Understanding the mechanism of nucleosome remodeling by SMARCAL1

Thesis submitted to
Jawaharlal Nehru University
for the award of the degree of

DOCTOR OF PHILOSOPHY

Ritu Bansal

SCHOOL OF LIFE SCIENCES
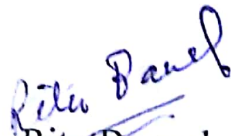**JAWAHARLAL NEHRU UNIVERSITY**
NEW DELHI-110067
INDIA
2017

JNU

CERTIFICATE

· This is to certify that the work embodied in this thesis entitled **"Understanding the mechanism of nucleosome remodeling by SMARCAL1"** has been carried out in the School of life Sciences, Jawaharlal Nehru University, New Delhi, India. This work is original and has not been submitted so far, in part or in full for the award of any other degree or diploma in any other university.

Dr. Rohini Muthuswami
Rohini Muthuswami
Associate Professor
(Supervisor)
Jawaharlal Nehru University
New Delhi - 110067

Ritu Bansal

(Candidate)

12.3.1?

Dr. S.K. Goswami
(Dean)

Prof Shyamal K. Goswami
Dean
School of Life Sciences
Jawaharlal Nehru University
New Delhi - 110067

ii

Dedicated to my parents
&
Friends

ACKNOWLEDGEMENT

At the verge of completing my thesis, I can truly say that I am thankful to all my teachers, friends and family for their unconditional love and support.

First and foremost, I would like to express my sincere regards towards my supervisor Dr. Rohini Muthuswami for her continuous support and encouragement during my Ph.D. I would always be thankful for her continuous motivation whenever my experiments failed and for her ideas to make an experiment work. It is a difficult thing to get a teacher who supports you in your tough time and guide you towards the right path. I would like to thank her for being that one teacher for me all during my five years of Ph.D. Her guidance helped me a lot in improving my writing skills and hence writing my thesis. Without her precious time and support, it would have been difficult for me to complete my research. I appreciate her efforts in molding me to go beyond my limits and aim for best. Here, I would also like to thank current Dean Dr. S.K. Goswami and previous Dean Dr. B.N. Malick.

I would like to thank Dr. Ajay Kumar Saxena and lab's seniors Dr. Sitaram Meena and Dr. Shantipal Gangwar for their support and encouragement that helped me a lot in my Ph.D. work.

I would also like to express my gratitude towards Dr. S. S. Komath, Dr. S.Gourinath, Dr. P.K.Yadav, Dr. Apoorva Sau, Dr. R.N.K. Bamezai, Dr. Alok Mondal, and Dr. Neelima Mondal for giving me the research ideas during my presentations and motivating me towards better science.

I am out of words to express my gratitude towards all the lab members for their affection, understanding and friendship during all my five years. I would like to thank my lab seniors Dr. Dominic and Dr. Pynskhem for providing a healthy environment and good guidance during my Ph.D. I would like to thank my immediate seniors Dr. Tapan, Dr. Meghna and Dr. Isha for providing a nice discussion over lab meetings and motivating me towards science. I would also like to thank my lab mates Dr. Soni, Ketki , Ramesh, Vijendra, Dr. Upasana and Pramita; for all the fun we had during last five years. It's not me who did it, it's actually we did it! I would also like to thank my juniors Rakesh, Prashant and Anindita for their love and support. I am also thankful to Shashi Ji, Bishan Ji and Sheenu Ji for their assistance.

me and giving a sweet gesture all the time. You all are the reason why I keep pushing myself to do best. I love you all!

Finally, I would also like to thank a good friend and senior Dr. Atul Kumar Agrawal for his continuous love, support and motivation. Thanks for being a good friend!

I am also thankful to all current and former JNU students who worked hard to make this campus beautiful and memorable one. All the activities like debates, hostel nights, and cultural programs have helped a lot in developing me as a person. Thank you JNU!

<div align="right">Ritu Bansal</div>

ABSTRACT

The function of SWI/SNF helicase proteins have widely been characterized in DNA-dependent molecular processes like DNA replication, transcription and DNA damage repair. The protein of my interest, SMARCAL1, is a member of SWI/SNF helicase family that plays role in DNA replication and repair by utilizing the energy released from ATP hydrolysis in the presence of DNA molecules containing double-strand to single-strand transition regions to mediate replication fork stabilization. Recently, our laboratory has shown that SMARCAL1 can also function as a transcription modulater.

ADAAD is the N-terminal proteolytic fragment of SMARCAL1 that possesses all the characteristic helicase motifs (Q, I, Ia, II, III, IV, V and VI) of SMARCAL1 necessary for its DNA dependent-ATPase function. In my thesis, I have characterized the role of the positively and negatively charged residues (D591, R592, H594 and R595) present in motif VI of ADAAD using site-directed mutagenesis. Using circular dichroism I have shown that motif VI plays a role in maintaining the conformational integrity of ADAAD. Ligand binding experiments showed that conformational alteration in the motif VI mutants leads to defect in the interaction of the mutant proteins with stem-loop DNA in the presence of ATP. Fluorescence spectroscopy also revealed that the positive charge of R595 residue was crucial for binding to stem-loop DNA in the presence of ATP indicating that this residue makes direct or indirect contact with the stem-loop DNA. My results, thus, indicate that the motif VI of ADAAD seems to behave in a similar manner as eIF4A, an RNA helicase, towards nucleic acid binding. Further, this is the first report

that signifies the role of motif VI of a member of SWI/SNF family, ADAAD, in maintaining the conformational integrity of the protein.

The crystal structure of eIF4A has shown the importance of inter-domain interaction between the two RecA-like domains constituting its helicase core through salt bridge stabilization between motif VI and motif II. By constructing double mutant protein H329D/D591H and studying its ATPase activity, I observed that unlike eIF4A, salt bridge is not formed between motif II and motif VI of ADAAD. However, H329D/D591H showed a synergistic defect in stem-loop DNA binding in the presence of ATP. I also constructed a double mutant protein W277A/R595H that again showed a synergistic defect in stem-loop DNA binding in the presence of ATP. From these experimental results, I concluded that both the RecA-like subdomains 1A and 2A work synergistically to interact with stem-loop DNA in presence of ATP.

To verify that whether this synergistic defect in stem-loop DNA binding was extended to intra-domain motifs of ADAAD, I studied the ligand binding and conformational alteration of RecA-like domain 2A associated double mutant R592A/F507A where the mutations were made in motif IV and motif VI respectively. Ligand binding experiments showed that the R592 residue of motif VI was not needed for stem-loop DNA binding in the presence of ATP; however, F507 residue of motif IV was found to be important for stem-loop DNA binding in the presence of ATP. The double mutant F507A/R592A was also found to show altered interaction with stem-loop DNA in the presence of ATP. Further, the secondary structure of double mutant showed alteration as compared to the single mutants R592A and F507A. These experiments, thus, led me to infer that the RecA-like 2A domain is crucial for maintaining the conformational integrity of the protein.

Earlier it has been observed in our lab that SMARCAL1 negatively regulates the transcription of c-*myc*; however, the mechanism for this negative regulation was still a puzzle. As SMARCAL1 is reported to bind the double-strand to single-stand transition regions of DNA, therefore, I analysed the secondary structures present in the c-*myc* promoter using QGRS-mapper and Mfold analysis. Circular dichroism spectroscopy showed that ADAAD induces a conformation alteration in these secondary structures of c-*myc* promoter DNA and this alteration is similar to the conformational changes induced in stem-loop DNA by the protein. The alterations in the DNA were dependent on the ATPase activity of the protein as K241A mutant that is unable to hydrolyze ATP was also unable to induce conformational alteration in the c-*myc* promoter.

Inherited, biallelic mutations in SMARCAL1 causes Schimke Immuno-Osseous Dysplasia (SIOD), which is characterized by renal failure, T-cell immunodeficiency and cancer predisposition. Therefore, I analyzed the mutation R820H (corresponds to R595H in ADAAD) present in SIOD patients for its biochemical properties, cellular localization and transcriptional regulation of c-*myc* and *dgcr8* gene. The ATPase activity assay showed that the mutant protein was unable to catalyze the ATP hydrolysis. CD spectroscopy indicated that conformation of the mutant protein had significantly been altered and this was further verified by molecular simulation experiments. The RMSD value showed a signified alteration in dynamic stability of R595H as compared to the wild-type ADAAD. Ligand binding experiment revealed that the interaction of R595H with stem-loop DNA in the presence of ATP was impaired. Conformational analysis indicated that R595H mutation was not able to induce the conformation alteration in both c-*myc* as well as *dgcr8* promoter DNA. Taken together, I conclude that mutation of R820 to H820 results in impaired conformational integrity of the protein leading to impaired stem-loop DNA binding in the presence of ATP and loss of ATPase activity. The non-

functional mutant protein is thus, unable to regulate c-*myc* and *dgcr8* transcription and thereby, possibly resulting in the development of SIOD disease.

Abbreviations

ACF	ATP utilizing chromatin assembly and remodeling factor
ADAAD	Active DNA-dependent ATPase A Domain
ATM	Ataxia telangiectasia mutated
ATP	Adenosine triphosphate
ATR	Ataxia telangiectasia and Rad3 related
BAF	BRG1 or BRM-associated factors
BRG1	Brahma related gene
BSA	Bovine serum albumin
CD	Circular Dichroism
Cdk	Cyclin-dependent kinases
CENP-A	Centromeric protein A
CHD	Chromodomain helicase DNA binding
DEAE	Diethylaminoethyl
DEPC	Diethylpyrocarbonate
DMEM	Dulbecco's modified eagle's medium
DNA	Deoxyribonucleic acid
DNA-PK	DNA-dependent Kinases
dNTP	Deoxy nucleotide triphosphate
dsDNA	Double-stranded DNA

DTT	Dithiotheritol
E. coli	*Escherichia coli*
EDTA	Ethylenediaminetetraacetic acid
EGTA	Ethylene glycol-bis (β-aminoethylether)-N,N,N',N tetraacetic acid
eIF4A	eukaryotic initiation factor 4A
ERCC6	Excision Repair 6, Chromatin Remodeling Factor
EtBr	Ethidium bromide
FBS	Fetal bovine serum
EP400	E1A binding protein 400
FITC	Fluorescein isothiocyanate
FP	Forward primer
FUN30	Function Unknown Now 30
GAL promoter	Galactose promoter
GFP	Green Fluorescence Protein
GST	Glutathione Sepharose Tag
HARP	HepA-related protein
HAT	Histone acetyl transferases
HCV	Hepatitis C Virus
HDAC	Histone deacetylase complex
HeLa	Henrietta Lacks
HMGN	High mobility group nucleosomal
HMTs	Histone methyl transferases
HP1α	Heterochromatin protein 1α

HR	Homologous recombination
HRP	Horse radish peroxidase
IB	Immuno blot
IgG	Immunoglobulin G
IP	Immunoprecipitation
IPTG	Isopropyl-β-thiogalactopyranoside
IR	Ionizing radiation
ISWI	Imitation switch
kb	Kilobases
kDa	Kilo Dalton
K_{cat}	Catalytic constant/Turnover number
K_d	Dissociation constant
K_M	Michaelis-Menten constant
LANA	Latency-associated nuclear antigen
LB	Luria-Bertini
MeCP2	Methyl-CpG binding protein
MNase	Micrococcal nuclease
NADH	Nicotinamide adenine dinucleotide
NER	Nucleotide excision repair
NLS	Nuclear localization signal
NS3	Nonstructural Protein 3
NURF	Nucleosome remodeling factor
NURD	Nucleosome remodeling deacetylase

OD	Optical density
PAGE	Polyacrymylamide gel electrophoresis
PBAP	Polybromo containing BAP
PBS	Phosphate buffer saline
PBST	Phosphate buffer saline tween-20
PCR	Polymerase chain reaction
PMSF	Phenylmethane sulfonyl fluoride
PVDF	Polyvinylidene difluoride
RapA	RNA polymerase associated protein
Rb	Retinoblastoma
RecA	Recombination protein A
RIPA	Radioimmunoprecipitation assay buffer
RMSD	Root mean square deviation
RNA	Ribonucleic acid
RP	Reverse primer
RPA	Replication protein A
RSC	Chromatin structure remodeling
RT-PCR	Real-time PCR
slDNA	Stem-loop DNA
ssDNA	Single-stranded DNA
SDS	Sodium dodecyl sulphate
SF2	Superfamily 2
SIOD	Schimke immunoosseus dysplasia

SMARCAL1	SWI/SNF related matrix associated actin dependent regulator of chromatin subfamily A-like protein 1
SMARCAD1	SWI/SNF related matrix associated actin dependent regulator of chromatin subfamily A, containing DEAD/H Box1
SNF	Sucrose non fermenting
SHPRH	SNF2, Histone-linker, PHD and RING Finger Domain-Containing Helicase
SUMO	Small ubiquitin like modifier
SWI	Mating type switch
SWI/SNF	mating type switching/Sucrose Non-Fermenting
SWR1	SWI2/SNF2 related ATPase
TBP	TATA binding protein
TF	Transcription factor
TRITC	Tetramethylrhodamine isothiocyanate
TTF2	Transcription Termination Factor 2
UV	Ultraviolet
ZRANB3	Zinc Finger Ran-Binding Domain-Containing Protein
β-Me	β- mercaptoethanol

LIST OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1: Review of literature

INTRODUCTION

The human genome is made up of roughly 3 billion base pairs of DNA packaged into 23 chromosomes in a haploid cell. The basic diploid nature of the human cells, except in sperm and ova, doubles the number of chromosomes and thus, the total amount of DNA is 6 billion base pairs in a cell. Since each base pair is 0.34 nm long, therefore, each diploid cell is made up of approximately 2 meters of DNA (Getzenberg et al., 1991). Furthermore, a human body is made up of 50 trillion of cells making the total length of DNA in a human body to around 100 trillion meters. After reading all these facts, the first question that arises in our mind is how this lengthy DNA is squeezed into a human body. So, here in my review of literature I will briefly elaborate the mechanism adopted to allow the packaging of DNA into a mitotic cell. I will also describe how chromatin is altered by different proteins to access this packaged DNA for catalyzing basic metabolic processes inside cell.

Histone-DNA octamer: nucleosome

The solution to above query lies in the fact that the lengthy DNA is compacted with the help of histone and non-histone proteins. An octamer of positively charged histone proteins (H3, H4, H2A, H2B) interacts with the negatively charged phosphodiester DNA backbone and the resulting DNA-protein complex is called as the nucleosome (Bakayev et al., 1975; Kornberg, 1974). This histone-DNA unit is also defined as the main unit for transcriptional control for gene activation and repression at the level of gene promoters, gene transcriptional unit and chromosomal bodies. The X-ray structure of the nucleosome has also been solved to gain an

insight into the histone-DNA interaction and its significance (Luger et al., 1997; Richmond and Davey, 2003). A regularly spaced nucleosomal array with linker histones or other nucleosomal proteins combine to make a 30 nm chromatin fiber (van Holde and Zlatanova, 1995; Woodcock and Dimitrov, 2001).

The histones have been found to weigh approximately the same mass as DNA in a nucleosome unit. Besides, they were viewed to be the diverse proteins necessary for regulation of gene expression (Woodcock and Dimitrov, 2001). Further, major experiments like acid-based extraction proved that these proteins were conserved across species i.e. H4 histone (DeLange et al., 1969; Phillips and Johns, 1965).

As stated earlier, the histone octamer is made up of H2A, H2B, H3 and H4 wherein the nucleosome contains H2A-H2B dimer and $(H3)_2$, $(H4)_2$ tetramer (Hereford et al., 1979; Thomas and Furber, 1976; Thomas and Kornberg, 1975). In addition, a fifth histone, H1, functions as a linker protein between two nucleosomes (Thomas and Kornberg, 1975). The repeating structure in the form of nucleosome beads on chromatin were identified with the help of electron microscopy and biochemical studies (Oudet et al., 1975). Further, micrococcal nuclease digestions studies led to the identification of core histone octamer (Finch et al., 1977). Finally, the X-ray crystallographic structure has also helped in our understanding of the nucleosome (Luger et al., 1997).

The 30 nm Chromatin fiber

The next level of compaction involves interaction between nucleosomes that drive the folding of the nucleosome array into 30 nm chromatin fibers (Woodcock and Dimitrov, 2001). Although the crystal structure of nucleosome core is available at a resolution of 1.9 A°, there is a strong need to understand the structure of chromatin fiber to study how different modification of chromatin fibers alters chromatin structure and lead to differential gene expression. Based on the biophysical and biochemical studies two different models have been proposed for chromatin fiber structural arrangement (Tremethick, 2007). The first model proposes a solenoid structure, where the nucleosomes are arranged one after another to fold into a one-start helix (Figure 1.1) (Robinson et al., 2006). The second model is referred as zigzag model, where two rows of nucleosomes lie beside each other in a zigzag fashion and linker DNA crisscrosses between each stack of nucleosome (Figure 1.1) (Bednar et al., 1998).
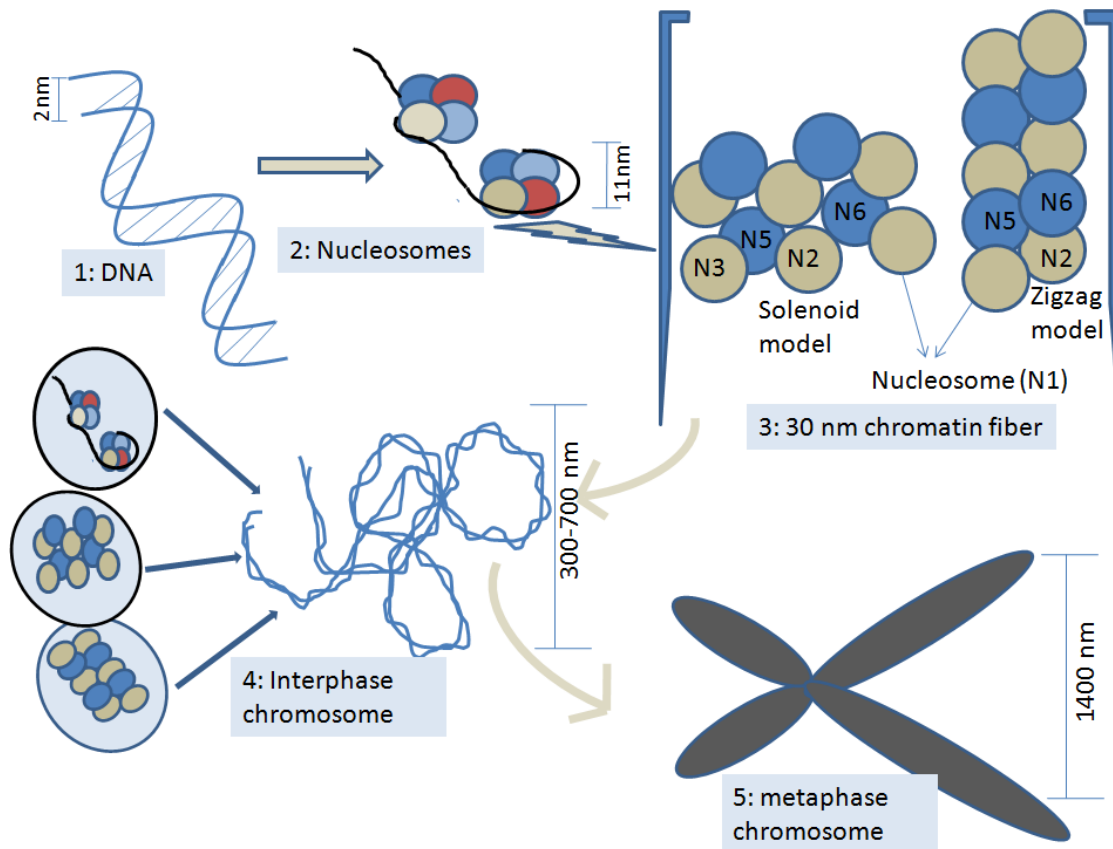
Figure 1.1: Model depicting multiple stages of chromatin folding: DNA, the 30 nm chromatin fiber and mitotic chromosome. The 30 nm chromatin fiber has shown both the arrangements of nucleosomes in chromatin.

Mechanism behind chromatin fiber formation

Due to the compact form of chromatin fiber, the crystal structure of the 30 nm fiber is not available yet. The crystal structure of nucleosomal core suggests an uneven charge distribution on the surface of nucleosome. A cluster of acidic amino acid residues of H2A interacts with the

N-terminal tail of histone H4 initiating from a neighboring nucleosome (Davey et al., 2002; Luger et al., 1997). Mutational studies performed with H4 histone encompassing 14-16 amino acid residues proved that this interaction was crucial not only for the formation of nucleosome crystal, but also for the compaction and generation of high-order chromatin fiber (Dorigo et al., 2004). Activation of transcription requires acetylation of histone H4 to inhibit the formation of 30 nm chromatin fiber by altering its interaction with acidic patch of H2A histones of neighboring nucleosomes (Shogren-Knaak et al., 2006; Shogren-Knaak and Peterson, 2006). Similar to H4 histone, many H2A variants have also been discovered that inhibit or promote the formation of compact chromatin domains leading to activation or inhibition of transcription respectively (Greaves et al., 2007; Valdés-Mora et al., 2012). For example, incorporation of histone variant H2A.Z that differs from H2A in just two amino acids promotes the formation of 30 nm chromatin fiber (Fan et al., 2004). In contrast, the histone variant H2A.Bbd has been shown to inhibit the formation of condensed chromatin fiber and thereby, facilitate the activation of transcription (Bao et al., 2004; Gautier et al., 2004). Other chromatin architectural proteins like HP1α, methyl-CpG binding protein (MeCP2) and the polycomb complex helps in constraining the primary chromatin structure into the secondary and tertiary chromatin domains (McBryant et al., 2006). Additionally, some viral proteins, for example latency-associated nuclear antigen (LANA) of Kaposi-sarcoma associated herpes virus, have been shown to mimic the structure and charge of H4 tail and hence, bind to the acidic patch of histone H2A resulting in chromatin compaction (Barbera et al., 2006).

