domain complex before and after simulation also reveals no major changes in terms of the secondary structure.

In the MM/GBSA approach the binding affinity of the 4FI ligand bound to the complex between SNRPG and the RING finger domain of RBBP6 was -27.96 kcal/mol. Detailed structural analysis showed that the peptide binding groove is much more flexible in the SNRPG~RING finger domain complex structure as compared to 4FI~SNRPG~RING finger domain complex. This suggests that the peptide is not able to adopt stable bound conformations. The obtained ΔG_{bind} value was consistent with the other observations. The study determined a -0.01 kcal/mol prediction value of the Gibbs energy difference ($-T\Delta S_{tot}$) suggesting that the reaction between the 4FI ligand and the SNRPG~RING finger domain complex was predictively spontaneous and favors the forward reaction.

Microscale thermophoresis (MST) analysis was used to determine the binding affinity and binding strength of the biophysical interaction between SNRPG and the RING finger domain of RBBP6. Detailed procedures on the expression and purification of the proteins used in the MST assay are embedded in Appendix I. The results of the binding affinity between the two proteins were a K_D value of 3.1596 nM under aqueous buffer conditions. The successful binding between SNRPG and the RING finger domain of RBBP6 is in alignment with the *in silico* studies, which predicted that the interaction between the two proteins is feasible. The results confirmed the findings obtained from the previous research using the Y2H technique (Chibi *et al.*, 2008; Kappo *et al.*, 2012). This study provides substantiated evidence that the two cancer implicated proteins indeed interact *in vivo* and may participate actively in processes leading to tumorigenesis and tumor development. The findings help to project the oncogenic potential of

SNRPG and its possible links to the RING finger domain of RBBP6 in cancer-cell protein networks. From a drug discovery perspective, SNRPG and the RING finger domain of RBBP6 are potential targets in PPI-focused smart-drug discovery. Inhibiting the SNRPG~RING finger domain complex using small molecule inhibitors may lead to greatly protracted remissions and even curative therapies for several diseases including cancer (Mabonga and Kappo, 2019b).

7.2 Conclusion

Accelerating the arrival of next-generation drugs is a grand challenge awaiting us. A concerted effort to target undruggable proteins may reap enormous benefits in terms of yielding new medicines for diverse human diseases. The potential of the proteome space as a drug discovery arsenal may unlock and bring under control all manner of pathologies, from neurodegenerative diseases to cancer and to infectious diseases. Researchers are becoming aware of the great benefits that could accrue by tackling these undruggable proteins towards PPI-focused anti-cancer drug design. Therapeutic peptidomimetics are at the forefront in the technology development curve as potential PPI-focused next generation drugs. A significant number of laboratories are trying to create collections of architecturally complex synthetic mimicries that might be more effective in binding to these challenging and undruggable protein families. In the course of time, an increasing number of proteins historically branded intractable will likely be considered druggable. The time may even come, far into the future, when all disease-relevant proteins will have been targeted successfully as potential adjuvants. Thus, the final panorama of human ingenuity will rapture as we reach the pinnacle of curing the incurable - when the majority of undruggable proteins will have been drugged.

This work describes for the first time the unprecedented evidence of the longstanding postulations of the PPIs between SNRPG and the RING finger domain of RBBP6. The critical roles played by the two proteins in tumorigenesis and tumour development are well documented. However, their oncogenic potential has been overlooked and downplayed for a long time. Regardless of the tremendous efforts to uncover their biological significance, SNRPG and the RING finger domain of RBBP6 remained beyond the targeting capabilities of established drug discovery technologies. Validated by overwhelming evidence of high-quality biological data, the study unleashes new information on the interaction between SNRPG and the RING finger domain of RBBP6. The study exhumes the two onco-proteins from the force majeure of being undruggable targets to the fascio effect of being therapeutic vulnerabilities in the drug discovery parlance. Many of the splicing machinery components remain attractive as drug targets from a biological perspective. Regardless of being previously regarded as "prototypically intractable", "difficult to drug" or "yet to be drugged", results from this study provide the first evidence that affirm the druggability of the interaction between SNRPG and the RING finger domain of RBBP6.

7.3 Recommendations for further studies

Despite the strong and interesting implications associated with SNRPG and the RING finger domain of RBBP6 as well as their significant prowess as potential smart-drug discovery targets in PPI-focused anti-cancer studies, the oncogenic potential of the two proteins remains yet to be proven by fully investigating the mechanisms and functional basis of their operations. Their evasive roles in linking the splicing machinery to tumorigenesis and tumour development remain a dark spot and yet to be tamed. The findings presented in this study offer a foundational

platform to prompt further investigations into understanding the mechanism of action between the two proteins.

The study recommends that further *in silico* studies be conducted on the putative interaction between the two proteins with the aim of validating the amino acid residues involved in the binding between SNRPG and the RING finger domain of RBBP6. However, instead of running the MD simulation for only 80 ns, it is recommended that the simulations be conducted for a period longer than 100 ns to enable more projected results of the complex-ligand interaction. It is additionally proposed that the structure of SNRPG~RING finger domain be solved using macromolecular X-ray crystallography studies, the binding (hot spot) residues be identified, and mutational studies conducted.

In conclusion, the study recommends that multiple MST assays be conducted at varying parameter ranges to facilitate the determination of thermodynamic parameters such as enthalpy and entropy changes of the binding event. More interaction assays between SNRPG and other therapeutically relevant regulatory splicing proteins are recommended. More so, interaction studies between SNRPG and other N-terminal domains of RBBP6 should be investigated. Another aspect of the work that should be further investigated is to synthesize the 4FI ligand and perform an additional MST assay to further validate the *in silico* binding studies done on the SNRPG~RING finger domain complex. This would give more weight to the results and determine if in reality there is interaction between the proposed inhibitor and the protein complex. The study also recommends that other techniques that include Isothermal Titration Calorimetry (ITC) analysis and Surface Plasmon Resonance (SPR) analysis be conducted in order to validate the interactions reported in this study.

It is further suggested that the identification and selection of additional inhibitory ligands to facilitate further research towards drug design and development. Above all, future studies should by all means consider synthesizing the imhibitor for ex-vivo and pre-clinical studies using cell lines and animal models or synthesize and test 4FI-derived compounds. Finally, it is recommended that the design and development of PPI-focused small molecule drugs be derived from the 4FI 'lead' compound as potential arsenals for the inhibition of SNRPG~RING finger domain interaction in PPI-focused anti-cancer drug discovery.

REFERENCES

Akram ON, Degraff DJ, Sheehan JH, Tilley WD, Matusik RJ, Ahn J-M and Raj GV (2014). Tailoring peptidomimetics for targeting protein–protein interactions. *Molecular Cancer Research* **12**(7): 967-978.

Beloglazkina A, Zyk N, Majouga A and Beloglazkina E (2020). Recent small-molecule inhibitors of the p53–MDM2 protein–protein interaction. *Molecules* **25**(5): 1211.

Bhandari GP, Angdembe MR, Dhimal M, Neupane S and Bhusal C (2014). State of non-communicable diseases in Nepal. *BMC Public Health* **14**(1): 1-9.

Chibi M, Meyer M, Skepu A, Rees DJG, Moolman-Smook JC and Pugh DJ (2008). RBBP6 interacts with multifunctional protein YB-1 through its RING finger domain, leading to ubiquitination and proteosomal degradation of YB-1. *Journal of molecular biology* **384**(4): 908-916.

Díaz-Eufracio BI, Naveja JJ and Medina-Franco JL (2018a). Protein–protein interaction modulators for epigenetic therapies *Advances in protein chemistry and structural biology* (Vol. 110.): Elsevier pp. 65-84.

Díaz-Eufracio BI, Naveja JJ and Medina-Franco JL (2018b). Protein–protein interaction modulators for epigenetic therapies. *Advances in protein chemistry and structural biology* **110**: 65-84.

Du L, Grigsby SM, Yao A, Chang Y, Johnson G, Sun H and Nikolovska-Coleska Z (2018). Peptidomimetics for targeting protein–protein interactions between DOT1L and MLL oncofusion proteins AF9 and ENL. *ACS Medicinal Chemistry Letters* **9**(9): 895-900.