The compact or open structure of chromatin seen in isolated chromatin fragments shows their association either with heterochromatin region or transcribing gene regions, respectively (Gilbert et al., 2004). *In vitro*, magnesium ion concentration has also been observed to play a role in the

formation of tertiary chromatin structures by inducing a self-association between two linker

histones in a nucleosomal array (Strick et al., 2001). *In vivo*, the presence of 3 mM magnesium

ion concentration in the nucleus during interphase and 17 mM during metaphase supports the

role of magnesium ion in condensation of chromosome (Strick et al., 2001). Although the

structure of 30 nm chromatin fiber has remained highly mysterious, still the ability to assemble it

in vitro or isolating it from nuclei has helped us in elucidating this secondary structure and

hence, helped us in understanding the higher-order chromatin compaction to a next level.

Condensation of chromatin fiber into mitotic chromosome

The 30 nm chromatin fiber needs to be further compacted at least 200-fold to finally fit inside the

nucleus during mitosis (Figure 1.1) (Khorasanizadeh, 2004). A zigzag model has been proposed

wherein it has been suggested that the chromosome compaction is a dynamic and energy-

dependent process catalyzed by ATP driven enzymes like topoisomerase II and condensins

(Rydberg et al., 1998). The process introduces a positive supercoiling/writhe in the closed

circular DNA by these condensin complex leading to chromosome condensation (Rydberg et al.,

1998). Many histone modifications in histone H3 and H1 are also associated with generation of

condensed form of mitotic or metaphasic chromosome. For instance, the replacement of histone

H3 by H3 variant CENP-A in centromeric nucleosomes has been shown to help in segregation of

chromosomes during mitosis (Sullivan and Karpen, 2004). Further, phosphorylation of histone

H3 and H1 tail have also been monitored. Though the role of H3 phosphorylation is not clearly

defined, the importance of H1 phosphorylation has been observed for the super packing of

mitotic chromosome (Dou et al., 2002).

Chromatin regulation through its modification

This tightly packaged DNA can be accessed for various DNA processes only by repositioning the nucleosomes. DNA accessibility is made possible by two different classes of enzymes-histone modifying enzymes and ATP-dependent chromatin remodelers.

I will begin by discussing about the histone modifying enzymes that covalently modify the histones and thus, alter the interaction between the nucleosome and the DNA (Figure 1.2) (Brownell et al., 1996; Taunton et al., 1996). A diverse range of modifications have been recognized that modify precise sites on the histones (Strahl and Allis, 2000). These modifications introduce a 'histone code' that ultimately leads to downstream events like activation or repression of transcription (Strahl and Allis, 2000). The modifications are recognized by chromatin-associated proteins possessing domains that can recognize and bind to the modified histone regions. For example, the bromodomain and chromodomain recognize acetylated or methylated histones respectively and thus, helps in gene regulation (Fischle et al., 2003b; Santos-Rosa and Caldas, 2005; Winston and Allis, 1999). In the following section, I will define the role of chromatin modifying enzymes and recognition of these histone proteins in transcription regulation.

Lysine Acetylation and deacetylation

Acetylation of histones leads to relaxation of histone-DNA contacts in a nucleosome, thus, promoting accessibility of the transcription activators to the DNA molecule (Kurdistani and Grunstein, 2003; Kurdistani et al., 2004). To assist in the process of histone acetylation many

chromatin modifiers containing histone acetyl transferase (HAT) domain have been discovered

(Figure 1.2) (Roth et al., 2001). These HAT domains reside in many multi-subunit protein

complexes. For instance, Gcn5 HAT is present in two complexes Ada and SAGA, both of which

perform different biological function (Grant et al., 1997). Structurally, Gcn5 HAT contains a

central conserved core that binds to the co-factor acetyl coenzyme A first and then to the

substrate H3 and H2B through a cleft present over the co-factor binding site (Marmorstein and

Roth, 2001). After association with the Gcn5 HAT, the histone H3 adopts a random coil structure

that results in the interaction between the histone molecule and Gcn5 leading to acetylation of

the lysine present in H3 (Marmorstein and Roth, 2001). This acetylation is promoted by the

phosphorylation of H3 serine. Together these two modifications promote transcription activation

(Clements et al., 2003). The effect of lysine acetylation is reversed by the action of many histone

deacetylases (HDACs), like Sir2, that binds to acetyl lysine of H4 peptide with the help of

conserved histidine present in active site of HDACs (Figure 1.2) (Marmorstein, 2001).

Lysine methylation

Generally, enzymes that catalyze the methylation of lysine contain a SET domain that first binds

to the co-factor S-adenosyl-L-methionine and then transfers this methyl group to the lysine of

histone H3 (Figure 1.2) (Dillon et al., 2005; Zhang et al., 2003). This covalent modification is

associated with epigenetic inheritance where some activators and repressors maintain it in the

mitotic and meiotic transcription pattern and pass this genetic state to daughter cells (Lachner et

al., 2003; Tachibana et al., 2007). Additionally, an example of epigenetic activation has also

been seen in *Drosophila* where the Ash1 protein methylates H3 lysine thus, creating a signal for

the recruitment of transcriptional activator Brahma complex (Beisel et al., 2002). The Brahma

complex inhibits HP1, a protein that has role in gene silencing, thereby resulting in

transcriptional activation (Beisel et al., 2002; Turner, 2002). Proteins like HP1 and polycomb,

through their chromodomain, can also bind to methyl lysine of histone H3 and thus, mediate

gene silencing and chromatin condensation (Jacobs and Khorasanizadeh, 2002; Lachner et al.,

2001). Thus, histone methylation can mediate both transcription activation and repression.

Lysine ubiquitination and SUMOylation

Attachment of ubiquitin molecule to the C-terminus of histone leads to either transcription factor

degradation through proteasomal degradation pathway or to recruitment of other modification

complexes (Pickart, 2004; Sun and Allis, 2002; Zhang, 2003). Ubiquitin attachment is conducted

in three steps involving E1 activation, E2 conjugation and finally E3 ligase enzymes (Figure 1.2)

(Robinson and Ardley, 2004). H2B ubiquitination and then deubiquitination in a sequential

manner leads to methylation of lysine in histone H3 and thus, these events lead to transcription

regulation (Figure 1.2) (Zhang, 2003) .

While ubiquitination of histone H2A and H2B is seen to be important for transcription

activation, sumoylation of histone H4 leads to transcriptional repression (Shiio and Eisenman,

2003). Sumoylation is a form of post-translational modification that involves the attachment of

small ubiquitin like modifier (SUMO) proteins to the lysine group of histone proteins and so

regulates many cellular functions like transcription, DNA repair and pre-amyloid oligomer

formation (Geiss-Friedlander and Melchior, 2007; Wilkinson and Henley, 2010).

Serine phosphorylation

The association of histone H3S28 phosphorylation with chromosome condensation during mitosis is well established (Hsu et al., 2000). Other phosphorylation sites on histone H3, H4, H2A and H2B have also been identified (Figure 1.2). Of these, Ser10 phosphorylation on H3 has been linked to transcription activation (Cheung et al., 2000; Nowak and Corces, 2000). The enzymes/kinases that can phosphorylate histone H3 include PKA, aurora kinase-B/lpl1, Msk1 and Rsk-2 all of which target the serine/threonine sites on histones that are surrounded by basic amino acid residue like lysine (Rossetto et al., 2012; Sawicka and Seiser, 2012). Phosphorylation, thus, achieved can lead to acetylation of neighboring lysine and, hence, inhibit methylation of lysine as observed in the case of Gcn5, which shows increased lysine acetylation because of its interaction with the closest phosphoserine (Fischle et al., 2003a; Nowak and Corces, 2004). Phosphorylation of lysine can be reversed by the action of protein phosphatase 1(PP1) family, for example, Glc phosphatase that operates on serine 10 of histone H3 (Figure 1.2) (Hsu et al., 2000).

DNA methylation

Besides covalent modification of histones in nucleosomes, the DNA also undergoes modification that correlates with genome imprinting, embryonic development and X-chromosome inactivation (Rothbart and Strahl, 2014). The nucleotide base cytosine is methylated in mammals and plants, resulting in repression of gene transcription (Bird, 2002). This cytosine methylation mainly occurs at the CpG dinucleotides in mammals and has been linked to the formation of

heterochromatin (Figure 1.2) (Bird, 2002). The methylation of cytosine is achieved by DNA methyl transferase that is recruited to the chromatin by histone H3 lysine methyl transferase, thereby proving an association between the two silencing events (Tamaru and Selker, 2001). The DNA methylation is recognized by proteins containing methyl-CpG binding domain, like for example MeCP2, that gets recruited at the methylated positions and represses transcription by drafting Sin3 HDAC complexes at these sites (Fuks et al., 2003).
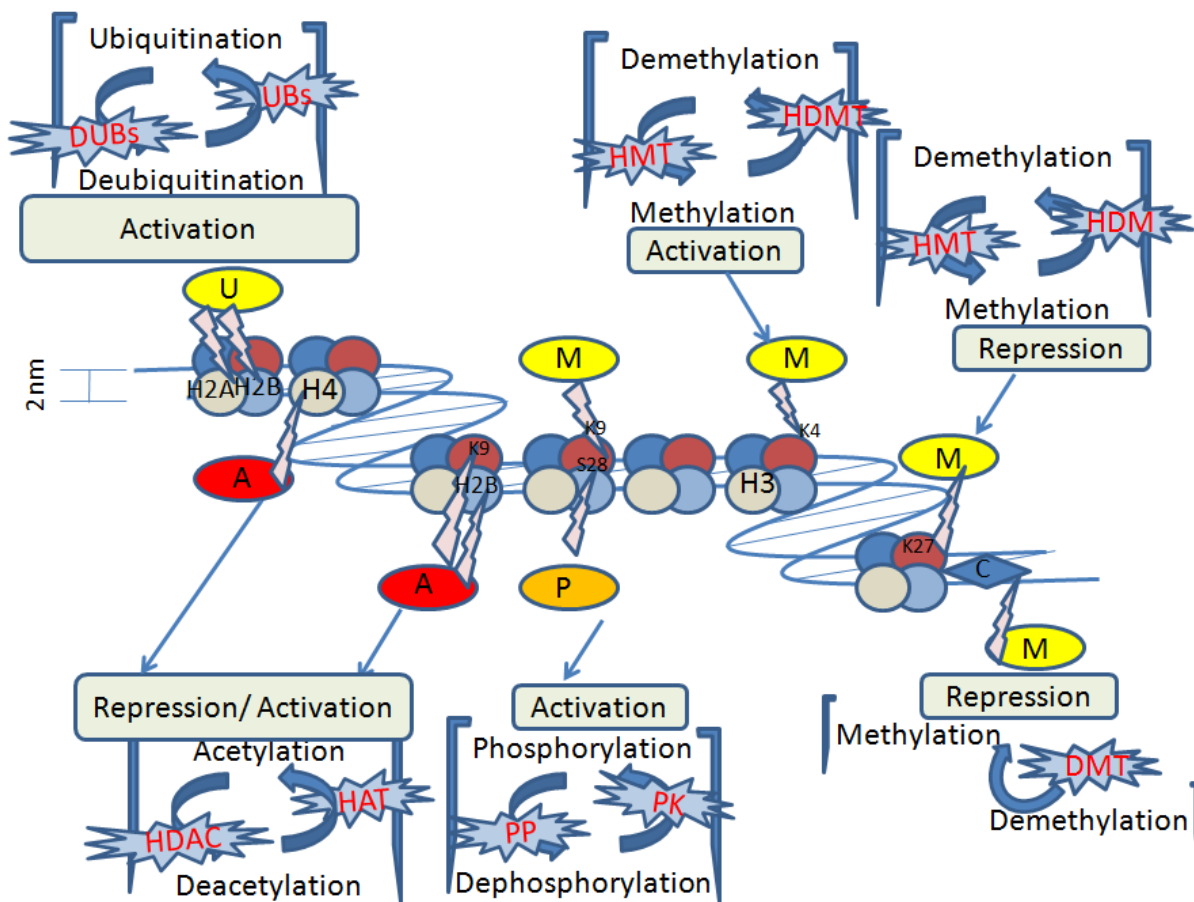
Figure 1.2: This figure depicts various post-translational modifications for gene expression regulation. Different histone modifying enzymes are shown with the sites of modifications (UBs- Ubiquitinating enzymes, DUBs- Deubiquitinating enzymes, HMT- Histone methyl

transferases, HDMT- Histone demethyltransferases, DMT- DNA methyl transferases, PK-Protein kinases, PP-Protein phosphatases, HAT- Histone acetyl transferases, HDAC-Histone deacetyl transferases).

ATP-dependent chromatin remodeling

The histone modifications and DNA methylation do not reposition nucleosomes; this is achieved by the ATP-dependent chromatin remodelers (Flaus et al., 2006; Narlikar et al., 2013). In general, both enzymes work in conjunction to either promote or inhibit the accessibility towards histones and DNA. For instance, many of these proteins work together with HDACs, methyl-CpG-binding proteins, histone acetylases and histone methyl transferase to regulate transcription (Luo and Dean, 1999). The importance of chromatin remodelers in cell growth and differentiation is proved by the fact that misregulation/mutation in any of these proteins leads to many kinds of cancer (Mirabella et al., 2016; Yan et al., 2008). Here, I will try to give a brief detail regarding ATP-dependent chromatin remodelers which utilize the energy of ATP to perform different functions like nucleosome sliding, nucleosome exchange and nucleosome eviction from their positions.

Discovery and early genetics

The first chromatin remodelers were identified by mutational studies in yeast during genetic screening that affected the transcription of genes *HO* and *Suc2* necessary for mating-type switching and growth on sucrose respectively and were referred as SWI/SNF proteins for

switching defective or sucrose nonfermenting respectively belonging to the Swi2/Snf2 family (Neigeborn and Carlson, 1984; Peterson and Herskowitz, 1992). SWI/SNF proteins function in a multi-subunit complex that work in association with many transactivators and histone acetylases like Gcn5 and therefore, helps in transcription activation (Hirschhorn et al., 1992). This is a 12 subunit complex that also contains the yeast proteins SWI3, SWI2/SNF2, SNF5 and SNF6 (Cairns et al., 1994; Tsukiyama and Wu, 1995). This multi-subunit complex disrupts nucleosome structure and increases accessibility towards DNA in an ATP-dependent manner (Cairns et al., 1994; Peterson and Herskowitz, 1992). Although these proteins are not necessary for cell viability, mutation in any of these proteins leads to alteration in transcription due to the disassembly of SWI/SNF protein complex (Holstege et al., 1998; Sudarsanam and Winston, 2000).

Another yeast complex termed RSC (Remodels the structure of chromatin) has also been purified in yeast that contains the DNA-dependent ATPase protein STH1 (Cairns et al., 1996). Unlike SWI/SNF protein complex, RSC protein complex is made of subunits that are necessary for mitotic growth (Angus-Hill et al., 2001; Cairns et al., 1996).

SWI/SNF proteins and their subfamilies

Around 25 years ago, proteins containing a short array of motifs required to hydrolyze ATP were grouped under the helicase family of proteins (Gorbalenya and Koonin, 1993). This helicase family has been further classified into three superfamilies-SF1, SF2 and SF3 (Gorbalenya and Koonin, 1993). The SF2 superfamily contains non-hexameric helicases, a subset of which

harness the energy of ATP to remodel the nucleosomes (Flaus and Owen-Hughes, 2004; Havas et al., 2000). Proteins that contain the same sequence characteristics and helicase motifs as Snf2p of *Saccharomyces cerevisiae* are classified as SNF2 family within the SF2 helicase superfamily (Flaus et al., 2006). These SNF2 families of proteins have been reported to help in ATP-dependent chromatin remodeling and DNA translocation (Flaus et al., 2006). Most of the Snf2 related proteins comes under the 'Snf2-like' grouping containing subfamilies Snf2p, Chd1, ISWI and Mi-2, which are the core subunits of ATP-dependent chromatin remodeling complexes (Flaus et al., 2006). The second category is 'Swr1-like' grouping that includes proteins like Swr1, EP400, Ino80 and Etl1. The third grouping is that of 'Rad54-like' proteins comprising of ATRX and Arip4 (Flaus et al., 2006). The 'SSO1653-like' grouping contains Mot1, SSO1653 and ERCC6 sub-family that are unique in their ability for recognizing non-chromatin substrates (Flaus et al., 2006). The final grouping includes SMARCAL1 and RapA subfamilies. These are classified as distant members because they lack the conserved Snf2 sequence hallmarks (Flaus et al., 2006).

The main property of chromatin remodelers is to mobilize/evict nucleosomes using their enzymatic activities (Hota and Bartholomew, 2011). Each remodeler can exist in more than one complex resulting in diversity. For example, *Saccharomyces cerevisae* contains around nine different chromatin remodeling complexes and humans contain at least 30 different chromatin remodeling complexes (Ho and Crabtree, 2010). This broad repertoire of complexes is recruited differentially at the genomic sites to perform different functions. The main reason underlying the complexity in the behavior of these chromatin remodelers is speculated to be the difference in their catalytic subunits and enzyme specific accessory subunits (Narlikar et al., 2013).

One of the key mechanism by which the remodelers differ from each other is the way they interact with the nucleosomes. Broadly, they can be divided into three categories (Hota and Bartholomew, 2011). The first category is defined on the basis of remodeler interaction with nucleosomal DNA through the DNA translocase domain of these proteins (Hota and Bartholomew, 2011). The other domains of the protein can also interact with the nucleosomal DNA and help in stabilizing interaction of the nucleosomes with the remodeler. The second category includes those remodelers that recognize the open face of the histone octamer that is not bound by DNA (Hota and Bartholomew, 2011). The third category are those proteins that interact through the extranucleosomal/linker DNA (Hota and Bartholomew, 2011) (Table 1.1).

Table 1.1: A Table showing Nuclear remodeling protein complexes known in humans along with their chromatin interacting functions.

Nucleosomal DNA interacting proteins	
Rad54	Rad54 protein helps in pairing between homologous DNA to carry out the homologous recombination. The crystal structure of Rad54 has revealed that it binds to translocating nucleosomal DNA (Dürr et al., 2005;) This function of Rad54 is important for chromatin metabolism (Thomä et al., 2005).
Chd1	Chd1 also helps in maintaining chromatin architecture. Chd1 consists of a chromodomain and a helicase/ATPase domain. The chromodomain of Chd1 are required for ATP-dependent nucleosomal mobilization. The crystal structure of Chd1 has also shown that it can bind to nucleosomal DNA (Hauk et al., 2010).
Isw2	Isw2 is a nuclear protein that helps in nucleosome sliding to remodel the chromatin. Experiments like photochemical crosslinking, DNA footprinting and structural modeling has shown that Isw2 interacts with nucleosomal DNA exiting from the histone octamer of a nucleosome and requires histone H4 tail for stabilizing its interaction with nucleosomes(Dang and Bartholomew, 2007).
Snf2	Snf2 proteins regulate many DNA processing activities by nucleosomal sliding and

	reposition. Experiments like photochemical crosslinking, DNA footprinting and structural modeling has displayed that Snf2 binds to the position where DNA is pointing towards histone octamer(Dang and Bartholomew, 2007). However, Snf2 makes fewer contacts with DNA and exposes 50bp of DNA by pushing the nucleosome away (Hota and Bartholomew, 2011).
ISWI	ISWI protein mobilizes nucleosomes to help in chromatin remodeling. By utililizing its HAND and SLIDE domain ISWI protein push the DNA towards entry side of nucleosome and hence smoothen the nucleosomal DNA progress towards histone octamer (Dang and Bartholomew, 2007).
Histone interacting proteins	
NoRC and ACF/CHRAC	These nucleolar remodeling complexes help in nucleosome sliding in an ATP-dependent manner. Acidic amino acid residues present in the ATPase domain of Isw2 containing complexes like NoRC and ACF help in binding to H4 tail of histones (Clapier et al., 2001; Shogren-Knaak et al., 2006; Strohner et al., 2001).
INO80	INO80 is a protein complex identified in yeast and possesses DNA-dependent ATPase activity. INO80 has been shown to modify H4 tail of histones and then bind to histone proteins and thereby helps in DNA replication and DNA damage response (Racki et al., 2014; Udugama et al., 2011).
SWI/SNF and RSC proein complex	SWI/SNF protein complexes get recruited at the promoter and helps in transcription activation and transcription repression. Snf2 and Sth1 subunit present in SWI/SNF and RSC protein complex, respectively, possesses Bromodomain to recognize acetylated histones (Hota and Bartholomew, 2011).
NURF	NURF protein which is an ISWI type remodeler contains a subunit, BPTF, that utilizies methylated histones with the help of its chromodomain and hence helps in recruiting NURF protein at methylated histone sites (Li et al., 2006; Wysocka et al., 2006).
Linker DNA interacting proteins	
Isw1 and Chd1	Isw1 and Chd1 rely on specific length of linker DNA, 30bp, for their efficient binding to nucleosomes (Gangaraju and Bartholomew, 2007).
Isw2 and INO80	Isw2 and INO80 require 20bp length of linker DNA for their binding to nucleosomes. For nucleosomal movement, INO80 require 53 bp of linker DNA while Isw2 require 70bp length of linker DNA (Kagalwala et al., 2004; Udugama et al., 2011).
SWI/SNF	SWI/SNF proteins do not depend on linker DNA binding for their remodeling activity.