Du X, Li Y, Xia Y-L, Ai S-M, Liang J, Sang P, Ji X-L and Liu S-Q (2016). Insights into protein–ligand interactions: mechanisms, models, and methods. *International journal of molecular sciences* **17**(2): 144.

GonzáLez-Abuín N, Martínez-Micaelo N, Blay M, Pujadas G, Garcia-Vallvé S, Pinent M and ArdéVol A (2012). Grape seed-derived procyanidins decrease dipeptidyl-peptidase 4 activity and expression. *Journal of Agricultural and Food Chemistry* **60**(36): 9055-9061.

Hanlon L, Avila JL, Demarest RM, Troutman S, Allen M, Ratti F, Rustgi AK, Stanger BZ, Radtke F and Adsay V (2010). Notch1 functions as a tumor suppressor in a model of K-ras–induced pancreatic ductal adenocarcinoma. *Cancer Research* **70**(11): 4280-4286.

Heneghan C, Blacklock C, Perera R, Davis R, Banerjee A, Gill P, Liew S, Chamas L, Hernandez J and Mahtani K (2013). Evidence for non-communicable diseases: analysis of Cochrane reviews and randomised trials by World Bank classification. *BMJ open* **3**(7).

Jana T, Ghosh A, Das Mandal S, Banerjee R and Saha S (2017). PPIMpred: a web server for high-throughput screening of small molecules targeting protein–protein interaction. *Royal Society open science* **4**(4): 160501.

Jonsson PF, Cavanna T, Zicha D and Bates PA (2006). Cluster analysis of networks generated through homology: automatic identification of important protein communities involved in cancer metastasis. *BMC Bioinformatics* **7**(1): 1-13.

Kappo MA, Eiso A, Hassem F, Atkinson RA, Faro A, Muleya V, Mulaudzi T, Poole JO, Mckenzie JM and Chibi M (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**(10): 7146-7158.

Mabonga L and Kappo AP (2019a). Peptidomimetics: A Synthetic Tool for Inhibiting Protein–Protein Interactions in Cancer. *International Journal of Peptide Research and Therapeutics*: 1-17.

Mabonga L and Kappo AP (2019b). The oncogenic potential of small nuclear ribonucleoprotein polypeptide G: a comprehensive and perspective view. *American journal of translational research* **11**(11): 6702.

Pelay-Gimeno M, Glas A, Koch O and Grossmann TN (2015). Structure-based design of inhibitors of protein–protein interactions: mimicking peptide binding epitopes. *Angewandte Chemie International Edition* **54**(31): 8896-8927.

Robertson NS and Spring DR (2018). Using peptidomimetics and constrained peptides as valuable tools for inhibiting protein–protein interactions. *Molecules* **23**(4): 959.

Salmaso V and Moro S (2018). Bridging molecular docking to molecular dynamics in exploring ligand-protein recognition process: An overview. *Frontiers in Pharmacology* **9**: 923.

Steinbrecher T and Labahn A (2010). Towards accurate free energy calculations in ligand protein-binding studies. *Current Medicinal Chemistry* **17**(8): 767-785.

Stevers LM, Sijbesma E, Botta M, Mackintosh C, Obsil T, Landrieu I, Cau Y, Wilson AJ, Karawajczyk A and Eickhoff J (2017). Modulators of 14-3-3 protein–protein interactions. *Journal of Medicinal Chemistry* **61**(9): 3755-3778.

Wu K-J, Lei P-M, Liu H, Wu C, Leung C-H and Ma D-L (2019). Mimicking strategy for protein–protein interaction inhibitor discovery by virtual screening. *Molecules* **24**(24): 4428.

Zhang G, Andersen J and Gerona-Navarro G (2018). Peptidomimetics targeting protein-protein interactions for therapeutic development. *Protein and peptide letters* **25**(12): 1076-1089.

APPENDIX I

Experimental Procedures

General stock solutions, buffers and media

Ammonium persulphate (APS): 10 % stock solution of APS was prepared by dissolving 1 g of APS in 10 ml of distilled water; this was stored at -20 ℃.

Ampicillin: 100 µg/ml (final concentration) stock solution was prepared by dissolving 1 g of ampicillin in 10 ml of distilled water. The solution was filter-sterilized and divided into 2000 µl aliquots and stored at -80 ℃.