The outcome of action of all these remodelers is to provide either a uniform space to nucleosomes or to disassemble nucleosomes. In this next section, I will discuss in detail the mechanism of mobilizing and disassembling nucleosomes by these ATP-dependent chromatin remodelers.

Interaction with nucleosomal DNA

The nucleosome remodelers belong to the SF2 family of helicases but do not possess any helicase activity. The crystal structure of the ATPase domain of very few remodelers has been characterized till now. The structure of ATPase domain of Rad54, alone as well as in complex with translocating dsDNA, from *Sulfolobus solfataricus* and zebrafish has been determined (Dürr et al., 2005; Thomä et al., 2005). In recent times, the crystal structure of ATPase domain of yeast Chd1 with chromodomains has also been solved and the solved structure shows a remarkable homology with the ATPase domain of Rad54 (Hauk et al., 2010). However, the two lobes of ATPase domain of Chd1 have a different orientation as compared to Rad54 and this difference does not allow Chd1 to bind to naked DNA. Thus, Chd1 can bind only to nucleosomal DNA (Hauk et al., 2010).

The interaction of the ATPase domains of Isw2 and Snf2 with nucleosomal DNA has also been determined with the help of photochemical crosslinking, structural modeling and DNA footprinting (Dang and Bartholomew, 2007). These studies have shown that Isw2 interacts with nucleosomal DNA exiting from the histone octamer while the Snf2 protein was found to bind at the position where DNA is pointing towards the histone octamer (Dang and Bartholomew, 2007; Hota and Bartholomew, 2011). The ISWI proteins also contain a HAND and SLIDE domain

which works in combination to push the DNA towards the entry side of the nucleosome and hence, smoothen the progress of the nucleosomal DNA around the histone octamer (Dang and Bartholomew, 2007). Further, the Isw2 requires histone H4 tail for chromatin remodeling but SWI/SNF does not (Dang et al., 2006). Mutations introduced in the H4 tail leads to defective nucleosomal remodeling by Isw2 as the interaction between the remodeler and the nucleosome is impaired (Dang et al., 2006). These experiments also showed that the H4 tail is needed for stabilizing the interaction with the ATPase domain of Isw2. The Isw2 also contains an auxiliary subunit, Itc, that is necessary for its interaction with the nucleosomal DNA (Kagalwala et al., 2004).

The Snf2 protein of the SWI2/SNF2 complex, makes fewer contacts with the nucleosomal DNA than Isw2. Its interaction mediated by the DNA translocase domain is localized to a 10 bp stretch of DNA (Dechassa et al., 2008). This way the Snf2 protein pushes the nucleosomes off the DNA exposing 50 bp of DNA (Kassabov et al., 2003).

The interaction of other proteins with nucleosomes DNA is yet to be characterized.

Interaction with histones

These proteins bind mainly to histones tail since these tails act as main marks for post-translational modifications. The H4 tail of histones is bound by the acidic amino acid residues, present in the ATPase domain of Isw2 containing protein complexes like NoRC and ACF/CHRAC (Clapier et al., 2001; Shogren-Knaak et al., 2006; Strohner et al., 2001). Some of the proteins like INO80 require this H4 tail to be modified for interaction, while proteins like

ACF loses its connection with nucleosomes possessing acetylated H4 tail at lysine 16 (Racki et al., 2014; Udugama et al., 2011).

The Snf2 and Sth1 subunit of SWI/SNF and RSC protein complex, respectively, possess the signature bromodomain motif that recognizes acetylated histones (Hota and Bartholomew, 2011). Other chromatin remodelers contain a PHD finger or a chromodomain that recognizes methylated lysine present in histones (Wysocka et al., 2006). BPTF, a protein subunit of NURF which is an ISWI type remodeler, recognizes the methylated histones and helps in recruitment of NURF complex at methylated nucleosomal sites (Li et al., 2006; Wysocka et al., 2006).

The SWI/SNF and RSC protein complexes can also interact with the globular regions of the H2A/ H2B docking domain (Hota and Bartholomew, 2011). Variants of H2A like H2A.Bbd and H2AL which are defective in this acidic domain cannot interact with SWI/SNF and RSC complex (Doyen et al., 2006; Syed et al., 2009). However, the H2A variant H2A.Z has an enhanced acidic domain which forms chromatin sites that promotes the recruitment of SWI/SNF proteins, like in case of GAL promoter (Lemieux et al., 2008). A globular region specific modification, acetylation of lysine 56 in histone H3, promotes the binding of SWI/SNF proteins, but has no effect on RSC complex recruitment and their remodeling at chromatin region (Neumann et al., 2009; Xu et al., 2005). All these data support the more extensive binding of SWI/SNF protein complex with the histones while the ISWI protein complex interacts more with the nucleosomal DNA.

Interaction with linker DNA

The chromatin remodelers require a definite length of linker DNA for interaction with the nucleosomes as well as for their mobilization (Hota and Bartholomew, 2011). The first remodelers identified to have a strong reliance on linker DNA was the ISWI family of proteins (Kagalwala et al., 2004; Zofall et al., 2004). These proteins were shown to arrange nucleosomes with uniform distribution of linker DNA between flanking nucleosomes, a process referred as nucleosome spacing (Tsukiyama et al., 1999). Isw1 and Chd1 require 30 bp of linker DNA for efficient binding while INO80 and Isw2 require approximately 20 bp linker DNA for binding (Gangaraju and Bartholomew, 2007). The importance of the length of the linker DNA for the remodeling activity was also proven by the fact that INO80 needs 53 bp of DNA for nucleosome movement while Isw2 requires 70 bp of length for optimal nucleosomal movement (Kagalwala et al., 2004; Udugama et al., 2011). The ACF complex that contains the ISWI proteins has also been shown to mobilize nucleosomes after interacting with linker DNA (Yang et al., 2006). The SWI/SNF proteins do not depend on the linker DNA for their remodeling activity.
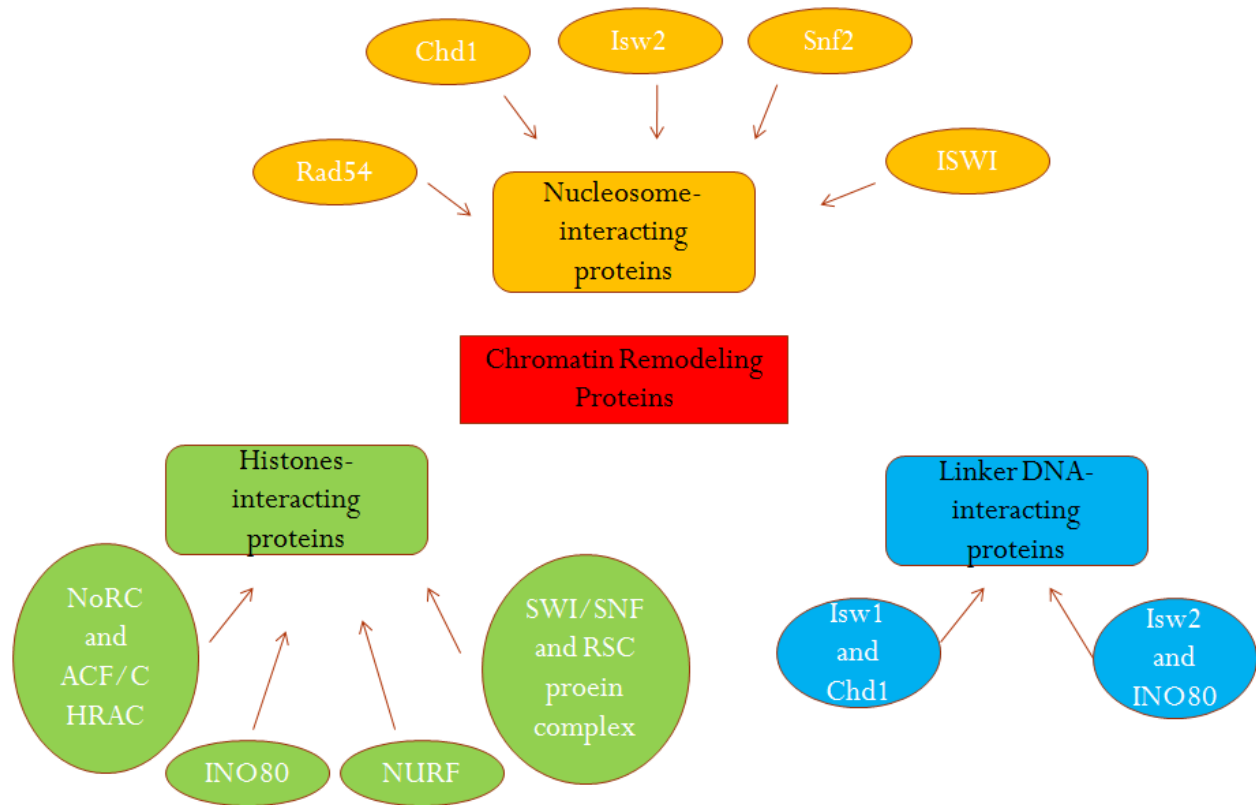
Figure 1.3: A Cartoon representing different Chromatin Remodeling Proteins and their interaction with nucleosomes and nucleosomal DNA.

Mechanism of ATP-dependent chromatin remodeling

The mechanism by which the ATP-dependent chromatin remodeling proteins distribute nucleosomes is still a paradox (Fan et al., 2003). The crystal structure of nucleosome reveals around 100 histone–DNA interactions stabilizes the nucleosome assembly (Luger et al., 1997; Richmond and Davey, 2003; Suto et al., 2003). Hence, mobilization of nucleosomes is strongly dependent on the ability of remodelers to overcome this energy barrier. Therefore, different mechanisms/models have been put forward to explain the nucleosome mobilization by the ATP-

dependent chromatin remodelers (Flaus and Owen-Hughes, 2004). In the following section, I'll discuss selective models that have been put forward to explain the mechanism of nucleosomal/chromatin remodeling.

Twist defect diffusion

This model states that small alterations in the DNA referred as twists generate torsion on the surface of DNA that is propagated along the DNA and forces the adjoining regions of nucleosomes to move forward and backward along the DNA. The crystal structure of nucleosome core particle validates the presence of these twists along the DNA (Luger et al., 1997; Richmond and Davey, 2003; Suto et al., 2003). Co-crystallization of the nucleosomal core with DNA has shown that the locations of these twists can be shifted (Richmond and Davey, 2003; Suto et al., 2003). The remodeling enzymes alter the superhelical torsion and relieve the torsion stress generated within DNA fragments by accelerating the rate of twist defect diffusion along a single direction (Figure 1.4) (Cairns, 2007).

Bulge diffusion

This model states that a bulge is created in the nucleosomes when DNA is unwrapped from the entry/exit site of nucleosome and gets wrapped at a distal position on a histone octamer (Flaus and Owen-Hughes, 2004). The bulging out of DNA results in incorporation of approximately 40 bp free DNA, resulting in nucleosomes sliding as proved by both *in vitro* and *in vivo* studies (Figure 1.3) (Kulić and Schiessel, 2003). The remodelers like RSC and SWI/SNF force the DNA

into loops/bulges by continuously pulling it such that the DNA moves towards the nucleosomal entry/exit site, thus effecting nucleosomal movement (Zhang et al., 2006).

The above models suggest that the histone octamer remains unaltered during nucleosomal remodeling. However, some recent findings show that the histone dimer in an octamer can be exchanged or removed during remodeling reactions (Flaus and Owen-Hughes, 2004).

Many chromatin remodelers have been identified that replace a dimer of the nucleosome with another (Flaus and Owen-Hughes, 2004). For example, Swr1p manipulates the histone composition by incorporating histone variants at many nucleosome positions (Kobor et al., 2004). Swr1p interacts with the histone variant Htz1p in *S. cerevisiae* and utilizes the energy of ATP to replace H2A type Hta1p-Htb1p dimer with Htz1p-Htb1p dimer (Figure 1.4) (Kobor et al., 2004; Wan et al., 2009). This variant incorporation by Swr1p prevents the spread of the hetrochromatanization in budding yeast (Wan et al., 2009).

All these reports suggest the importance of ATP-dependent chromatin remodelers in altering the chromatin by utilizing different mechanism. In the following section, I will be focusing on the different members of the ATP-dependent chromatin remodeling protein family.

A: Twist defect diffusion

B: Bulge diffusion

Nucleosome repositioning/Sliding

C:DNA unwrapping

Nucleosomal site exposure
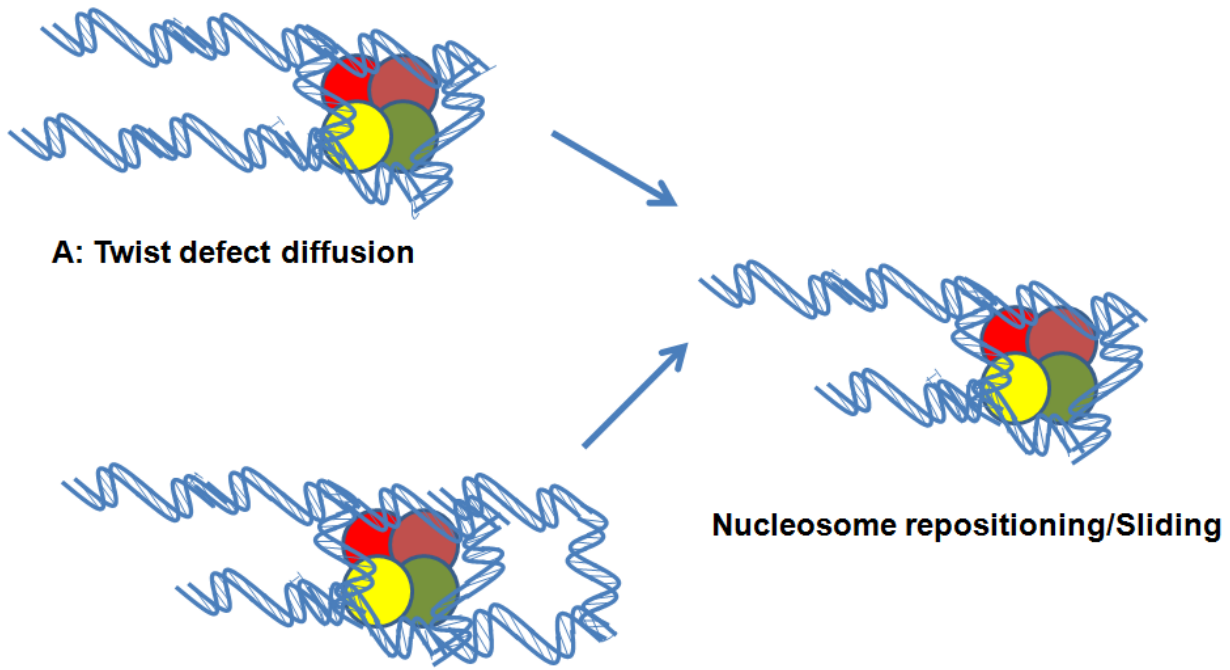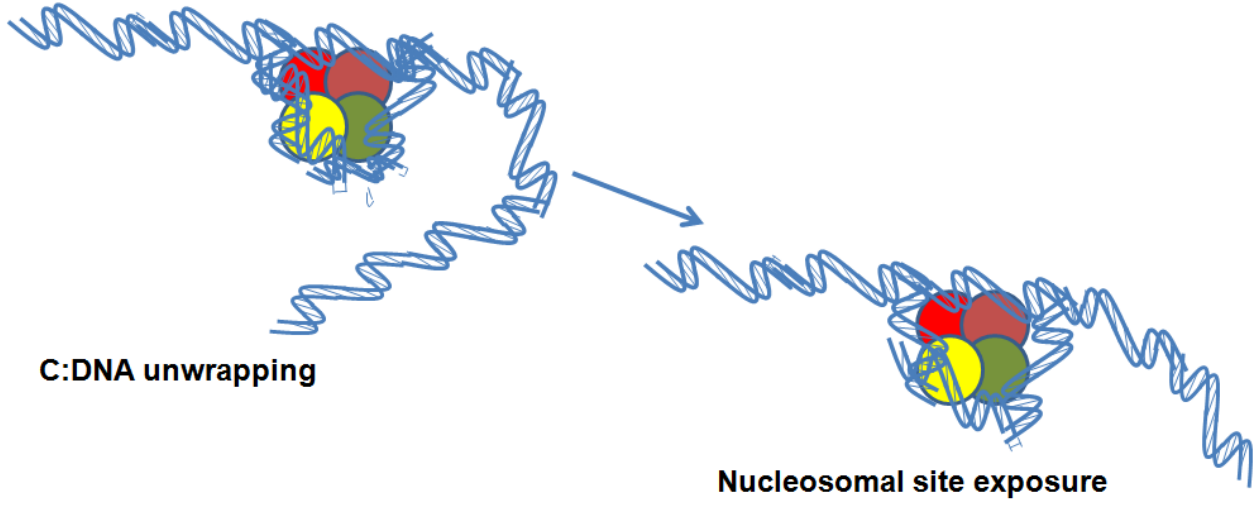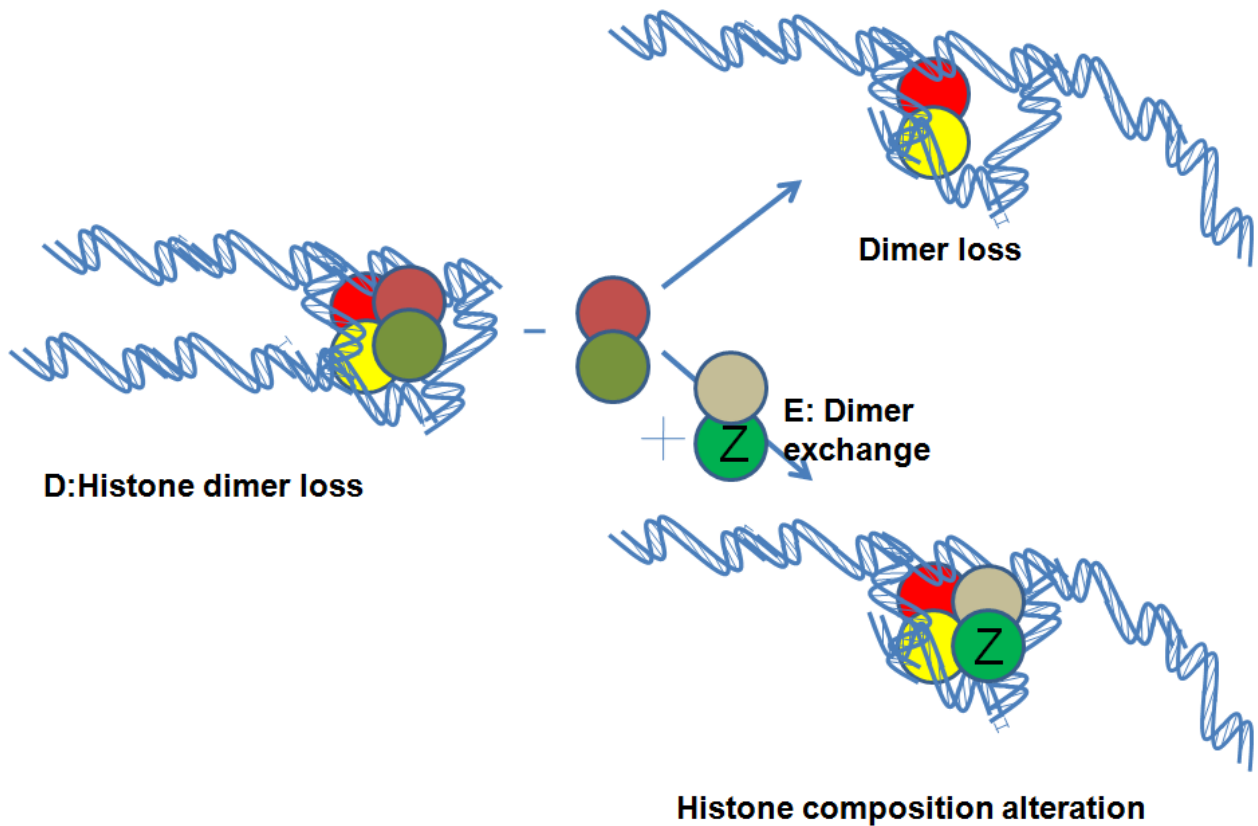
Figure 1.4: Mechanism of action of different chromatin remodelers: (a) Twist defect diffusion along the nucleosomal surface, (b) Bulge diffusion resulting in nucleosome repositioning, (c) DNA unwrapping, (d) Histone dimer loss during nucleosomal remodeling and, (e) Histone dimer exchange leading to change in histone composition.

Snf2 subfamily

The Snf2 is a component of a 11 subunit complex called SWI/SNF that plays a functional role in transcription of many genes (Becker and Hörz, 2002). The homologues of the Snf2 proteins have

been identified in many different organisms (Laurent et al., 1992). These include the RSC complex that contains the paralogue Sth1 in *S. cerevisiae* (Cairns et al., 1996) and, orthologues Brahma in *Drosophila melanogaster* (Tamkun et al., 1992) and, hBRM and BRG1 in humans (Muchardt and Yaniv, 1993). In addition to the catalytic ATPase subunit, the Snf2 family of proteins contains a bromodomain and an AT-hook region (Figure 1.5). The bromodomain has been shown to bind to the acetylated histone tails of the nucleosome and the AT-hook region has been shown to help in binding to the AT rich regions of DNA and thereby, both these auxiliary domains assist in targeting the protein to the chromatin (Aravind and Landsman, 1998; Boyer et al., 2004). The Snf2 family members help in regulation of transcription either through their disruptive function on nucleosomes leading to nucleosome repositioning or by the alteration of histone-DNA contacts resulting in removal of histone from the nucleosomes (Becker and Hörz, 2002). These proteins are involved in activation of inducible genes involved in metabolism and mating-type switching (Winston and Carlson, 1992). They have also been shown to regulate late mitosis both in yeast and *C. elegans* (Krebs et al., 2000). The BRM protein of *D. melanogaster* is a component of two complexes-BAP and PBAP-both of which co-localize to the hyperacetylated chromatin region and repress the expression of genes involved in fly development (Mohrmann et al., 2004). In mammalian cells, hBRM and BRG1 play an important role in development, cell differentiation and tumor suppression (Hendricks et al., 2004; Neely and Workman, 2002). The Snf2 protein complexes have been shown to interact with histone deacetylases, histone methyl transferases, methyl-DNA binding proteins, RNA polymerase II, histone chaperones and cohesion complex, and thus, can regulate a wide variety of processes (Dunaief et al., 1994; Harikrishnan et al., 2005; Sif et al., 2001).

ISWI subfamily

The ISWI proteins identified through their sequence similarity to Snf2p protein of *D. melanogaster* are present as the catalytic core of NURF and ACF/CHRAC chromatin remodeling complexes (Elfring et al., 1994; Ito et al., 1997; Tsukiyama et al., 1995; Varga-Weisz et al., 1997). The ISWI protein complexes are found in yeast, flies and higher eukaryotes. Unlike the SWI2/SNF2 proteins, the ISWI protein contains SANT and SLIDE auxiliary domains (Figure 1.5). These two domains are important for interactions with histone tail and nucleosomal DNA respectively (Grüne et al., 2003). Biochemical studies have shown that the ISWI proteins play a role in nucleosome repositioning rather than in nucleosome disruption (Längst et al., 1999). Importantly, these proteins require a specific stretch of histone H4 tail near the DNA surface for interaction as well as nucleosome repositioning (Clapier et al., 2001). The ISWI family members participate in many nuclear processes and are involved in functional interaction with various protein complexes. In yeast, ISW1a and ISW1b both coordinate to regulate the elongation and termination by RNA polymerase II and mRNA processing during transcription (Morillon et al., 2003). The ACF protein complex composed of ISWI and Acf1, in flies and mammals, is involved in assembling and mobilizing the nucleosomes (Ito et al., 1997). NURF, on the other hand, is involved in randomization of already spaced nucleosomes. Thus, the ISWI proteins regulate transcription initiation and elongation. They also participate in DNA replication and chromatin assembly (Dirscherl and Krebs, 2004; Morillon et al., 2003; Tsukiyama, 2002).

Chd1 subfamily

The Chd1 proteins were so named due to the presence of chromodomain (Chromatin Organizer Domain) motif in the N-terminal region of the protein (Woodage et al., 1997). These proteins also contain the signature ATPase motifs of the SNF2 family (Woodage et al., 1997). In addition, domains like SANT-like, BRK, DNA binding and PHD (plant homeo domain) have also been identified in members of this subfamily (Figure 1.5) (Marfella and Imbalzano, 2007). Chromodomain helps the Chd1 proteins to bind to DNA, RNA and methylated histones (Marfella and Imbalzano, 2007). The Mi-2 proteins of the Chd1 subfamily link DNA methylation to chromatin remodeling and histone deacetylation, thus play a role in transcriptional repression (Längst and Manelyte, 2015). Other members like CHD5 have been shown to play a role in neural development (Ho and Crabtree, 2010).

Swr1 subfamily

The Swr1 subfamily is made up of proteins like Ino80, Swr1, EP400 and Etl1. The Swr1p protein from yeast is a part of Swr1p complex that mediates exchange of H2A dimers with the variant H2A.Z dimers (Figure 1.5) (Krogan et al., 2003; Mizuguchi et al., 2004). The homologue of yeast Swr1, Domino from *D. melanogaster,* is also involved in histone variant exchange in an acetylation-dependent manner (Kusch et al., 2004). Ino80 protein from *S. cerevisae* was first identified for its role in transcriptional regulation of inositol biosynthesis (Bachhawat et al., 1995; Ebbert et al., 1999). In addition to nucleosome repositioning, the specific feature that

separates Ino80 protein complexes from other Snf2 family related proteins is its ability to separate DNA strands in a classical helicase assay due to the presence of two RuvB helicases (Jin et al., 2005; Shen et al., 2000). The Ino80 proteins also participate in DNA damage repair pathway as the Ino80 deleted strains show defect in double-strand break repair (Morrison et al., 2004; van Attikum et al., 2004). The *INO80* mutants also showed sensitivity towards hydroxyurea, ultraviolet light and ionizing radiation supporting their role in DNA replication and DNA damage repair (Morrison et al., 2004; van Attikum et al., 2004). Finally, Etl1 protein members like human SMARCAD1 and yeast FUN30 are widely expressed proteins that function in developmental pathways (Adra et al., 2000; Clark et al., 1992; Schoor et al., 1999). FUN30 mutants exhibit decreased sensitivity towards ionizing radiation and the deletion of FUN30 results in temperature sensitivity(Barton and Kaback, 1994; Ouspenski et al., 1999).

Rad54 subfamily

The Rad54 proteins were identified because their inactivation led to increased sensitivity towards ionizing radiation (Tan et al., 2003). The role of Rad54 proteins in mediating Rad51-dependent single-strand invasion into DNA duplex during homologous recombination mediated DNA repair has been well-studied (Klein, 1997; Krogh and Symington, 2004). These proteins catalyze ATP hydrolysis in the presence of double-stranded DNA and translocate along the DNA to induce a change in the nucleosome accessibility (Alexeev et al., 2003; Alexiadis and Kadonaga, 2002). The crystal structure of helicase region of zebra fish Rad54 has been determined in recent times

and I will describe it in detail in later section (Thomä et al., 2005). The ATRX subfamily of proteins, so named because mutations in ATRX causes Alpha Thalassemia, an X linked genetic disorder (Gibbons et al., 1995). ATRX plays a role in regulation of heterochromatin structure as well as transcription by increasing the accessibility of nucleosomal DNA, although the precise mechanism is still not known (Gibbons et al., 2000; McDowell et al., 1999).

Rad5/16 subfamily

The Rad5/16 group consists of subfamilies Rad5/16, Ris1, Lodestar and SHPRH (Flaus et al., 2006). Except for Lodestar subfamily, the proteins of the Rad5/16 group are characterized by the presence of a RING finger (Chen et al., 2005). This motif has been shown to help in interaction with protein complexes like Ubc13p-Mms2p (Hoege et al., 2002). The helicase region of Rad5p participates in double-stranded DNA break repair (Chen et al., 2005). The Lodestar proteins are recognized as cell-cycle regulated proteins involved in mitosis (Girdham and Glover, 1991). TTF2 proteins, a human homologue of Lodestar subfamily, participates in termination of elongating RNA pol I and pol II, perhaps by clearing the RNA polymerases from the DNA template at the entry into mitosis (Jiang et al., 2004). The role of SHPRH possessing PHD finger is known in transcription regulation, but the exact function is still to be studied (Aasland et al., 1995).

SSO1653-like subfamily

The SSO1653-like group consists of subfamilies ERCC6, Mot1 and SSO1653 (Flaus et al.,

2006). Mot1 proteins of yeast and its homologues BTAF1 and TAF172 proteins are present

across fungi and other higher eukaryotes (Chicca et al., 1998). Mot1p does not play a canonical

Snf2 related role in nucleosome repositioning but interacts with TBP (TATA-binding protein)

and helps in recycling its DNA bound states (Auble et al., 1997; Dasgupta et al., 2005). Human

ERCC6 proteins, also called as CSB (Cockayne Syndrome B), and its homologue Rad26p in *S.

cerevisae* assist in transcribing RNA polymerase II to dissociate or surpass from blocked DNA

lesions (Svejstrup, 2003). These proteins, thus, play a role in nucleotide excision repair

(Svejstrup, 2003). ERCC6 proteins also help in nucleosome spacing by harnessing the energy

released after ATP hydrolysis for this activity (Citterio et al., 2000). The biological role of

Sso1653p, an archaeal member of this group, is still unknown, although its role in generating

DNA torsion and DNA-dependent ATP hydrolysis has been studied (Liu et al., 1995).

SMARCAL1/ RapA subfamily

SMARCAL1 and RapA subfamily of proteins are classified as distant members of the SNF2

family of proteins. SMARCAL1 subfamily of proteins contains two subtypes, both containing

exclusive auxiliary domains in addition to the classical helicase motifs. The first subtype

contains the proteins related to human SMARCAL1 and possesses HARP domains immediately

N-terminal to the centrally localized helicase region. Biallelic mutation in SMARCAL1 causes a

disease Schimke immuno-osseous dysplasia (SIOD), revealing the biological importance of SMARCAL1 (Boerkoel et al., 2002). I will be elaborating on SMARCAL1 later in this chapter.

The second subtype is present in the genomes of plants, protists and animals and its domain-wise organisation is like that of ZRANB3 protein in humans. ZRANB3 and ZRANB3-like proteins contain helicase region at the N-terminus followed by a Zinc finger domain similar to the one present in Ran binding proteins, and an endonuclease domain at the C-terminus (Plambeck et al., 2003).

Proteins belonging to the RapA subfamily are predominantly found in bacterial and archaeal organisms. These proteins contain the characteristic helicase-like region with a definite spacing of at least 160 amino acid residues between motif III and motif IV, but the central region is devoid of conserved features and hence, varies from other SNF2 family of proteins (Flaus and Owen-Hughes, 2001). RapA, also called as HepA, releases the stalled polymerases from the transcription sites and helps in polymerase recycling under high salt condition (Muzzin et al., 1998; Sukhodolets et al., 2001).
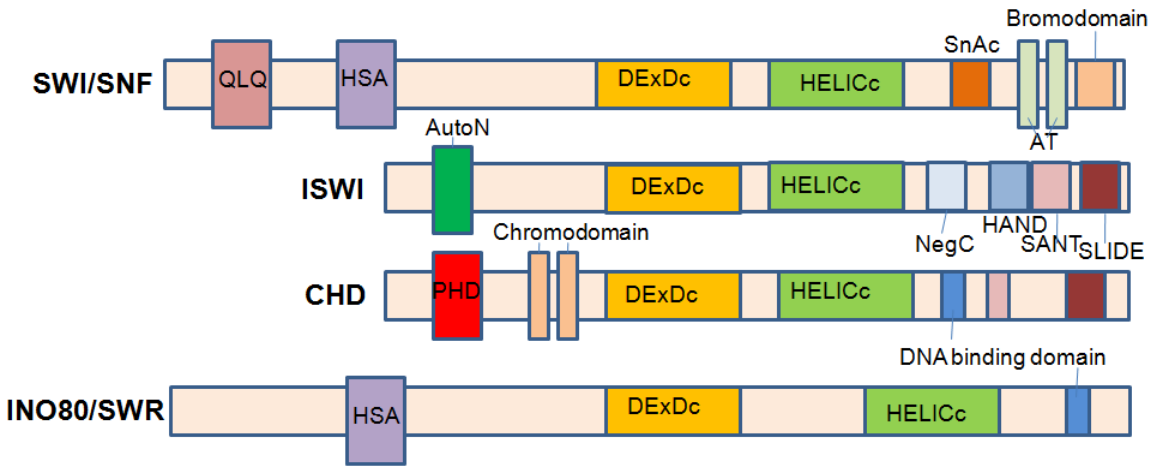
Figure 1.5: The domain organization of some of the SNF2 family proteins.

Crystal structure of ATPase core domain

The crystal structure of ATPase domain of *S. solfataricus* (SSO1653*),* zebrafish Rad54,

S.cerevisiae Chd1 and, *E. coli* RapA has been solved (Dürr et al., 2006; Shaw et al., 2008; Hauk

et al., 2010; Thomä et al., 2005). According to these studies, the catalytic core is made up of two

domains with each containing a RecA type α/β sub domains mentioned as 1A and 2A,

respectively (Dürr et al., 2005). The seven helicase motifs necessary for ATPase activity occupy

the positions in loop regions of these subdomains (Dürr et al., 2005). Motifs Q, I, IA, II and III

are present in subdomain 1A while motifs IV, V and VI are present in subdomain 2A. In addition

to 1A and 2A subdomain, 1B and 2B subdomains are also present in RecA domains 1 and 2

(Dürr et al., 2005). Subdomains 1B and 2B helps in translating energy released from ATP-

mediated conformational changes between subdomains 1B and 2B into DNA translocation (Dürr

et al., 2005). A deep cleft, which acts as an active site cleft, is created between domains 1 and 2

34

(Figure 1.6) (Dürr et al., 2005). Motifs I, II and III occupy this cleft and form the ATP binding site. The remaining motifs IV, V and VI occupy the outside positions and thus, the protein attains a closed conformation (Dürr et al., 2005). In the presence of nucleotides, domain 2 undergoes an 180° flip to attain an open conformation as observed in the crystal structure of ATP bound helicases (Figure 1.6) (Dürr et al., 2005).

The structure and amino acid sequences of motif I, which binds ATP, and motif II, which interacts with magnesium ion are highly conserved between Rad54, SWI/SNF protein and other classical helicases (Dürr et al., 2005). The remaining motifs like, motif 1a, TxGx and III in subdomain 1A possess some structural differences between Rad54 and other canonical helicases. Motif III of Rad54 is found to be structurally similar to the other SF2 helicases where it acts as a sensor for ATP hydrolysis (Dürr et al., 2005). This motif has been implicated in binding to both the gamma-phosphate of ATP as well as DNA. It can also interact with domain 2 of the helicase lobe. Motif IV is highly conserved in SF1/SF2 helicases and SWI/SNF proteins (Dürr et al., 2005). This motif has been found to interact with the amide backbone of DNA at the start of double helix through its highly basic residues lysine and arginine (Dürr et al., 2005). Rad54 possesses a lysine moiety at the spatially same position to provide a potential DNA contact (Dürr et al., 2005). The roles, sequences and lengths of helicase motifs V and VI differ among SF1 and SF2 helicases. In PcrA and Rep, members of SF1 family, the conserved amino acids of motif V interact with DNA (Korolev et al., 1997; Velankar et al., 1999). SF2 helicase members like UvrB, HCV and eIF4A did not show specific interactions with DNA or ATP (Caruthers et al., 2000; Theis et al., 1999; Yao et al., 1997). However, Rad54 possesses conserved amino acid residues lysine/arginine (lys 568), serine/ threonine (Ser 566 and 567) and glycine (Gly 570) that make contact with DNA (Dürr et al., 2005). Motif VI has been observed to possess a conserved

arginine in both SF1and SF2 helicases. While SF1 helicase members, Rep and PcrA, has shown the association of the conserved arginine with ATP, SF2 helicases like eIF4A, HCV and UvrB has shown that this arginine mediates DNA binding and inter-domain communication with motif II which leads to conformational changes in the protein during ATP hydrolysis (Caruthers et al., 2000; Theis et al., 1999; Yao et al., 1997). However, the crystal structure of Rad54 has shown that the equivalent arginine, arginine 600 of motif VI, interacts with the sulphate ion bound to ATP-gamma phosphate binding site of Motif Ia exactly as in PcrA (Dürr et al., 2005).

The duplex DNA was found to bind to Rad54 along the entrance of the active site cleft between domain 1 and domain 2 where ATP binding site is present (Figure 1.6) (Dürr et al., 2005). Predominantly, subdomain 1A binds to double-stranded DNA via the phosphate chains present in the minor groove of DNA (Dürr et al., 2005). The crystal structure of Rad54 has shown that DNA is initially present in the B-DNA conformation with both the strands bound to each other tightly, but as the enzymes proceed along the translocation they induce widening of the minor groove (Dürr et al., 2005). Motifs Ia and II of subdomain 1A were found to be involved in DNA binding. Domain 2 also showed an interaction with DNA phosphate backbone (Dürr et al., 2005). However, the binding towards domain 2 was much weaker as analyzed by the binding constant affinity of purified domain 1 and domain 2 towards DNA (Dürr et al., 2005). Binding towards DNA was not increased by ATP or by its non-hydrolysable analog, indicating that DNA binding by Rad54 is not affected by ATP (Dürr et al., 2005). However, ADAAD (N-terminal truncated, ATPase domain of bovine homolog of SMARCAL1) has shown that stem-loop DNA binding is stimulated upon binding to ATP (Nongkhlaw et al., 2009). Still, the exact DNA binding site and ATP hydrolysis mechanism needs to be identified. I will describe the detailed biochemical studies of ADAAD later in my thesis. Members of DExx box helicases show the presence of two

ATP-driven processes for catalyzing helicases activity-translocation on the downstream single-stranded DNA and destabilization of upstream double helical DNA (Dürr et al., 2006). SsoRad54cd, when compared with NS3 and PcrA, shows the presence of same RecA-like domains and helicase motifs I, II and III (Figure 1.6) (Dürr et al., 2005). Even, 3′-5′ DNA strand and 3′-5′ oligo-dU was also found to occupy the same positions in SsoRad54cd and NS3, respectively (Figure 1.6) (Dürr et al., 2005). However, some additional domains are also present in NS3 and PcrA helicases that make contact with the nucleotide bases of single-stranded DNA which are absent in Rad54 (Dürr et al., 2005). Rad54 lack single-stranded DNA binding domain as well as the ssDNA-stimulated ATPase activity (Dürr et al., 2005). Additionally, double-stranded DNA was found to present firmly along the entire side of enzyme (Figure 1.6) (Dürr et al., 2005). The helix destabilizing region, necessary to unwind double helix, was also found to be absent in SsoRad54 (Dürr et al., 2005). All these data prove the importance of SWI/SNF proteins in DNA translocation, but not in duplex destabilization.
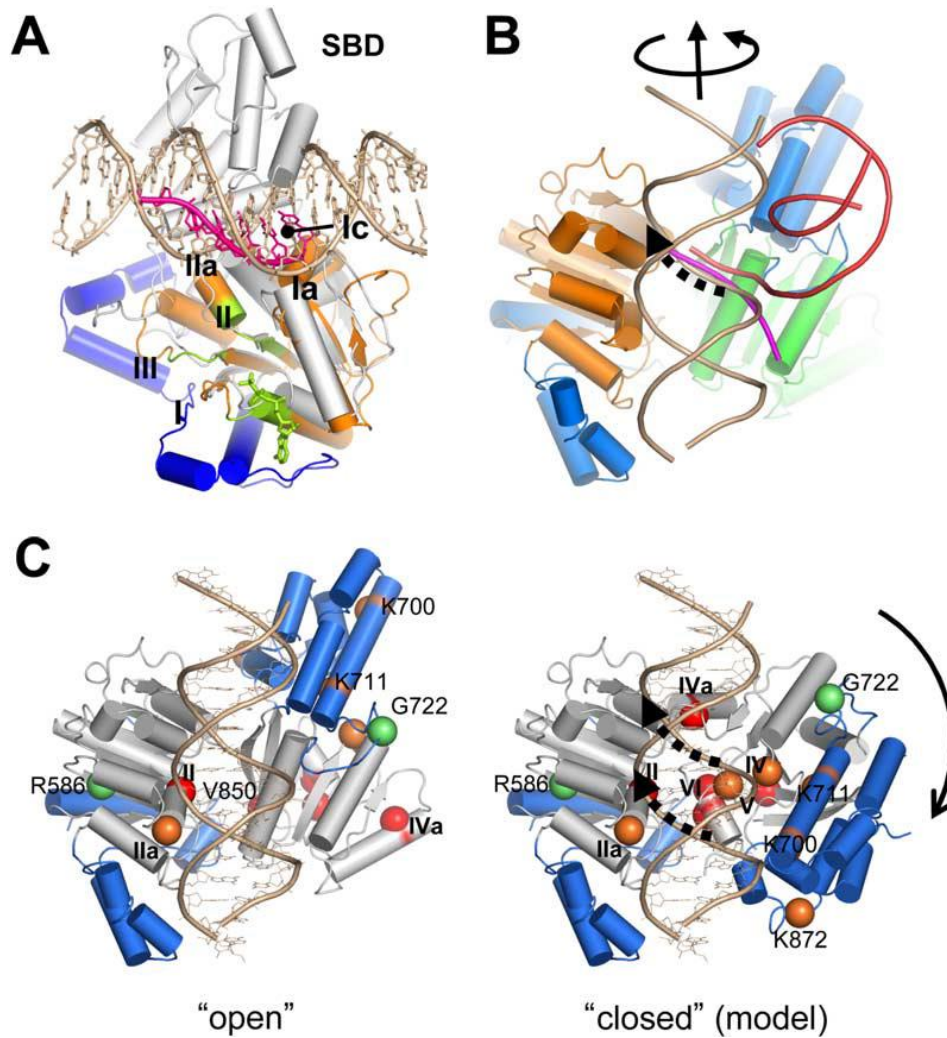
Adapted from: (Dürr et al., 2005)

Figure 1.6: Comparison of different helicases and a model depicting ATP-induced

conformational changes. (A) Overlay of models of Sso-Rad54cd domain 1 with the similar

regions of Hepatitis C virus helicase NS3. The ATPase motifs and DNA strands (3'-5' dsDNA

strands of SsoRad54cd (brown) and oligo-dU of NS3 helicase (magenta)) of both helicases

superimpose well. One additional domain, ssDNA binding domain (SBD), is present in NS3

helicase. SBD is absent in Sso-Rad54cd and hence helicase activity is absent in SWI2/SNF2

proteins. (B) Comparison of models of DNA bound to different helicases. The helicases used for

comparison are Sso-Rad54cd (model plus brown ds DNA), NS3 (only oligo-dU shown in magenta), and PcrA (only ssDNA/dsDNA shown in red). Double-stranded DNA is transported along the SWI2/SNF2 proteins in a same way as single-strand DNA is scanned by helicases (dashed arrows). (C) Crystallographic structure Sso-Rad54cd which is in open conformation (Left) and model for the putative closed conformation after binding to ATP and DNA (Right) (Dürr et al., 2005).