Ampicillin agar plates: 3.4 g Bacteriological agar, 3.2 g Tryptone, 2.0 g Yeast and 1.0 g NaCl was dissolved in 200 ml distilled water. This was autoclaved allowed to cool down before ampicillin was added.

Bis-Acrylamid: 29.2 g of acrylamide and 0.8 g of N, N-Methylene bis acrylamide was dissolved in 100 ml of distilled water.

Cell lysis buffer: 2.5 ml Tris (50 mM; pH 8.0), 100 µg/ml lysozyme, 250 µl PMSF (0.5 mM), 100 µg/ml DTT (5 mM) and 7.5 ml NaCl (100 mM). The buffer was made up to 50 ml by an addition of distilled water.

Chloramphenicol: A 34 mg/ml stock solution was prepared in ethanol. The solution was stored at 4 ℃ until needed.

Coomassie staining solution: 1 g Coomassie® Brilliant Blue was dissolved in 450 ml (45 %) methanol and 100 ml (10 %) glacial acetic acid and the volume brought to 1000 ml by the addition of 450 ml distilled water.

Destaining solution: 500 ml (50 %) methanol and 70 ml (7 %) glacial acetic acid were mixed and the volume was made up to 1000 ml by the addition of 430 ml distilled water.

Dithiothreitol (DTT): 1 M stock solution was prepared in 0.01 M sodium acetate, pH 5.2. The solution was filter-sterilized, divided into 1 ml aliquots and stored at -80 °C.

Isopropyl β-D-1-thiogalactopyranoside (IPTG): 1 M stock solution was prepared by dissolving 2.38 g of IPTG in 8 ml of double distilled water and brought to a final volume of 10 ml. This was divided into aliquots and kept at -20 °C.

PBS: this was prepared by the addition of 8 g/l NaCl, 0.2 g/l KCl, 1.44 g/l Na2HPO4 and 0.24 g/l KH2PO4, pH 7.4 and sterilized by autoclaving.

PMSF (100mM): 0.174 g of PMSF was dissolved in 10 ml of ethanol and stored at - 20º C.

2 M Tris (pH 8.8): 242.2 g of Tris base was dissolved in 800 ml of ddH2O. The pH was adjusted to 8.8 by adding concentrated HCl. The final volume was made up to 1 L with distilled water and the solution was sterilized by autoclaving.

2 M Tris (pH 6.8): 242.2 g of Tris base was dissolved in 800 ml of ddH2O. The pH was adjusted to 6.8 by adding concentrated HCl. The final volume was made up to 1 L with distilled water and the solution was sterilized by autoclaving.

YT Broth: this was prepared by the addition of 16 g/l Tryptone powder, 10 g/l Yeast extract, 5 g/l NaCl and 4 g/l Glucose. Autoclaved and stored at room temperature, normally used within three days.

2X sample buffer: 100 mM Tris-HCl pH 6.8, 4 % SDS, 20 % Glycerol and 0.2 % Bromophenol Blue. The buffer was stored at room temperature, prior to use, 200 mM freshly prepared DTT was added.

30 % glycerol: 30 ml of glycerol was mixed with 70 ml of distilled water 10 %

Sodium Dodecyl Sulphate (SDS): 50 g SDS was dissolved in 400 ml of distilled water. The pH was adjusted to 7.2 and the volume adjusted to 500 ml with distilled water. Filter-sterilize, if needed.

5X running buffer: 15.1 g Tris-base, 72 g Glycine and 10 % SDS were combined and the volume brought to one litre with distilled water.

15 % SDS resolving/ separating gel recipe: 2812 µl Bis, 1425 µl Tris 8.8, 28.13 (10 %) µl SDS, 1350 µl distilled water, 75 µl APS and 15 µl TEMED.

SDS stacking gel recipe: 650 µl Bis, 1250 µl Tris 6.8, 25 (10 %) µl SDS, 3000 µl distilled water, 100 µl APS and 10 µl TEMED.