SMARCAL1 and its functional role

As stated earlier, SMARCAL1, a distant member of Snf2 family of proteins, possesses the helicase-like motifs flanked by specific auxiliary domains (Flaus et al., 2006). It has DNA-dependent ATPase activity and like the other members of the Snf2 family lacks classical helicase activity (Coleman et al., 2000). SMARCAL1 possesses two HARP domains at the N-terminal positions, immediately upstream of helicase motifs (Coleman et al., 2000). Mutations in the helicase-like motifs of SMARCAL1 causes an inherited, biallelic, autosomal recessive genetic disease SIOD which is characterized by spondyloepiphyseal dysplasia, growth defects, renal dysfunction, T-cell immunodeficiency and cancer predisposition (Boerkoel et al., 2002). Genome-wide linkage mapping and positional candidate approach led to the identification of SMARCAL1 as the gene responsible for SIOD (Boerkoel et al., 2002; Clewing et al., 2007; Santangelo et al., 2014).

SMARCAL1 has been shown to recognize DNA molecules that possess double-strand to single-strand transition regions like stem-loop DNA and fork DNA (Muthuswami et al., 2000). Additionally, the protein possesses ATP-stimulated strand annealing activity by which

SMARCAL1 rebinds RPA bound single-stranded DNA present at DNA damage sites (Yusufzai and Kadonaga, 2008).

Earlier, SMARCAL1 has been proposed to regulate the transcription of c-*myc* and c-*kit* by altering their promoter DNA regions (Baradaran-Heravi et al., 2012). In our lab, SMARCAL1 has been shown to play an important role in transcription regulation of c-*myc*. In serum starvation condition, SMARCAL1 has been shown to stabilize the secondary structures at the promoter regions of c-*myc* gene, thereby, resulting in transcriptional inhibition of c-*myc* (Sharma et al., 2015). I will describe the role of SMARCAL1 in c-*myc* transcription regulation in detail in my third chapter.

SMARCAL1 has also been shown to play a vital role in DNA damage response, cell cycle progression and stalled replication fork restart (Bansbach et al., 2009). SMARCAL1 travels along the DNA replisome during S-phase and gets concentrated at stalled replication fork sites (Ciccia et al., 2009). Not only SMARCAL1 occupies the positions at DNA damaged sites, but also it repairs DNA damage by catalyzing fork regression and replication fork restoration (Ciccia et al., 2009; Mason et al., 2014). RPA, a single-strand DNA binding protein, directs SMARCAL1 to damaged sites to catalyze replication fork restart (Bansbach et al., 2009). Mutations in SMARCAL1 causes hypersensitivity of cells towards replication stress and accumulation of double-stranded breaks at damage sites (Couch et al., 2013). In cells lacking SMARCAL1, stalled forks become the substrates for endonucleases that yield double-strand breaks making the cells prone towards chromosomal disarrangements and apoptosis (Couch et al., 2013).

Too much activity or too less activity of SMARCAL1 causes the formation of double-strand breaks (Couch et al., 2013). SMARCAL1 maintains its fork regression activity only in ATR proficient cells as ATR regulates the optimal activity of SMARCAL1 (Couch et al., 2013). In particular, ATR phosphorylates serine residue of SMARCAL1 at 652 positions after SMARCAL1 interacts with replication fork and hinder its remodelling activities (Couch et al., 2013). Treatment of cells with ATR inhibitor causes fork collapse and cell death because of the dysregulation of SMARCAL1 (Couch et al., 2013).Ni^{2+}

In addition to generation of stalled replication forks, one more source of replication stress is the formation of TTAGGG repeats of telomeric DNA (Chan et al., 2009; Günes and Rudolph, 2013). Telomeres thus formed, leads to generation of unusual DNA structures leading to replication fork inactivation (Cox et al., 2000). SMARCAL1 has been reported to resolve this replication stress by promoting the replication at telomeric sequences (Chan et al., 2009; Günes and Rudolph, 2013). SMARCAL1 role in maintaining telomeric DNA stability has been further proved by its inactivation which leads to the accumulation of circular telomeric DNA (Chan et al., 2009; Günes and Rudolph, 2013). The SMARCAL1 activity at telomeres is different from its genome maintenance activity as SMARCAL gets localized at telomeres through its high affinity DNA binding site even in the absence of RPA (Chan et al., 2009; Günes and Rudolph, 2013).

Thus, all these studies signify the importance of SMARCAL1 in maintenance of genomic stability both in DNA damage repair and in endogenous source of replication stress occurring at the telomeres.

Active DNA-dependent ATPase A (ADAAD)

To further study the biochemical activities of SMARCAL1, the C-terminal proteolytic fragment

of bovine homolog of SMARCAL1 containing the catalytic ATPase domain was generated in

Dr. Joel Hockensmith's laboratory, and termed as Active DNA-dependent ATPase A Domain

(ADAAD) (Muthuswami et al., 2000) (Figure 1.8). ADAAD has been shown to recognize DNA

in a structure-specific manner, whereby it interacts predominantly with double-stranded to

single-stranded transition regions, structural DNA elements seen in DNA-dependent metabolic

processes like replication, transcription and repair (Muthuswami et al., 2000).

Previous studies performed in our lab have shown that ADAAD can bind to stem-loop DNA and

ATP, both, independently of each other to form a binary protein complex E.DNA or E.ATP

respectively (Nongkhlaw et al., 2009). Then, binding to the other ligand induces a ternary

complex formation E.DNA.ATP or E.ATP.DNA that can result in ATP hydrolysis (Nongkhlaw

et al., 2009). Although both the ligands can bind first to form a binary complex, there is a

preferred sequence to the ADAAD-ligand binding, where stem-loop DNA binds first to the

protein followed by ATP binding, resulting in ATP hydrolysis (Nongkhlaw et al., 2009).

Based on these observations made in ADAAD and the crystal structure of Rad54, a model has

been proposed to explain the ATP hydrolysis mediated by this protein in the presence of DNA. It

has been proposed that ADAAD in the absence of the ligands is present in an open/inactive

conformation (Figure 1.7) (Nongkhlaw et al., 2009). Interaction with the effector DNA induces a

conformational change that result in 10-fold increased affinity for ATP. Binding of both ATP

and DNA induces a closed/active conformation that is competent for ATP hydrolysis (Figure

1.6). Furthermore, biochemical studies performed with mutations present in SIOD patients

(A468P, I548N and S579L) have proved that ligand binding induces a conformational change in

the protein necessary for ATPase function of the protein (Gupta et al., 2015). Although this plausible mechanism has been suggested for DNA-dependent ATP hydrolysis, yet the amino acid residues and the site which contacts DNA remains a mystery.
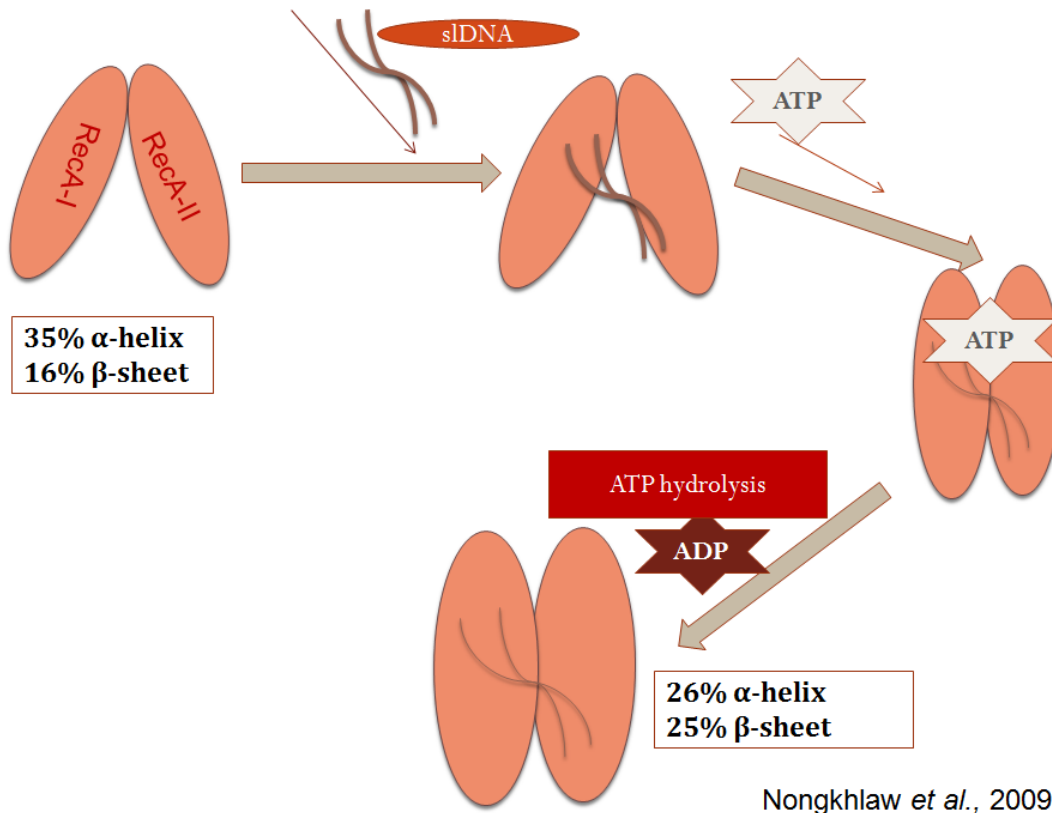
Figure 1.7: A model based on previous studies which show an open conformation/inactive form of the ADAAD protein. Stem-loop DNA binding followed by ATP binding leads to a conversion into closed conformation which is an active form of protein that catalyzes ATP hydrolysis (Nongkhlaw et al., 2009).

As stated earlier, ATP hydrolysis and ligand binding is an intrinsic property of helicases. The conserved helicase motifs have been shown important for mediating a coupling between ATP binding, DNA binding and ATP hydrolysis. The function of these motifs has been studied in detail in DEAD box proteins but not in SWI/SNF family of proteins. In our laboratory, motifs Q and I, conserved amino acids present in RecA-like domain 1A, have been found to be important for ATP hydrolysis but not for ligand, stem-loop DNA and ATP, binding (Nongkhlaw et al., 2012). Thus, we have got some biochemical information regarding the importance of RecA-like domain 1A of ADAAD, but we lack the information about the role of RecA-like domain 2A in ligand binding and ATP hydrolysis (Figure 1.8). Hence, the main objective of my Ph.D. is to understand the biochemical and functional importance of motif VI present in RecA-like domain2A (Figure 1.8).
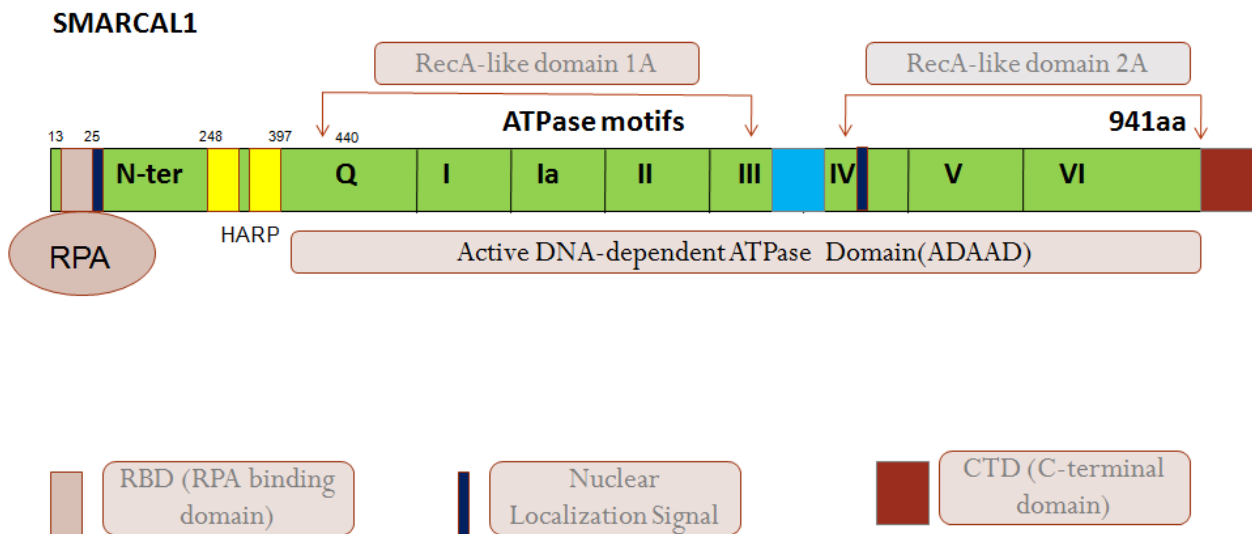
Figure 1.8: A Schematic representation of SMARCAL1 to understand its different domains.

Hypothesis and Objectives

A protein member of DEA(D/H)-box RNA helicase family, eIF4A plays role in the initiation of protein synthesis in eukaryotes. Specifically, it participates in melting of secondary structures present in mRNA that can hamper the translation initiation process. Due to the weak helicase activity of eIF4A, it works in cooperation with other helicases like eIF4E and eIF4G to perform its translation initiation activities. Still, the RNA binding and ATPase activity of eIF4A are necessary for the helicase activity (Caruthers et al., 2000). Mutational studies conducted in eIF4A have shown the importance of motif VI in RNA binding and ATP hydrolysis (Pause et al., 1993; Pause and Sonenberg, 1992). The crystal structure of eIF4A also suggests the role of motif VI in ATP catalysis in the presence of ATP and DNA. The importance of amino acid residues of motif VI in forming a salt bridge with amino acids present in motif II too has been observed in crystal structure of eIF4A (Caruthers et al., 2000). The crystal structure of the SWI/SNF protein family member, Rad54, has shown the importance of motif VI in mediating inter-lobe communication and ATP hydrolysis (Thomä et al., 2005). Therefore, my main objective is to characterize the role of motif VI in ADAAD. As I could not purify the full length SMARCAL1, I used ADAAD as an experimental system for my biochemical studies. ADAAD is a N-terminal truncated proteolytic fragment of bovine homologue of SMARCAL1 which contains all the seven helicase motifs present in proteins of helicase superfamily 1 and 2. Further, I will also be studying the role of this motif in inter-domain and intra-domain interactions.

SMARCAL1 has been shown to be important for transcriptional regulation of c-*myc* (Sharma et al., 2015). Therefore, my second objective is to understand the mechanism for this transcriptional regulation by SMARCAL1, using ADAAD as an experimental system. Mutations in

SMARCAL1 causes a multisystem disorder SIOD (Boerkoel et al., 2002). One of the mutations, R820H, lies in motif VI. The biochemical role and the functional consequence of the motif VI associated mutation present in SIOD patients have not been investigated yet. Therefore, my final objective is to study how the mutation in motif VI impacts ligand binding, ATP hydrolysis, and thereby the function of the protein.

Based on this information, I have designed the following objectives for my study:

- To investigate the role of motif VI of ADAAD in ATP and DNA binding and how it is coupled to the activation of ATPase function of this protein.

- To study the inter-domain and intra-domain interactions, which can together induce a conformational change in ADAAD leading to DNA binding and ATP hydrolysis.

- To study the role of SMARCAL1 in transcription.

- To investigate the role of R820H mutation that has been identified in SIOD patients in DNA damage and repair.

2: Materials and Methods

Chemicals

Chemicals used for biochemical studies were of analytical grade and were purchased from Qiagen (India), G-biosciences (USA), SRL (India), Merck (India), Millipore Sigma (USA), or GE Healthcare (USA). Restriction enzymes were purchased from New England Biolabs (USA), and Merck (India). Gel extraction kit and plasmid extraction kit were purchased from Qiagen (USA) or Thermo Scientific (USA). Bradford dye for protein estimation was purchased from Millipore Sigma (USA). Protein and DNA molecular weight marker were purchased from New England Biolabs (USA). Ni^{2+}-NTA resin was obtained from EMD Millipore (USA). DEAE-sepharose beads were purchased from GE Healthcare (USA).

Chemicals used for mammalian HeLa cell culture like DMEM media, penicillin-streptomycin-amphotericin B antibiotic cocktail, sodium bicarbonate, doxorubicin sodium chloride, protease inhibitor cock tail, luminol and *p*-Coumaric acid were purchased from Millipore Sigma (USA). All these chemicals were of culture grade. Fetal Bovine Serum (FBS) was obtained from GIBCO (Life technologies, USA).

X-ray films, developer and fixer were obtained from Kodak, USA. TRIzol reagent was obtained from Thermo Scientific (USA), chloroform and isopropanol for RNA preparation were purchased from Qualigens (India). RT-PCR kit was purchased from Applied Biosciences (USA) or the individual products were obtained from Thermo Scientific, USA. SYBR green mix and optical PCR tubes for RT-PCR were obtained from either Applied Biosystems (USA) or KAPA Biosystems. Protein A/G beads were obtained from Merck (India). Transfection reagent, Turbofect was purchased from Thermo Scientific, USA.

Antibodies

Antibodies used for western blotting and co-immunoprecipitation experiment include BRG1 (Millipore Sigma; Catalog #B8184), β-actin (Millipore Sigma #A1978), RPA (Cell Signaling Technology, Cat #2208) and RNAPII (Cell Signaling Technology; Catalog # 2629S). Polyclonal SMARCAL1 antibody was raised against N-terminal HARP domain by Merck (Catalog # 106014). HRP-conjugated anti-mouse IgG (Catalog # HPO5) and anti-rabbit IgG (Catalog # HPO3) antibodies were obtained from Merck (India).

Strains

The bacterial strains JM109 obtained from Merck (India) were used for cloning purpose. BL21 (DE3) cells purchased from EMD Millipore were used for expression studies. Other *E. coli* strains like BL21 like codon plus, Rosetta and Gold, used for ADAAD crystallization, were a kind gift from Dr. Ashok Patel (IIT Delhi).

Primers and oligonucleotides

The primers for creating site-directed mutants were designed using the primer design software primer3 and oligocalc (http://www.basic.northwestern.edu/biotools/ oligocalc.html). The primers as well as PAGE-purified oligonucleotides for ATPase and CD studies were synthesized by Millipore Sigma (USA). The list of primers used in this study along with the annealing temperatures used in the PCR is given in Table 2.1. The oligonucleotides used for

binding studies and ATPase assays are given in Table 2.2. The list of primers used for RT-PCR is shown in Table 2.3.

Table 2.1: List of primers used for making the site-directed mutants. P1 and P2 are the primers used in amplification of the plasmid for site-directed mutagenesis. FP is the forward primer and RP is the reverse primer. 'X°C' is the annealing temperature for these primers.

ADAAD mutation	Primer Sequences	Annealing Temperature ('X' °C)
D591H P1	5'-GCTGAGCATCGGGTGCACCGCATC-3'	50
D591H P2	5'-GATGCGGTGCACCCGATGCTCAGC-3'	
R592A P1	5'-GCTGAGGACGCGGTGCACCGCATC -3'	50
R592A P2	5'-GATGCGGTGCACCGCGTCCTCAGC-3'	
H594A P1	5'-GACCGGGTGGCACGCATCGGACAA-3'	50
H594A P2	5'-TTGTCCGATGCGTGCCACCCGGTC-3'	
R595K P1	5'- GACCGGGTGCACAAGATCGGACAA-3'	50
R595K P2	5'- TTGTCCGATCTTGTGCACCCGGTC -3'	
R595H P1	5'-CACCATATCGGACAATTGAG-3'	50
R595H P2	5'-CTCAATTGTCCGATATGGTG-3'	
R595A P1	5'-GACCGGGTGCACGCCATCGGACAA-3'	50
R595A P2	5'-TTGTCCGATGGCGTGCACCCGGTC-3'	

Table 2.2: Oligonucleotide used for binding studies, circular dichroism and ATPase assay.