Recombinant Expression and Purification of RING finger domain protein

Bacterial strain used

E. coli BL21 Star^{TM}pLysS (DE3) (Stratagene): Fomp T hsdS_B ($r_B^- m_B^-$) gal dcm rne131 (DE3)

Preparation of competent E. coli BL21 (DE3) pLysS cells for transformation

E. coli BL21 (DE3) pLysS cells were streaked out on a Luria agar plate and incubated at 37 ºC for 16 hours. A single colony was picked and used to inoculate 5 ml LB broth without antibiotics which was then incubated at 37 ºC overnight. The following morning, the culture was scaled up to 50 ml and incubated at 37 ºC until the optical density at 600 nm (OD_{600}) reached between 0.4–0.6. The cells were then harvested by centrifugation at 5000 rpm for 10 mins, re-suspended in 10 ml ice cold 0.1M $MgCl_2$ and left to incubate on ice for 30 mins. The cells were harvested again at 4000 rpm for 10 mins and the cells resuspended in 10 ml ice cold 0.1M $CaCl_2$. The cells were harvested again for the last time and resuspended in 3 ml ice cold CaCl2 and 3 ml 30% glycerol. The made competent cells were then aliquoted, snap frozen in liquid nitrogen and stored at -80 ºC until required for use.

Transformation of E. coli BL21 (DE3) pLysS cells using pQE30~SNRPG and pGEX-6P-2~RING constructs

Codon optimized cDNA encoding full-length of SNRPG (UniProtKB - P62308 (RUXG_HUMAN)) and RING Finger domain of RBBP6 (pdb3tzg) were purchased from GenScript (USA) according to the published sequences (NCBI ID: NM_001317165.1 and BVU_2266) and codon harmonized forms of the genes of SNRPG and RING Finger domain of RBBP6 were constructed. The DNA segments encoding regions were PCR amplified using 144 *BamHI* and *XhoI* restriction sites (due to their ability to use the same digestion buffer, hence restriction can be done once at 4 ºC overnight). The genes were cloned into the pQE30 and pGEX-6P-2 protein expression vectors respectively.

Frozen competent cells of the strain *E. coli* BL21 (DE3) *pLysS* were thawed on ice and to 100 µl of these cells, 2 µl of plasmid DNA was added to make a transformation mixture (designated the experimental tube). Another tube was set up with only competent cells and no plasmid DNA (designated the control tube). The mixtures were incubated on ice for 20 minutes after which the cells were heat-shocked by incubation in a water-bath for 5 minutes at 37 °C. Thereafter, the mixtures were placed back on ice for an additional 5 minutes after which 900 µl of pre-warmed Luria broth without ampicillin was added to both tubes, followed by a further incubation step at 37 °C for 2 hours. A 50 µl sample of the transformed cells were plated on Luria agar plates containing 100 µg/ml of ampicillin and incubated at 37 °C overnight.

Small-scale expression screening of E. coli BL21 (DE3) pLysS transformed cells with pQE30~SNRPG and pGEX-6P-2~RING constructs

Single colonies of *E. coli* BL21 (DE3) *pLysS* cells were used to inoculate 5 ml cultures of Luria broth containing 100 µg/ml ampicillin which were then incubated for 4 hours at 37 °C with vigorous shaking. Thereafter, 1 ml was removed from each culture and transferred to a clean 15 ml tube to serve as the un-induced control, and another 1 ml was transferred to a clean 15 ml tube and 0.5 mM isopropyl *β*-D-thiogalactoside (IPTG) added to serve as the induced culture. Both un-induced and induced cultures were incubated for a further 2 hours at 37 °C with vigorous shaking after which the cells from both cultures were harvested by centrifugation at 16 100 x g for 10 minutes. The cell pellets were re-suspended in 50 µl 2X sample buffer and heated for 10 minutes at 90 °C. Thereafter, the samples were spun down for 1 minute at the highest speed and 20 µl of the supernatant from each sample was quickly loaded onto a 15 % SDSPAGE gel for analysis and these are shown in Figures A1 and A2 below. The remaining 3 ml of the best expressing culture was used to make glycerol stocks which were later used for large-scale recombinant protein expression.

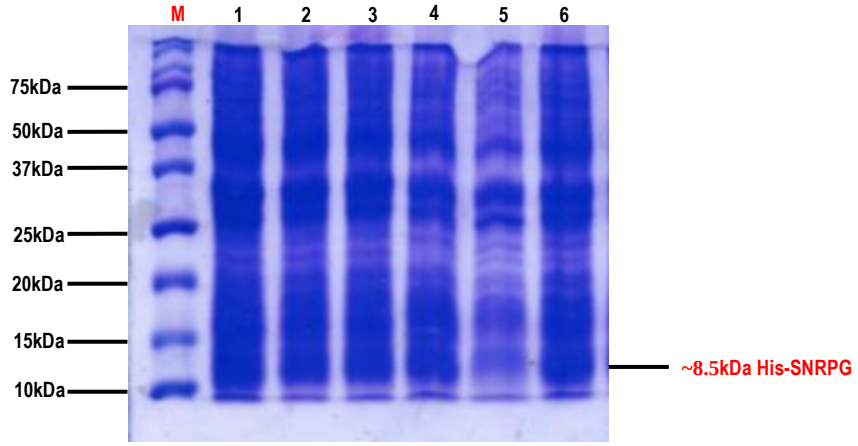


Fig. A1: Small scale expression profile of colonies transformed with pQE30~SNRPG. Lane 1 shows protein ladder, even-numbered lanes depict total bacterial cell lysates of cultures without induction by IPTG, while odd-numbered lanes indicate induction of protein expression with 0.5mM IPTG. The screening showed leaky expression of the SNRPG His-tagged protein running higher than its expected molecular weight of 8.5kDa which may have been due to incomplete denaturation of the protein. The addition of β-mecaptoethanol or DTT and boiling the samples at a higher heat for longer may have solved this problem.

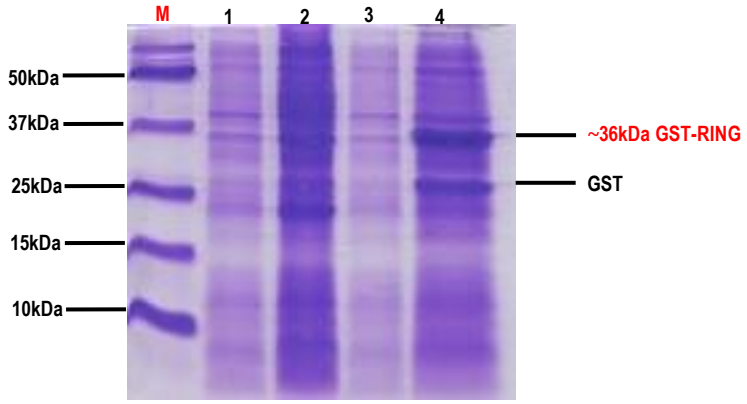


Fig. A2: Small scale expression profile of colonies transformed with pGEX-6P-2~RING. Lane 1 shows protein ladder, lanes 1 and 3 indicate un-induced cultures while lanes 2 and 4 show cultures induced with 0.5mM IPTG. Better expression of the protein was observed in lane 4, which showed expression of the RING GST-tagged protein at approximately 36kDa which corresponds in size to the predicted molecular weight of the protein.

Recombinant production of SNRPG and the RING Finger domain of RBBP6

A 100 µl sample of the best expressing clone from the expression screen was used to inoculate 100 ml of Luria broth containing 100 µg/ml ampicillin. This culture was grown overnight on a rotary shaker at 37 °C, shaking at 210 rpm. The following day, the culture was used to set up a 1:1000 dilution i.e. 1 ml of the overnight culture was used to inoculate 1 L of Luria broth containing 100 µg/ml of ampicillin. This was left incubating at 37 °C until the OD_{600}

was between 0.4–0.6. The cultures were then induced with 0.5 mM IPTG concentration and the proteins were expressed at 25 °C overnight. For the RING finger domain protein 100 µM $ZnSO_4$ was added prior to IPTG induction to ensure correct folding of the protein. Several protein expressions followed of up to 20 L per cycle. Following expression, the cells were harvested the next day by centrifuging the cells for 20 mins at 20 000 xg. Resulting pellets were resuspended in their appropriate lysis buffers which contained 50 mM Tris-HCL, 100 mM NaCl and 10 mM Imidazole for the SNRPG protein and PBS at pH7 containing 100 µg/ml Lysozyme, 1 mM PMSF, 1 mM DTT, 1 % Triton X-100 and 100 µM $ZnSO_4$ for the RING finger domain protein respectively. The re-suspended cell pellets were stored in -80 °C until required for purification. Results exhibiting expression of the SNRPG protein is shown below (Fig. A3). Expression of the RING finger domain protein is shown along with the purification profile of the protein (Fig A5).