Stem-loop DNA (Oligonucleotide)	5'-GCGCAATTGCGCTCGACGATTTTTTAGCGCAATTGCGC-3'

Table 2.3: List of primers used for performing RT-PCR with overexpressed samples of full-length SMARCAL1 and SMARCAL1 mutants K464A and R820H.

Gene amplicon	Primer Sequences	Annealing temperature ('X' °C)
GAPDH P1	5'-CGGAGTCAACGGATTTGGTCGTAT-3'	60
GAPDH P2	5'-GGAACATGTAAACCATGTAGTTGAGG-3'	
BRG1 P1	5'-GGTACCGTGGTTATTGATGGCTGGTAAG-3'	60
BRG1 P2	5'-AGATCTATGGCGTGACACTTGCTAACAT-3'	
SMARCAL1 P1	5'-GGTACCGGTGTGTCCTGCCCGCTGCTTC -3'	60
SMARCAL1 P2	5'-AGATCTCTTACCAGCCATCAATAACCAC -3'	

Construction of site-directed mutants

Site-directed mutants were made by PCR amplification using the mutagenic primers listed in Table 2.1. pCP101, which is a pET-14b plasmid containing the gene encoding for ADAAD, was used as a template. To create R820H mutation in *SMARCAL1*, GFP-*SMARCAL1* in pcDNA 3.1 background was used (Sharma 2016). The PCR reaction mixture contained 0.2 mM dNTPs, 200

51

μM each of the primer, 1.5 mM MgSO$_4$, 1X high fidelity buffer and 1 unit of high fidelity PCR enzyme mix (25 µl reaction). The PCR conditions for creating site-directed mutants are given in Table 2.4. After amplification, 2 µl of 10X buffer and 5 unit of DpnI (MBI Fermentas, USA) were added to 20 µl of the PCR. The reaction mixture was incubated at 37°C for 16 hr to digest the parental template DNA.

Table 2.4: PCR cycle for amplification of site-directed mutants. The annealing temperature (X°C) for various mutants is given in Table 2.1.

Temperature	Time duration	Number of cycles
95°C	5 min	x1
95°C	45sec	x16
X°C	45sec	
68°C	9.12 min	
72°C	10 min	x1

Transformation

Dpn1 digested mix (20µl) was added to 50 µl of JM109 competent cells and this mixture was gently tapped and incubated on ice for 30 min. After incubation, the mixture was given a heat shock treatment by incubating at 42°C for 90 sec. Subsequently, this mixture was quickly incubated on ice, and then 950 µl of sterilized LB media was added and incubated for 1 hr at

37°C. After incubation, the cells were centrifuged at 5000 rpm for 5 min. The excess media (900 µl) was removed and the rest (100 µl) was used to resuspend the pellet and then spread on LB agar containing ampicillin (100µg/ml). The plates were incubated at 37°C for 12-16 hr.

Screening of the Transformants

The transformants were verified by colony PCR. For performing colony PCR, individual colonies were picked and mixed in 20 µl water, followed by 5 min heating at 95°C for lysis. This mixture was incubated with 0.2 mM dNTP, 1X Taq DNA polymerase buffer, 0.4 µM each of forward and reverse gene specific primers and 1U of Taq DNA polymerase enzyme; and allowed for PCR reaction (Table 2.5). The PCR positive clones were selected and further confirmed by restriction digestion using EcoRI and XhoI. Finally, the site-directed mutation was confirmed by Sanger sequencing.

Table 2.5: Colony PCR to confirm the presence of insert.

Temperature	Time duration	Number of cycles
95°C	5 min	x1
95°C	1 min	x16
55°C	1.5 min	
72°C	1 min	
72°C	5 min	x1

Overexpression and purification of GST-tagged ADAAD

Overexpression and purification of GST-tagged ADAAD was done by transforming the appropriate plasmid into the expression host *E. coli* BL21 (DE3) competent cells. A 10 ml primary culture was grown by inoculating a transformant colony into sterile LB media having ampicillin (100µg/ml). The cultures were kept in an incubator shaker at 37°C for 12-16 hr overnight. Next day, 1% (v/v) of the primary inoculum was added to 500 ml sterile LB media containing 100µg/ml ampicillin. The cells were grown at 37°C till OD reached 0.6 and then shifted to 16°C and protein expression was induced by addition of 0.5 mM IPTG. The cells were grown for 16 hr and harvested by centrifuging at 6000 rpm for 5 min at 4ºC. The cell pellet was resuspended in lysis buffer containing 50 mM Tris-Cl pH 8.0, 150 mM NaCl, 150 mM MgCl$_2$, 0.1% (v/v) Triton X-114, 0.2 mg/ml lysozyme, 10 mM β-mercaptoethanol and 0.5 mM PMSF. The cell lysate was homogenized properly and then incubated on a rocking platform at 4ºC for 1 hr. After incubation, the cell lysate was sonicated for 10 cycles (15 sec ON, 45 sec OFF). The lysate was centrifuged at 12000 rpm for 30 min at 4°C. The supernatant was then transferred into a fresh flask and 60% (w/v) ammonium sulphate was added to precipitate the proteins. The precipitate was harvested by centrifugation at 10,000 rpm for 10 min at 4°C. The protein pellet was resuspended in the buffer containing 50 mM Tris-Cl pH 8.0, 2 M NaCl, 150 mM MgCl$_2$ and 5 mM β-mercaptoethanol and centrifuged at 40,000 rpm for 2 hrs at 4°C. After centrifugation the supernatant was dialyzed against the buffer containing 50 mM Tris-Cl pH 8.0, 150 mM NaCl, 150 mM MgCl$_2$ and 5 mM β-mercaptoethanol till the conductivity of the supernatant was the same as that of the dialysis buffer. After centrifugation

the supernatant was passed 3-4 times through the column containing 1 ml glutathione beads pre-equillibrated with dialysis buffer. After washing, prescission protease was added to the column for the cleavage of the GST-tagged ADAAD protein from the glutathione beads and incubated for 5-6 hr with the beads and cleavage buffer at 4°C. The flow through obtained was further passed through the DEAE-sepharose beads to remove the GroEL contamination. 10% SDS-PAGE was run to check the purity of the protein.

Overexpression and purification of His-tagged ADAAD and site-directed mutants

The appropriate plasmid was transformed into *E. coli* BL21 (DE3) competent cells. One transformed colony was picked and inoculated in 10 ml LB media containing ampicillin (100µg/ml). Next day, 1% (v/v) primary culture was inoculated into fresh LB media and allowed to grow at 37°C till the OD reached 0.6. The cells were induced with 0.5mM IPTG and grown overnight at 16°C. The cells were harvested after spinning at 6000 rpm for 5 min and resuspended in lysis buffer (10 ml lysis buffer was added per 1 gm of cell pellet) containing 50 mM Tris-Cl pH 8.0, 150 mM NaCl, 150 mM MgCl$_2$, 0.1% (v/v) Triton X-114, 0.2 mg/ml lysozyme, 10 mM β-mercaptoethanol and 0.5 mM PMSF. The cell suspension was incubated at 4°C for 1 hr and the cells were lysed using sonicator (15sec on and 45 sec off pulses; 6 times). The supernatant and the pellet were separated by centrifuging the cells at 9500 rpm at 4°C for 45 min. The pellet was discarded and the supernatant was passed through 1 ml of Ni^{+2}-NTA beads pre-equilibrated with the dialysis buffer. The column was then washed with ~25 ml wash buffer (containing 30 mM imidazole) before elution. The protein was eluted using buffer containing 250 mM imidazole, 50 mM Tris-Cl pH 8.0, 100 mM NaCl and 5 mM β-mercaptoethanol. After

elution the fraction were analyzed by 10% SDS-PAGE. The protein was dialyzed against dialysis buffer (20 mM Tris-SO$_4$ pH 7.4, 1 mM EDTA, 50 mM K$_2$SO$_4$ and 20% glycerol).

DEAE ion-exchange chromatography

The DEAE-sepharose beads (1 ml) were packed into 1 ml plastic column and then equilibrated with the dialysis buffer. The dialyzed protein was then passed though the DEAE-sepharose column. The fractions were collected and analyzed by 10 % SDS-PAGE.

Protein estimation

Protein estimation was done using Bradford assay in a 96 well microtiter plate. The color was monitored at 590 nm using Spectromax microplate reader (MTX Lab Systems, Inc**,** USA).

ATPase Assay

NADH coupled oxidation assay was used to monitor the ATPase activity. Typically, 1.4 µg of protein in 1X REG Buffer containing 25 mM Tris-OAc pH 7.5, 6 mM MgOAc, 60 mM KOAc, 1 mM MgSO$_4$, 5 mM β-mercaptoethanol, 1.4 mg/mL phosphoenolpyruvate, and 10 U pyruvate kinase and lactate dehydrogenase along with 2 mM ATP, 0.1mg/ml NADH and 10 nM stem-loop DNA in a 250 µl reaction volume. The reaction was incubated for 15 min at 37ºC and the absorbance of NADH was measured at 340 nm using microplate spectrophotometer (MTX Lab Systems, Inc**,** USA) (Figure 2.1)
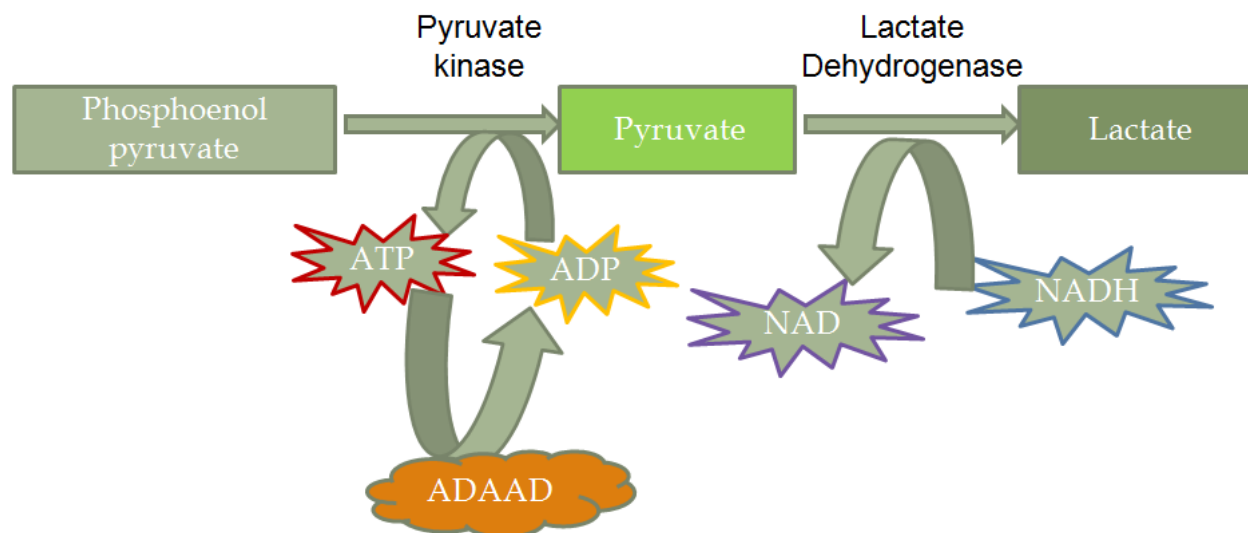
Figure 2.1: Diagrammatic representation of NADH-coupled oxidation assay to measure the ATP hydrolysis of ADAAD.

Fluorescence studies

Steady-state fluorescence titrations were assayed using Cary Eclipse Spectrofluorometer, USA. Excitation and emission slit width was kept at 5 nm and 10 nm, respectively. The protein was excited at 295 nm and the emission was monitored at 340 nm. All the experiments were executed at room temperature. Titrations of ligands were done in buffer containing 50 mM Tris-SO$_4$ pH 7.5, 5 mM β-mercaptoethanol and 1 mM MgSO$_4$. The binding data was fit to one-site saturation for the ligand and protein interaction. The following equation was used to calculate the dissociation constant:

$(K_d).\Delta F/F_o = B_{max}[L]/(K_d + [L])$

Where ΔF is $F_c - F_{ini}$

As shown in earlier studies, there is no detectable ATP hydrolysis in the absence of REG buffer and at 25^0C suggesting that the binding and conformational changes measured in our experiments are only because of the protein-ligand binding and there is no detectable enzymatic activity during these experiments. Here F_c is fluorescence intensity calculated after correcting dilution, F_{ini} is the initial fluorescence. B_{max} is the maximal binding, [L] is the ligand concentration, and K_d is the dissociation constant. All the fluorescence spectra were corrected for inner filter effect and dilutions (Figure 2.2)
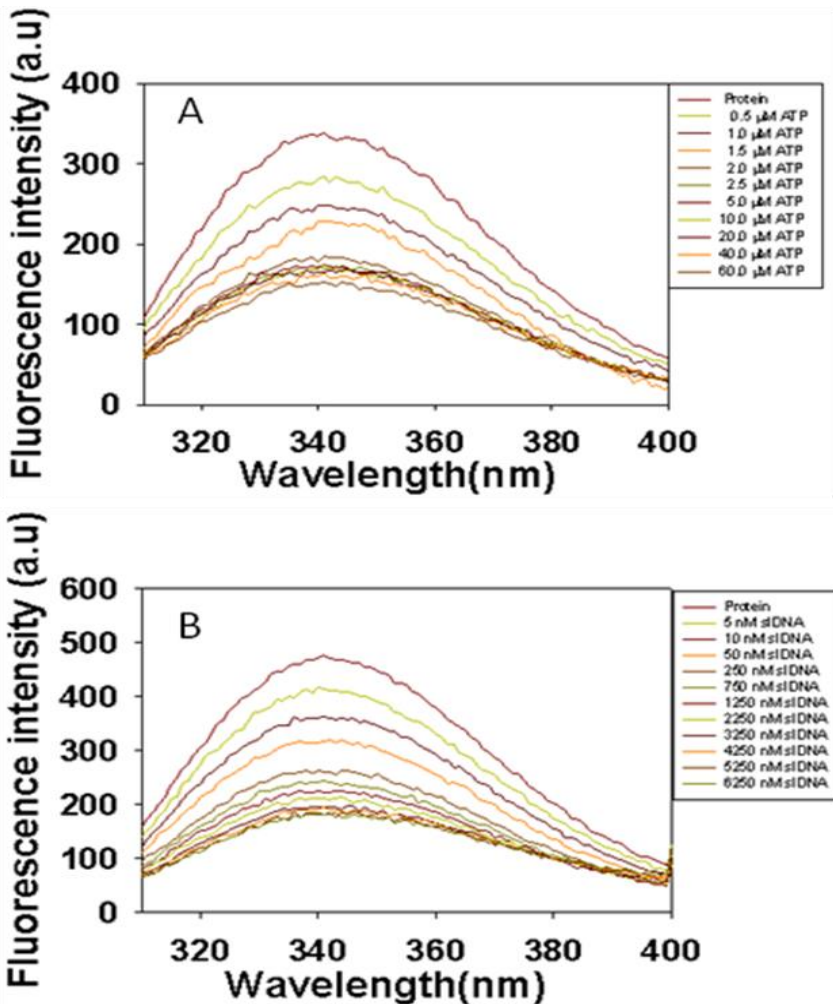
Figure 2.2: Representative fluorescence spectra showing fluorescence quenching of wild-type ADAAD in presence of (A) ATP and (B) stem-loop DNA.

Circular Dichroism (CD) Spectroscopy

Far UV CD spectra were obtained by Chirascan (Applied Photophysics) using a 1 mm cuvette. The reaction contained 0.1mg/ml protein, reaction buffer (20 mM Tris-Cl pH 7.5, 1 mM EDTA, 100 mM NaCl, 1 mM $MgCl_2$, and 1 mM DTT), 20 µM ATP and 2 µM DNA. The spectra recorded were corrected for dilution, buffer and ligand added. The CD scan rate was 1.425 s/nm. Three accumulations were measured for each scan. The secondary structure for each protein was calculated using Dichroweb software (http://dichroweb.cryst.bbk.ac.uk/html/home.shtml). The conformation of $G_E C_E$ DNA in the absence and presence of ADAAD and ATP was also monitored by CD spectroscopy. The spectra of 500 nM DNA in 1 mM sodium phosphate buffer (pH 7.0) were obtained in the range of 200 to 300 nm at 0.5 nm intervals at 37°C. The DNA was then incubated with 1 µM ADAAD, 2 mM ATP and 10 mM Mg^{+2} to check the conformational change in the presence of the protein. To assess the importance of ATP hydrolysis in the conformational change, 50 mM EDTA was added to the reaction. For each reaction, 5 scans were performed and for each condition scans were subtracted from the buffer reading. The CD values in both cases are reported in terms of mean residue ellipticity (θ), and was calculated by the formula $[\theta] = (S \times mRw)/(10cl)$, where S is the CD signal in millidegrees, mRw is the mean residue mass, c is the concentration of the protein in moles/liter and l is the cell length. I have measured the Circular Dichroism spectra both in the absence of ligands and in the presence of DNA and ATP. Both DNA and ATP are CD active only at higher concentrations. ATP is CD active in millimolar range and DNA is CD active in micromolar range. To correct the errors even

by low concentrations of ATP and DNA, I have subtracted the buffer +DNA+ ATP reading from Protein +DNA+ ATP reading.

Cell culture

HeLa cells obtained from NCCS, Pune were used for all the experiments. HeLa cells were grown in Dulbecco's Modified Eagle Medium (DMEM) media which were prepared by dissolving DMEM powder in 1L of autoclaved distilled water. Sodium bicarbonate (3.74 g) was added to the mixture and pH was adjusted to 7.4. The media was filtered and supplemented with 10% Fetal Bovine Serum (FBS) and 1% penicillin-streptomycin-amphotericin B (PSA) antibiotic cocktail.

Serum Starvation Experiment

HeLa cells grown with 10% FBS were used as unstarved cells. For serum starvation, the cells were grown in the presence of 0.4% FBS for 48 hr. These cells were released from serum starvation by adding 10% FBS again and collected after 2 hr.

Transfections and Immunofluorescence

HeLa cells were seeded on a coverslip in 35 mm culture dish. Wild-type SMARCAL1, Zeo-GFP control and mutant (K464A, R820H) plasmids were transfected into HeLa cells using either lipofectamine or turbofect. Transfected cells were grown at $37^{\circ}C$ for 36 hr and taken out for further processing. The cells were washed 3-4 times with 1X PBS, fixed with 1:1 methanol,

acetone mixture for 10 min and permeabilized with 0.5% (v/v) Triton X-100. After

permeabilization, the cells were washed with 1X PBS followed by blocking with 2% BSA at

37°C for 1 hr. The cells were again washed with 1X PBS and incubated with required amount of

primary antibody in 2% BSA at 37°C for 1 hr. The cells were then washed with 1X PBS 3-4

times, incubated with a mixture of TRITC or FITC-conjugated secondary antibodies and

DAPI/Hoechst 33342 at a dilution of 1:1000 in 2% BSA at 37°C for 1 hr. The cells were washed

with 1X PBS and the coverslips containing cells were mounted on slides and observed in

confocal microscope (Olympus) under 60X oil immersion objective.

Protein occupancy assay

Protein occupancy assay was executed using a method described earlier (Infante et al., 2012). In

this assay, 2×10^7 to 3×10^7 HeLa cells were cross-linked with 37% (v/v) formaldehyde (1%

final concentration) for 15 min at room temperature. The cross-linking reaction was quenched by

adding 2.5 M glycine (125 mM final concentration). The cells were washed with 10 ml of ice-

cold PBS twice at 4°C and scraped out from the culture plate with the help of cold and clean cell

scraper. The cells were resuspended in NPS buffer (0.5 mM spermidine, 0.075% (v/v) NP-40, 50

mM NaCl, 10 mM Tris-Cl pH 7.5, 5 mM MgCl$_2$, 1 mM CaCl$_2$ and 1 mM β-mercaptoethanol).