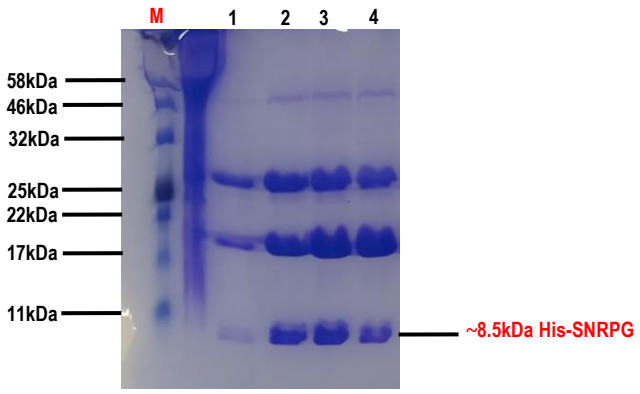


Fig. A3: Recombinant expression of SNRPG protein. Lanes represent progressive expression of the protein from 5µL to 20µL in a 2L expression. The last two lanes represent expression of the protein from a 1L expression.

Extraction

Frozen cell pellets were thawed on ice and treated with 1X EDTA-free protease inhibitor cocktail tablets. The cells were lysed by three to five cycles of freeze-thawing by incubating the cell suspensions at -80 °C for 10 minutes, followed by incubation at 37 °C for 10 minutes. The total

bacterial lysate containing cell debris and other particulate matter was then cleared by centrifugation at 20 000 xg for 30 minutes at 4 °C. The clear supernatant containing soluble protein was then decanted into fresh 50 ml tubes. Sodium azide was added to a final concentration of 0.02 % to prevent bacterial growth before storage at 4 °C.

Affinity Purification: SNRPG protein

A Nickel-NTA column using 10 ml of HisSelect Nickel-linked agarose beads was packed into a column by pouring the beads into a plastic column and allowing the storage ethanol to flow out. The beads were washed several times with dH_2O and then allowed to pack well by leaving the column in 4° C overnight. The next day, the beads were equilibrated with 5CV of equilibration buffer containing 50 mM Tris (pH 8) and 40 mM Imidazole. The total bacterial cell lysate was then poured down the column and the flow through collected and kept on ice. The column was washed again with 5CV equilibration buffer and the protein was eluted using elution buffer containing 50 mM Tris (pH 8) and 300 mM Imidazole. Samples were taken at every step for SDS-PAGE analysis. After purification, the beads were stored in 30% ethanol and left at 4 until further required while 0.02 % of 10 % NaNz was added to elution fractions.

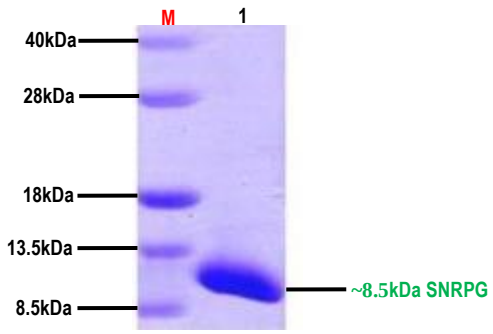


Fig. A4: Affinity Purification of SNRPG protein. Lanes represent protein ladder and the purified SNRPG protein running at the expected size of 8.5kDa.