The cells were divided into six 300 µl aliquots and were digested with 0, 1, 2.5, 10, 20 and 30 U

of MNase for 15 min at 37°C. Digestions were stopped by shifting the tubes to 4°C and adding

18 µl of 500 mM EDTA and 7 µl of 200 mM EGTA. The RNA and bound proteins were

degraded by treating the samples with 60 µl 10% SDS, 10 mg/ml proteinase K and then

incubating the cells with 10 µl of 10 mg/ml RNAse for 15 min at 37°C. The DNA was extracted

twice with phenol (saturated with 0.1 M Tris-Cl pH 7.5) and once with equal volume of

chloroform. The extracted DNA was precipitated with 0.1 volumes of 3 M NaOAc, pH5.3, and

2.5 volume 100% ice-cold ethanol and resuspended in TE buffer. The DNA was analyzed on a

1% agarose gel and the mononucleosome-sized DNA or protein-bound DNA with size less than

200 bp, was purified from the gel. After purification, the purified DNA was subjected to

quantitative real time RT-PCR. For each c-*myc* primer set, amplification of template was

analyzed and compared between MNase digested and undigested samples. Figure 2.2 shows the

graphical representation of the assay.
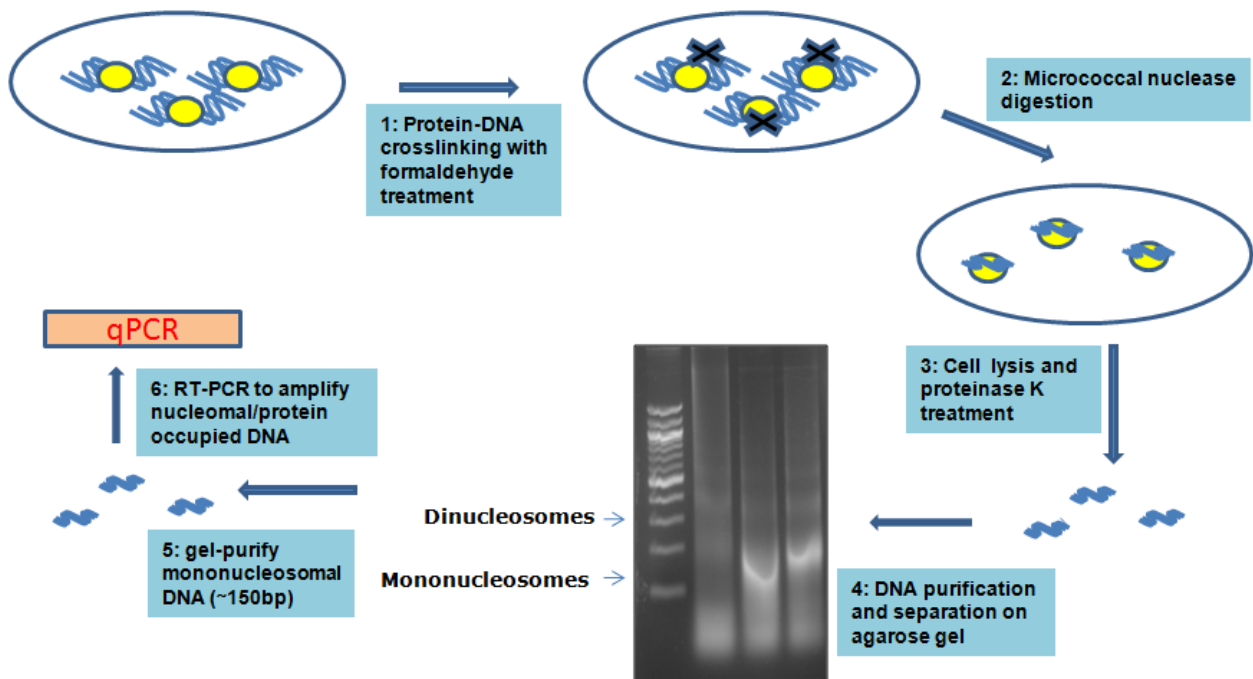
Figure 2.3: Methodology adopted for protein occupancy assay.

3: Characterization of the role of motif VI of ADAAD in ligand binding and ATP hydrolysis

INTRODUCTION

As described in the previous chapter, helicases are ubiquitous enzymes that harness the energy derived from ATP hydrolysis to unwind DNA. The crystal structure of Rep and PcrA have shown the location of all conserved motifs that are present dispersed throughout the 200 to 700 amino acids of helicase domain in the primary amino acid sequence (Subramanya et al., 1996; Velankar et al., 1999).

The helicases, as described in the previous chapter, are divided into many subfamilies. Of these, SF1 and SF2 proteins possess two domains: RecA-like domain 1 and RecA-like domain 2, which are further divided into two subdomains-RecA-like subdomains 1A and 1B; and RecA-like subdomains 2A and 2B (Korolev et al., 1998). Out of all four subdomains of helicase proteins, the conserved motifs occupy the base positions at RecA subdomain 1A and 2A and form the nucleotide interaction site (Dürr et al., 2005; Korolev et al., 1998). The amino acid residues present in the conserved motifs of helicase family are necessary for the interaction with ATP and DNA as well as for ATP hydrolysis and helicase activity (Korolev et al., 1998). However, divergent amino acid regions present in between these motifs provide the specific function to each of the protein (Korolev et al., 1998). In this section, I will highlight the function of different motifs of SF1 and SF2 helicases.

Motif Q is present 15-20 amino acid residues upstream of motif I. The importance of motif Q has been reported in SF1 helicases like UvrD, Uvr B, PcrA and Rep (Ahmad et al., 2012; Korolev et al., 1998; Sinha et al., 2009; Theis et al., 1999; Velankar et al., 1999). In SF2 helicases like yeast protein eIF4A (translation initiation factor and a RNA helicase) and viral helicase NS3, this motif has been shown to be essential in ATP binding and ATP hydrolysis (Gallivan and

McGarvey, 2003; Tanner et al., 2003). Recently, the motif Q has also shown to play a role in adenine recognition and thereby, ATP hydrolysis (Nongkhlaw et al., 2012; Tanner, 2003). Biochemical studies performed with ADAAD has shown the role of motif Q in ATP hydrolysis but not in ATP and DNA binding (Nongkhlaw et al., 2012).

Motif I contains the conserved consensus sequence AXXGXGKT, of which lysine is known to interact with magnesium ion when NTP-Mg^{+2} complex binds to the protein (Velankar et al., 1999). This lysine interacts with β-phosphate of ATP and stabilizes the transition state necessary for ATP catalysis (Velankar et al., 1999). The function of motif I has prominently been demonstrated in nucleic acid unwinding and ATP hydrolysis in SF1 proteins like UvrD, Upf1p and UL5 wherein mutation of the conserved lysine to alanine results in ATPase-deficient proteins with abrogated helicase activity (George et al., 1994; Graves-Woodward and Weller, 1996; Weng et al., 1996). The role of motif I in ATPase activity and nucleic acid unwinding has been found conserved in both SF1 helicases and SF2 helicase proteins (Gross and Shuman, 1996; Heilek and Peterson, 1997; Kim et al., 1998; Pause and Sonenberg, 1992; Sung et al., 1988). However, in ADAAD and other SWI2/SNF2 proteins motif I is shown to be important only for ATP hydrolysis and not for ATP binding (Nongkhlaw et al., 2012).

Motif Ia: Specific mutations introduced in the motif Ia (PTRELA) of UL9 of HSV-1 has shown its involvement in ssDNA binding (Marintcheva and Weller, 2003). The mutant proteins exhibited no change in intrinsic ATP hydrolysis but a substantial change in ssDNA-stimulated ATPase activity (Marintcheva and Weller, 2003). These results have shown the essentiality of motif Ia in viral growth as it contributes to ssDNA binding and helicase activity of the protein (Marintcheva and Weller, 2003). The role of motif Ia in ATP-dependent chromatin remodeling

proteins has not been completely elucidated though mutation of a conserved tryptophan in this motif of ADAAD results in abrogated ATPase activity (Gupta 2015). I will describe the role of this motif in detail in later chapter.

Motif II, also called as Walker B box, is a highly specific form of the ATPase B motif that is crucial for the ATP hydrolysis activity of helicases. The consensus sequence is DEAD and therefore, proteins possessing this sequence are also called as DEAD box proteins (Gorbalenya and Koonin, 1993; Tanner and Linder, 2001; Wassarman and Steitz, 1991). In SF1 helicases, the first aspartate and glutamate residue of this motif interacts with NTP-associated magnesium ion resulting in water molecule activation and hence, ATP hydrolysis (Story and Steitz, 1992). Mutations introduced in aspartate of DEAD sequence lead to defect in ATPase as well as helicase activity, signifying its role in coupling ATP hydrolysis and DNA unwinding (Brosh and Matson, 1995; Weng et al., 1996). With the increasing number of discovery of helicases, motif II of SF2 helicases has been further divided into DEAD, DExH, DEAH and DEAH* subgroups, where DEAH* is the subgroup with highest sequence variation in the conserved motifs (Figure 3a) (Tanner and Linder, 2001). This motif has also been seen to be important in coupling ATPase activity and helicase activity of eIF4A, the best studied member of DEAD protein family of RNA helicases of SF2 family (Caruthers et al., 2000; Pause and Sonenberg, 1992; Rogers et al., 2002).

Motif III is essential for maintaining helicase activity of SF1 and SF2 helicase superfamily of proteins (Brosh and Matson, 1995; Graves-Woodward and Weller, 1996; Pause and Sonenberg, 1992). In SF1 helicases, motif III is shown to interact directly with the nucleotide bases of nucleic acids by hydrogen bonding and base-stacking interactions (Velankar et al., 1999). Like

motif Ia, motif III also interacts with the loop region of aspartate and glutamate residues present in motif II, thereby, translating the energy released from ATP to DNA molecule (Korolev et al., 1997). The motif III residues of Rep protein of SF1 helicase superfamily make direct contact with single-stranded DNA molecule (Korolev et al., 1997). Surprisingly, mutations introduced in motif III in SF1 helicase proteins like UL5 of HSV-1 results in abrogation of helicase activity without affecting the ATPase activity of the protein (Graves-Woodward et al., 1997). The association of motif III with the gamma phosphate of ATP has been shown in PcrA-ADP-DNA complex (Velankar et al., 1999). In SF2 helicase like NS3, the motif III forms a loop that connects domain 1 and 2 and transduces the conformational changes generated by ATP binding resulting in hydrolysis (Kim et al., 1998). The energy released after ATP hydrolysis is then utilized in DNA binding and translocation (Kim et al., 1998). These results signify that motif III is involved in coupling energy released from DNA-dependent ATP hydrolysis to double-strand DNA unwinding and DNA translocation.

Motif IV in SF1 helicases has been reported to be important for NTP binding. The conserved amino acid, arginine, of motif IV in PcrA has been shown to play a role in ATP binding (Velankar et al., 1999). Mutation in this conserved arginine to alanine results in altered K_m for ATP interaction (Velankar et al., 1999). In addition to this arginine, a conserved phenylalanine present in this motif has also been shown to be required for ATP-dependent binding to RNA substrates in DEAD-box proteins (Banroques et al., 2008). This has further been supported by crystal structure of Rep protein, where tyrosine of motif IV has been shown to make contacts with the adenine ring of ATP (Korolev et al., 1997). All these findings suggest an involvement of motif IV in ATP binding and hence, ATP hydrolysis.

Motif V of SF1 and SF2 helicases share sequence similarity and occupy a similar position in the 3D structures of helicases. Mutational analysis of UvrB, UL5 and NS3 has revealed that this motif interacts with the sugar-phosphate backbone of nucleic acid (Graves-Woodward and Weller, 1996; Kim et al., 1998; Moolenaar et al., 1994). In contrast, in few RNA helicases mutations in this motif did not affect DNA binding but resulted in loss of ATPase and helicase activity (Fernández et al., 1997). In conclusion, motif V appears to participate in both DNA binding and ATP hydrolysis.

Motif VI with consensus sequence HRIGRXXR has been proposed to be present at ATP-binding cleft (Thomä et al., 2005). Several studies have shown a defective DNA binding when the conserved amino acid residues of motif VI were altered. Mutational analysis performed with *E. coli* UvrD, a DNA helicase and member of SF1 family, has shown that this motif is required for single-strand DNA binding, ATP catalysis and ligand-induced conformational alterations (Hall et al., 1998). The crystal structure of Rep protein has shown the association of motif VI with motif IV as well as with motif III (Korolev et al., 1997). These interactions help in transducing the conformational changes and energy released after NTP binding to nucleic acid interaction (Korolev et al., 1997). Crystal structure of PcrA has shown that as result of ATP-induced conformational changes, the arginine residue of motif VI interacts with gamma phosphate of ATP molecule (Velankar et al., 1999). Also due to its close proximity to ATP and DNA-binding site, the motif VI mediates a ligand-induced conformational change in the protein leading to the movement of the protein along the DNA substrate (Graves-Woodward et al., 1997; Weng et al., 1996). Mutations in motif VI of Upf1p results in a defect in RNA binding, ATP hydrolysis and RNA unwinding (Weng et al., 1996). The crystal structure of Rep protein shows that motif VI is

important for inducing the conformational changes in the DNA binding sites and thus, is responsible for ATP hydrolysis (Korolev et al., 1997).

In SF2 helicases, conflicting roles have been reported for motif VI. Several enzymes show defect in nucleic acid binding after mutation/removal of motif VI while in DEAD box RNA helicase eIF4A, mutation in only one-out of three arginine of conserved motif VI results in nucleic acid binding defect (Fernández et al., 1995; Li and Broyles, 1995; Moolenaar et al., 1994; Pause and Sonenberg, 1992). Conversely, crystal structure of NS3 helicases has shown that motif VI arginine is not able to contact DNA molecule; instead arginine makes contacts with phosphate group and transduces the ATP-released energy to DNA unwinding (Yao et al., 1997).

All these studies signify the communication of motif VI with NTP and nucleic acid that ultimately leads to conformational change in the protein allowing for it to translocate along the DNA molecule.
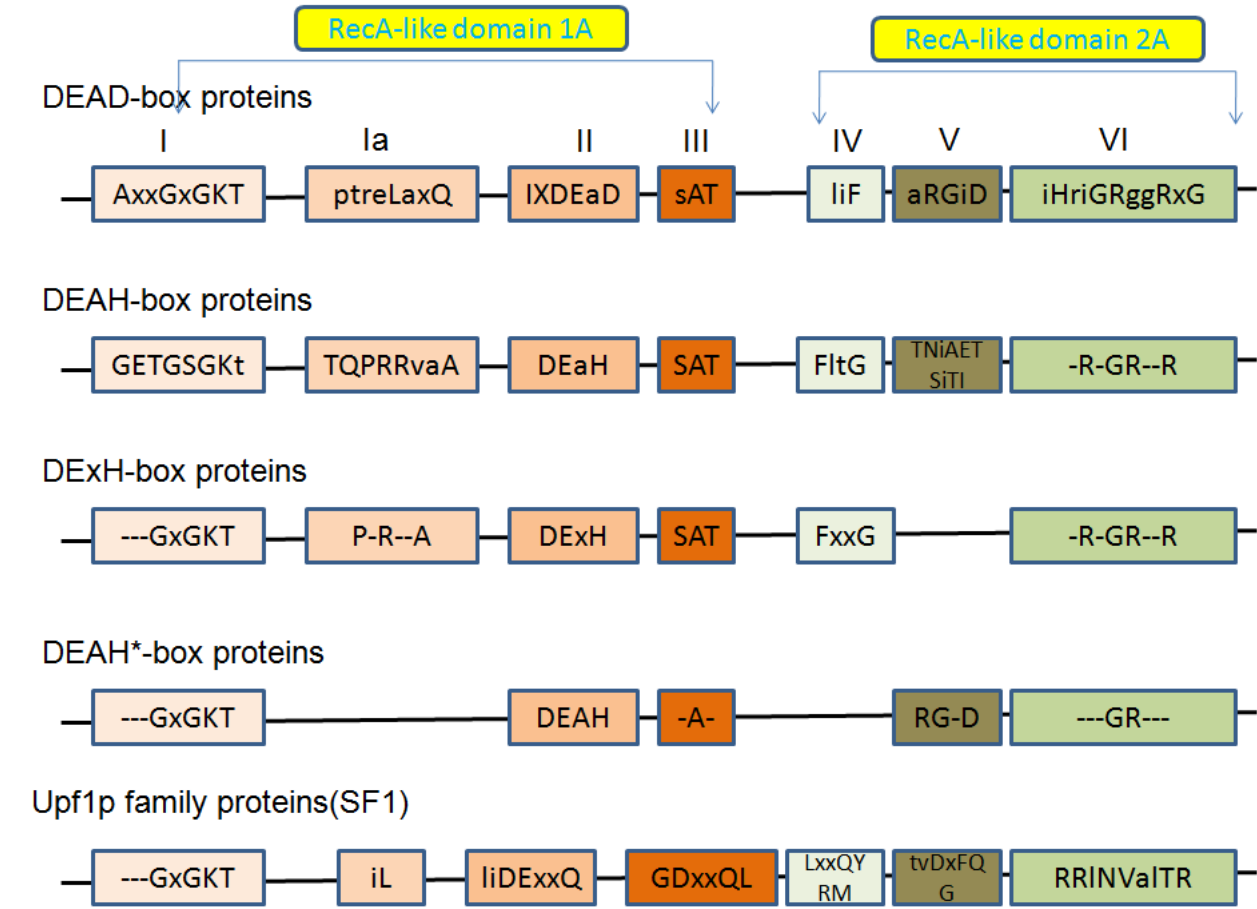
Figure 3a: Modular representation of SF1 helicases and DEAD-box containing SF2 helicases, depicting conserved motifs (I, Ia, II, III, IV, V and VI) shown in colored boxes (Tanner and Linder, 2001; Wassarman and Steitz, 1991). Motif I, Ia, II and III are present in RecA-like domain 1A, while motif IV, V and VI are present in RecA-like domain 2A. The proteins belonging to DEAH* box shows the highest variations in the conserved motifs. The N-terminal and C-terminal extensions vary in their lengths in each of the family of helicases. The motif Q is not shown in this figure.

Hypothesis and objectives

ADAAD is a proteolytic fragment of bovine homologue of SMARCAL1, whose role in ATP binding and DNA binding has been well-characterized (Muthuswami et al., 2000; Nongkhlaw et al., 2009). The protein has been shown to maximally hydrolyze ATP in the presence of DNA molecules containing double-stranded to single-stranded transition regions (Muthuswami et al., 2000; Nongkhlaw et al., 2009). Although the roles of motifs Q and I of RecA-like subdomain 1A are known in ligand binding and ATP hydrolysis, we still lack information regarding the function of motifs present in RecA-like subdomain 2A (Figure 3b) (Nongkhlaw et al., 2012). For my Ph.D, I decided to focus on motif VI present in the RecA-like subdomain 2A. Motif VI, DRAHRIGQ, is a conserved motif present in SWI2/SNF2 proteins and RNA helicases. It possesses positively charged amino acid residues arginine and histidine, negatively charged amino acid residue aspartate and neutral amino acid residue alanine, isoleucine and glycine. As mentioned above, studies in eIF4A and other helicases have shown the significance of motif VI in RNA binding and ATP hydrolysis (Pause et al., 1993; Pause and Sonenberg, 1992). The positive charged arginine residue of eIF4A has been shown to possess an interaction with RNA. From these literature studies, I hypothesized the same electrostatic interaction between positive charged residues histidine and arginine of ADAAD and negative charged phosphate backbone of DNA and therefore I created the site directed mutations R592A, H594A, R595A where I replaced the positive charged with neutral amino acid residue alanine. Additionally, I created a site directed mutation R595K to check the importance of side chain of arginine in interaction with DNA. Therefore, as my first objective, I have investigated the role of motif VI of ADAAD

71

in ATP and DNA binding and how it is coupled to the activation of ATPase function of this protein.

Based on this hypothesis, I designed the following experiments:

- Bioinformatic analysis to identify the conserved residues of motif VI.

- Create and purify site-directed mutants-R592A, H594A, R595K and R595A.

- ATPase assays to determine the impact of these mutations on ATP hydrolysis.

- Analysis of ligand binding using fluorescence spectroscopy.

- Analysis of protein structure in the absence and presence of ligands using CD spectroscopy.

- Calculation of approximate values of alpha structure, beta structure and random coil in the wild-type and mutant proteins using dichroweb analysis.
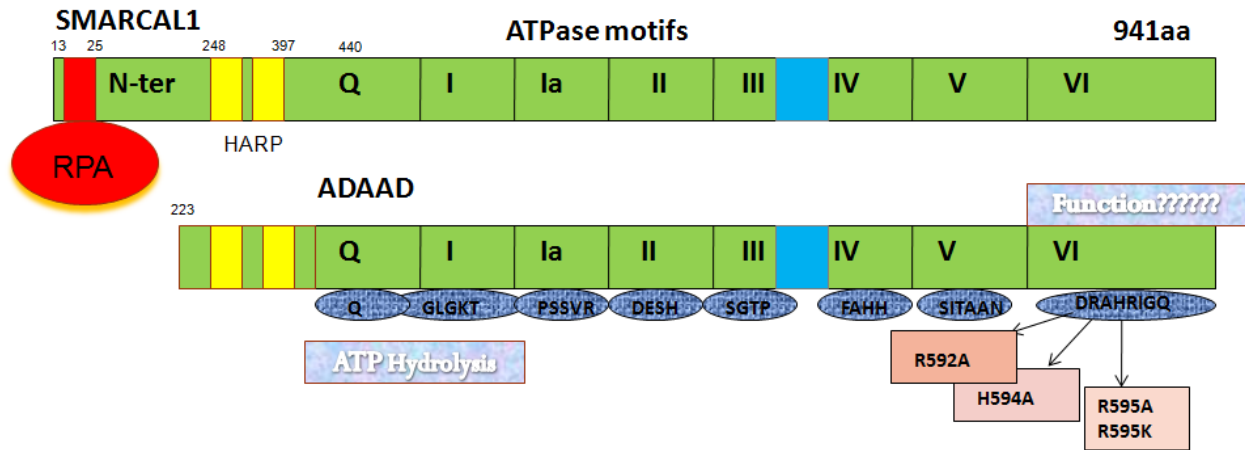
Figure: 3b: Schematic representation of SMARCAL1 and of ADAAD, the C-terminal proteolytic fragment of bovine homologue of SMARCAL1. The conserved motifs present in ADAAD as well as the role of motif Q and I in ATP hydrolysis has been highlighted here (Nongkhlaw et al., 2012). The conserved amino acid residues of motif VI selected for my study along with the mutations made are also shown.