Affinity Purification: RING finger domain protein

A 5 ml glutathione-agarose (SIGMA® Aldrich) column was prepared by weighing out the required amount of the glutathione-linked agarose beads, suspending it in deionised water and allowing it to swell overnight at 4 °C. The following day, the slurry misxture formed was poured into a 15x1 cm Econo® Chromatography Column (Amersham Pharmacia) of 20 mm diameter and allowed to pack well by standing at 4 °C overnight. Thereafter, the column was washed with dH_2O and equilibrated with 3 column volumes (3 CV) of PBS pH 7.4. The bacterial lysate was then poured unto the pre-equilibrated column and allowed to pass through by gravity flow. The resultant flow-through was collected and kept on ice. The column was then washed with 5 column volumes (5 CV) of PBS pH 7.4, after which the bound GST-RING Finger domain protein was eluted with 3 column volumes (3 CV) of elution buffer (50 mM Tris-HCl, pH 8.0 containing 15 mM reduced glutathione). The eluted GST-RING domain fusion protein was stored at 4 °C after the addition of 1 mM PMSF, 1 mM Dithiothreitol (DTT) and 0.02 % sodium azide. After protein elution, the glutathione-agarose column was washed with 3 column volumes (3 CV) of 2 M NaCl to remove any non-specific bound proteins on the beads and 5 column volumes (5 CV) of deionized water to clean up the affinity column further before storage at 4 °C. The expressed SNRPG protein was purified using a Nickel-NTA column recharged with cobalt. The eluted protein was then subjected to SDS-PAGE electrophoresis for analysis.

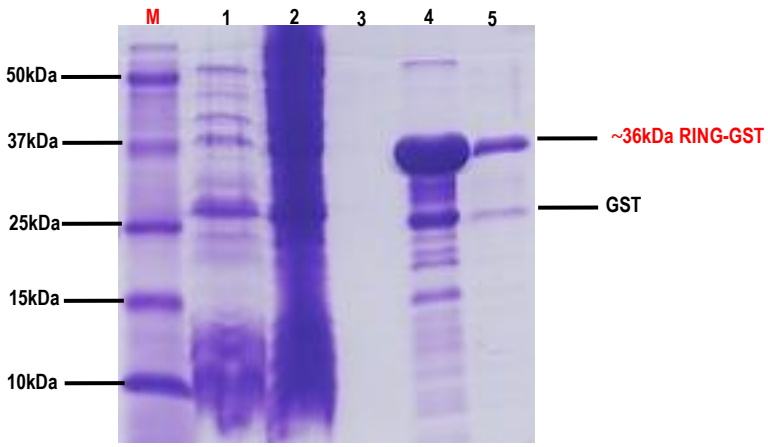


Fig. A5: Affinity Purification of RING finger domain of RBBP6. Lane M- protein marker; Lane 1 – Flowthrough; Lane 2 – Lysate (protein expression); Lane 3 - Clean Wash; Lane 4 - Elute 1 (in 20ml); Lane 5 – Elute 2 (in 10ml)

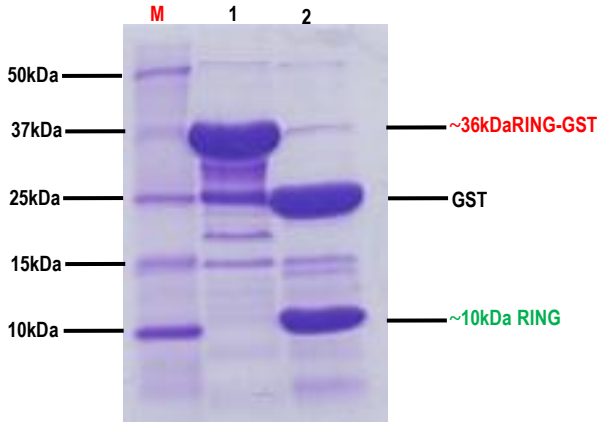


Fig. A6: Cleavage of RING finger domain of RBBP6. Lane M- protein marker; Lane 1 – Pooled elute before cleavage with 3C-Protease; Lane 2 – Pooled elute after cleavage completion

Protein concentration determination

The protein concentration was determined using a NanoDrop® ND2000 spectrophotometer (Thermo Fisher Scientific) by measuring the absorbance at 280 nm. About 0.5-2 µl sample of the elution buffer (respective to each protein) was placed directly on top of the detection surface of the NanoDrop® ND2000 spectrophotometer (Thermo Scientific) to blank it. Thereafter, 0.5-2 µl of the purified protein samples was placed directly on top of the detection surface in order to determine the protein concentration of His-SNRPG and the GST-RING finger domain protein.