RESULTS

The motif VI of SWI2/SNF2 proteins has the conserved DRAHRIGQ sequence.

I began my studies by performing multiple sequence alignment to identify the conserved residues of motif VI in SWI2/SNF2 proteins. Clustal W analysis showed that the motif VI in SWI2/SNF2 proteins consisted of DRAHRIGQ sequence (Figure 3.1a). In ADAAD, these residues were D591, R592, H594, and R595. Further, R820 in SMARCAL1, which corresponds to R595 of ADAAD, is mutated to histidine in SIOD patients.

I created site-directed mutants of the positively charged residues-R592, H594 and R595A. In each case, the residue was replaced with alanine, a neutral residue (Figure 3.1b). As R595 appeared critical and mutations in SMARCAL1 resulted in SIOD, therefore, I also mutated R595 to histidine and lysine creating R595H and R595K respectively. However, I will be describing the R595H results in my last chapter.
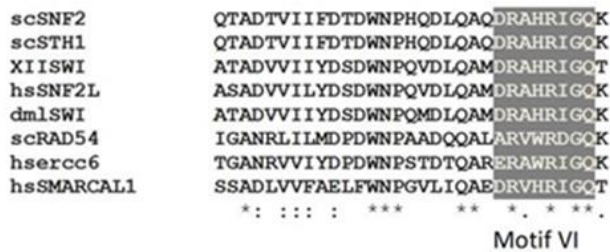
ADAAD	bSMARCAL1	hSMARCAL1
D591H	D814H	D816H
R592A	R815A	R817A
H594A	H817A	H819A
R595H	R818H	R820H
R595A	R818A	R820A
R595K	R818K	R820AK

(a)   (b)

Figure 3.1: (a) Clustal W analysis showing the conserved amino acids present in motif VI of SWI2/SNF2 proteins. ATP-dependent chromatin remodeling proteins from *Saccharomyces*

cerevisae, Xenopus itombwensis, Homo sapiens and *Drosophila melanogaster* were used for the

analysis. (b) Mutations introduced into Motif VI (DRAHRIGQ) are shown. The numbers of each

amino acid residue indicate the positions of ADAAD used for site-directed mutagenesis with

respect to human SMARCAL1, bovine SMARCAL1 and ADAAD.

The motif VI is crucial for the ATPase activity of ADAAD

The mutant proteins (R592A, H594A, R595A and R595K) were purified to homogeneity as

explained in Materials and Methods (Figure 3.2). The ATPase activity of these mutants was

studied and compared to the wild-type protein using NADH coupled oxidation assay (Figure

3.3a). R592A, R595A and R595K were not able to hydrolyze ATP suggesting that these residues

are essential for ATPase activity (Figure 3.3a).

H594A showed 20% ATPase activity as compared to the wild-type protein and therefore, I

calculated the kinetic parameters for this mutant. The K_M value was found to be 36.04 ± 4.54

nM, approximately thirty-fold higher than the wild-type ADAAD ($K_M = 0.88 \pm 0.22$nM),

suggesting that the residue might be playing a role in stem-loop DNA binding and ATP

hydrolysis (Figure 3.3b) (Table 3.3) (Gupta 2015). Moreover, V_{DNA} was also altered suggesting

that H594A that the rate of ATP hydrolysis was slower as compared to the wild-type protein

(Figure 3.3b). The catalytic efficiency, which is calculated as K_{cat}/K_M, was also found to be ~25-

fold lower in case of H594A. These results, therefore, indicate the importance of H594 residue in

ATP hydrolysis (Table 3.3).
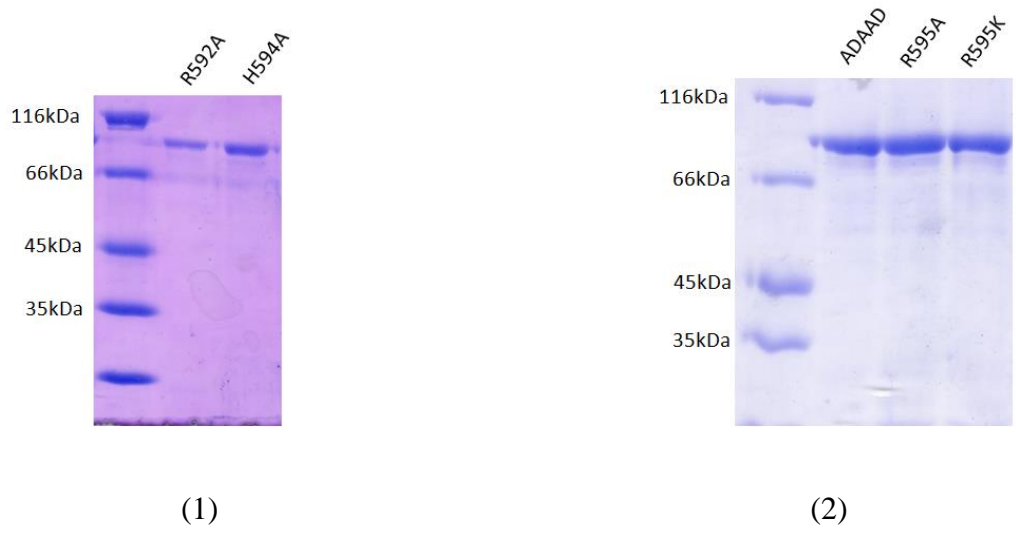
Figure 3.2: Coomassie Brilliant Blue staining showing the purified motif VI mutants: (1) R592A, H594A and, (2) R595A and R595K.
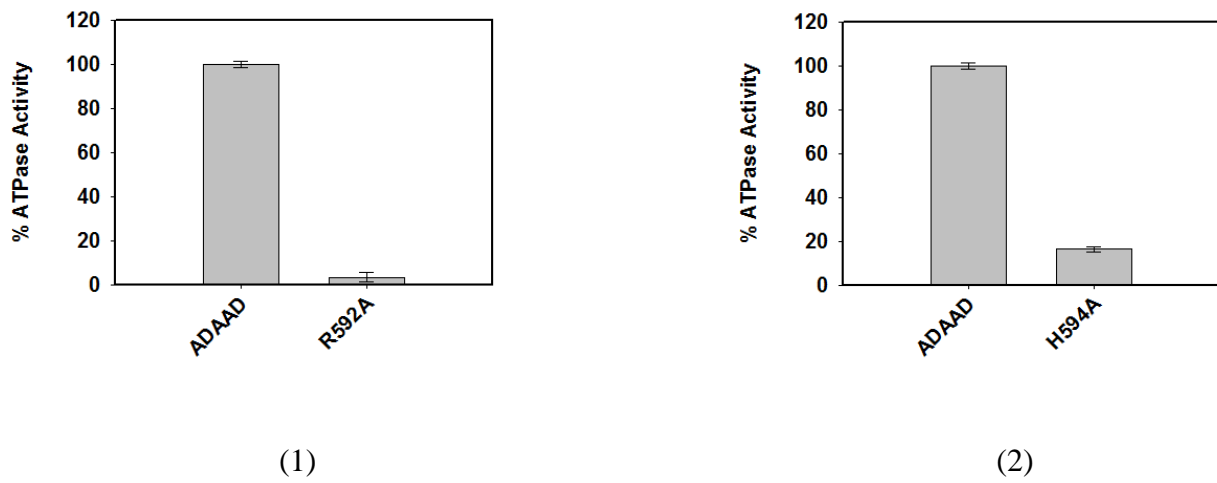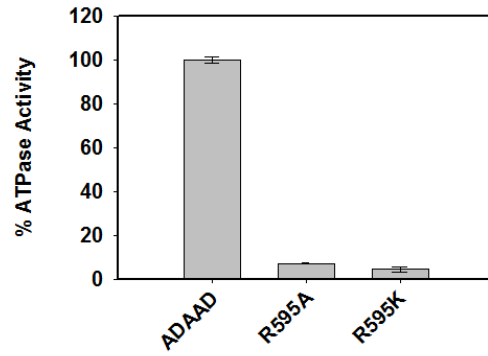
(3)

(a)
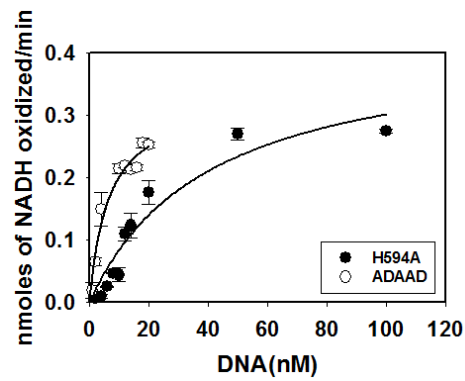
(b)

Figure 3.3 (a) Percentage ATPase activity was calculated for motif VI mutants in the presence of stem-loop DNA with respect to wild-type ADAAD. Each analysis represents the average of two independent experiments. (b) Calculation of Michaelis-Menten constant for wild-type ADAAD and H594A mutant. Each analysis represents the average of two independent experiments, whose standard deviations are shown in the figures. The protein concentration used in these analyses was 0.1 μM.

Arginine 592 is important for conformational integrity

The loss of ATPase activity in R592 could be either due to conformational change in the protein or due to impaired ligand binding. To investigate the reason for this defect in ATPase activity, the binding of ATP to the mutant protein in the absence and presence of stem-loop DNA was studied. The dissociation constant for the mutant protein, R592A-ATP interaction was similar to the wild-type ADAAD-ATP interaction both in the absence and presence of stem-loop DNA, suggesting that R592 residue does not play a role in ATP binding (Figure 3.4 a and b) (Table 3.1). Next, the interaction of the stem-loop DNA with the mutant protein was studied in the absence and presence of ATP. The dissociation constant for R592A-stem-loop DNA binding in the absence and presence of ATP was also approximately the same as wild-type ADAAD-stem-loop DNA binding (Figure 3.4c and d) (Table 3.1).

Therefore, the loss in ATPase activity appeared to be due to a change in the conformation of the protein, and so, I studied the conformation of the mutant protein in the absence and presence of ligands using CD spectroscopy. The conformation of the protein was found to be significantly altered as compared to wild-type ADAAD protein both in the absence and presence of ligands (Figure 3.5a and 3.5 b). Analysis using dichroweb software showed that R592A has a very high alpha helical content of protein both in the absence and presence of ligands as compared to the wild-type protein (Table 3.2).
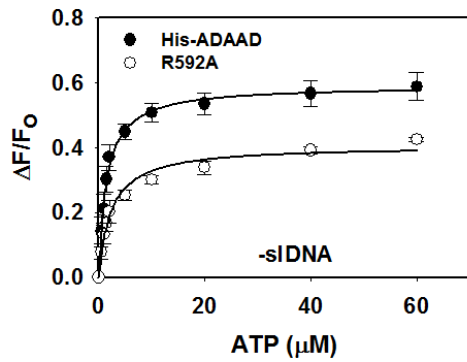
(a)
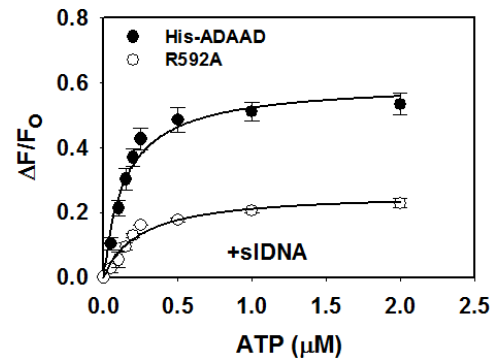
(b)
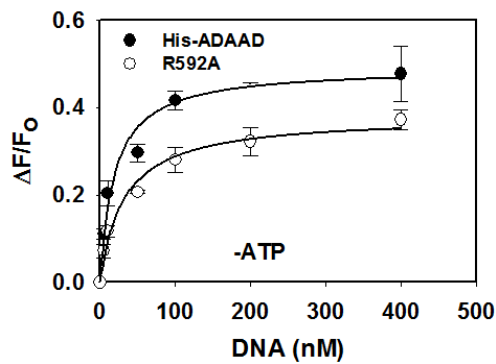
(c)
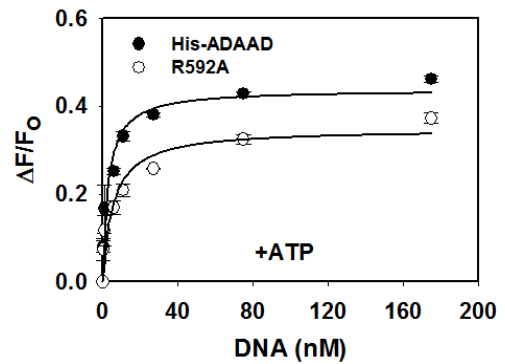
(d)

Figure 3.4: Binding parameters of R592A with respect to wild-type ADAAD using fluorescence spectroscopy was calculated for (a) ATP binding in the absence of stem-loop DNA, (b) ATP binding in the presence of stem-loop DNA, (c) stem-loop DNA binding in the absence of ATP, (d) stem-loop DNA binding in the presence of ATP. Each analysis represents the average ± standard deviation of two independent experiments. The protein concentration used in these analyses was 0.5 µM. All the data were fitted using one-site saturation model.

(a)   (b)

Figure 3.5: Global conformational changes recorded by CD spectroscopy in R592A as compared to the wild-type protein in the absence and presence of ligands. Comparison of CD structures for (a) protein alone and, (b) protein incubated with stem-loop DNA in the presence of ATP. The protein concentration used in CD analyses was 0.1 mg/ml.

Histidine 594 is important for an interaction with stem-loop DNA in presence of ATP

In case of H594A mutant, the interaction of ATP in the absence and presence of stem-loop DNA was found to be approximately the same as wild-type ADAAD (Figure 3.6a and 3.6b) (Table 3.1). Next, I investigated the stem-loop DNA binding in absence of ATP, and again I found no change in stem-loop DNA binding due to H594A mutation (Figure 3.6c) (Table 3.1). Finally, I measured the stem-loop DNA binding in presence of ATP and found that it was decreased by approximately two-fold, suggesting the importance of histidine 594 in stem-loop DNA binding after saturation with ATP (Figure: 3.6d) (Table 3.1). CD spectroscopy analysis showed that the conformation of H594A was similar to that of the wild-type ADAAD in the absence of ligands

(Figure 3.7a). However, in the presence of ligands a drastic change in conformation was observed in the mutant protein as compared to wild-type ADAAD (Figure 3.7b). Dichroweb analysis showed no change in alpha helical and beta sheet structure of the mutant protein in the absence of ligands (Table 3.2). However, a significant two-fold increased alpha helical content was observed in H594A after incubating the protein with stem-loop DNA in the presence of ATP (Table 3.2).

Thus, the impaired ATPase activity exhibited by H594A is not due to conformational changes in the protein. Rather, the impaired stem-loop DNA binding in presence of ATP leads to a conformational change in the protein resulting in slower ATP hydrolysis.
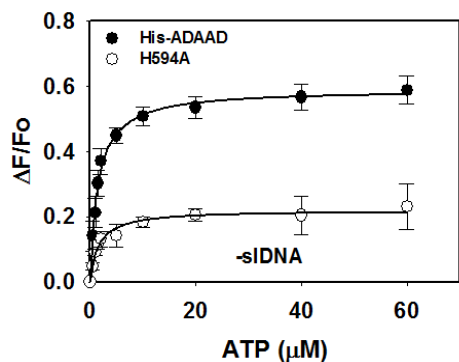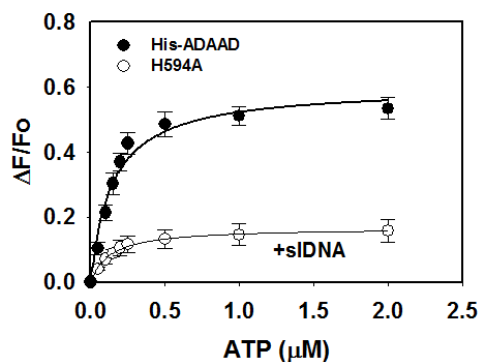
(a)   (b)
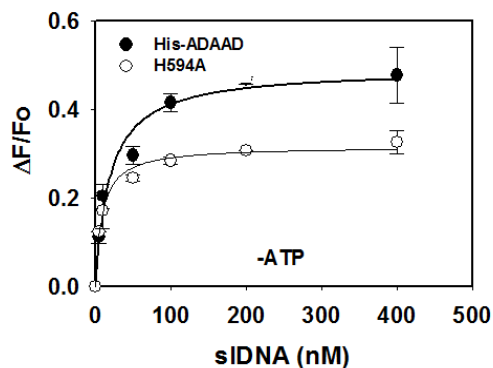
(c)
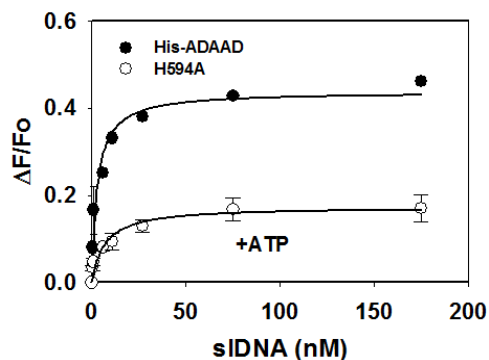
(d)

Figure 3.6: Binding plots for H594A mutant protein. (a) ATP binding in the absence of stem-loop DNA, (b) ATP binding in the presence of stem-loop DNA, (c) stem-loop DNA binding in the absence of ATP and, (d) stem-loop DNA binding in the presence of ATP. Each analysis represents the average of two independent experiments, whose standard deviations are shown in the figures. The protein concentration used in these analyses was 0.5 µM. All the data were fitted using one-site saturation model.
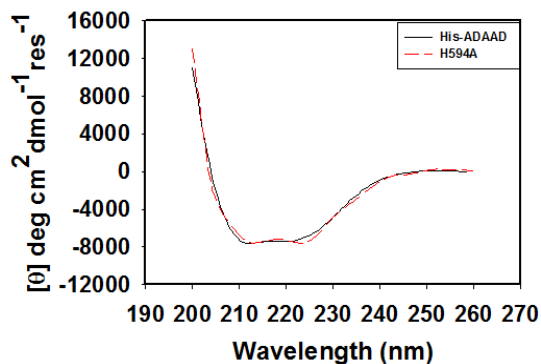
(a)
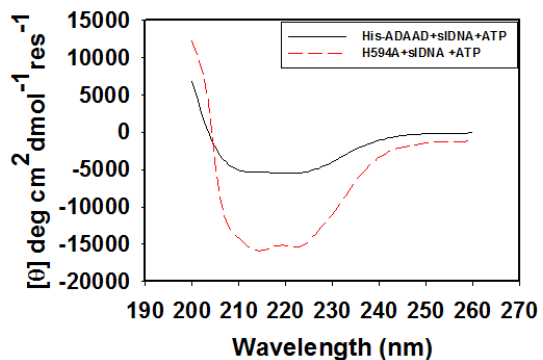
(b)

Figure 3.7: Global conformational changes in motif VI mutant H594A as compared to the wild-type ADAAD protein was recorded by CD spectroscopy. Comparison of secondary structure parameters for (a) protein alone and, (b) protein incubation with stem-loop DNA in presence of ATP. The protein concentration used in CD analyses was 0.1 mg/ml.

Arginine 595 residue is critical for maintaining the conformational integrity of ADAAD

Finally, I analyzed the role of R595 in ligand binding using fluorescence spectroscopy. In case of both R595A and R595K, binding results showed no change in ATP binding in the absence and presence of stem-loop DNA, suggesting that R595 was not involved in ATP binding (Figure 3.8a and 3.8b) (Figure 3.10a and 3.10b) (Table 3.1). Further, stem-loop DNA binding in the absence of ATP was also found to be unchanged for the two mutants as compared to the wild-type protein (Figure 3.8c and 3.10c) (Table 3.1). However, to my surprise, stem-loop DNA binding in the presence of ATP was approximately 10-fold defective in case of R595A and 2-fold weaker in case of R595K binding (Figure 3.8d and 3.10d) (Table 3.1). This result signifies the importance of positive charge of arginine in interaction with stem-loop DNA in the presence of ATP. CD spectroscopy showed that the conformation of both R595A and R595K was changed as compared to the wild-type protein even in the absence of the ligands (Figure 3.9a and 3.11a). Addition of stem-loop DNA to the protein after saturating with ATP also resulted in a greater change in secondary structure as compared to the wild-type ADAAD (Figure 3.9b and 3.11b). dichroweb analysis showed that mutation of arginine results in increase in the alpha helical of the protein to two-fold both in the absence and presence of stem-loop DNA after saturating the protein with ATP (Table 3.2). This study, hence, confirms the role of positive charge of arginine

83

in maintaining the conformation of the protein, thereby in interaction with stem-loop DNA in the presence of ATP. Loss in conformation integrity results in weaker interaction with stem-loop DNA in the presence of ATP and therefore, loss in ATPase activity.

(a)

(b)

(c)

(d)
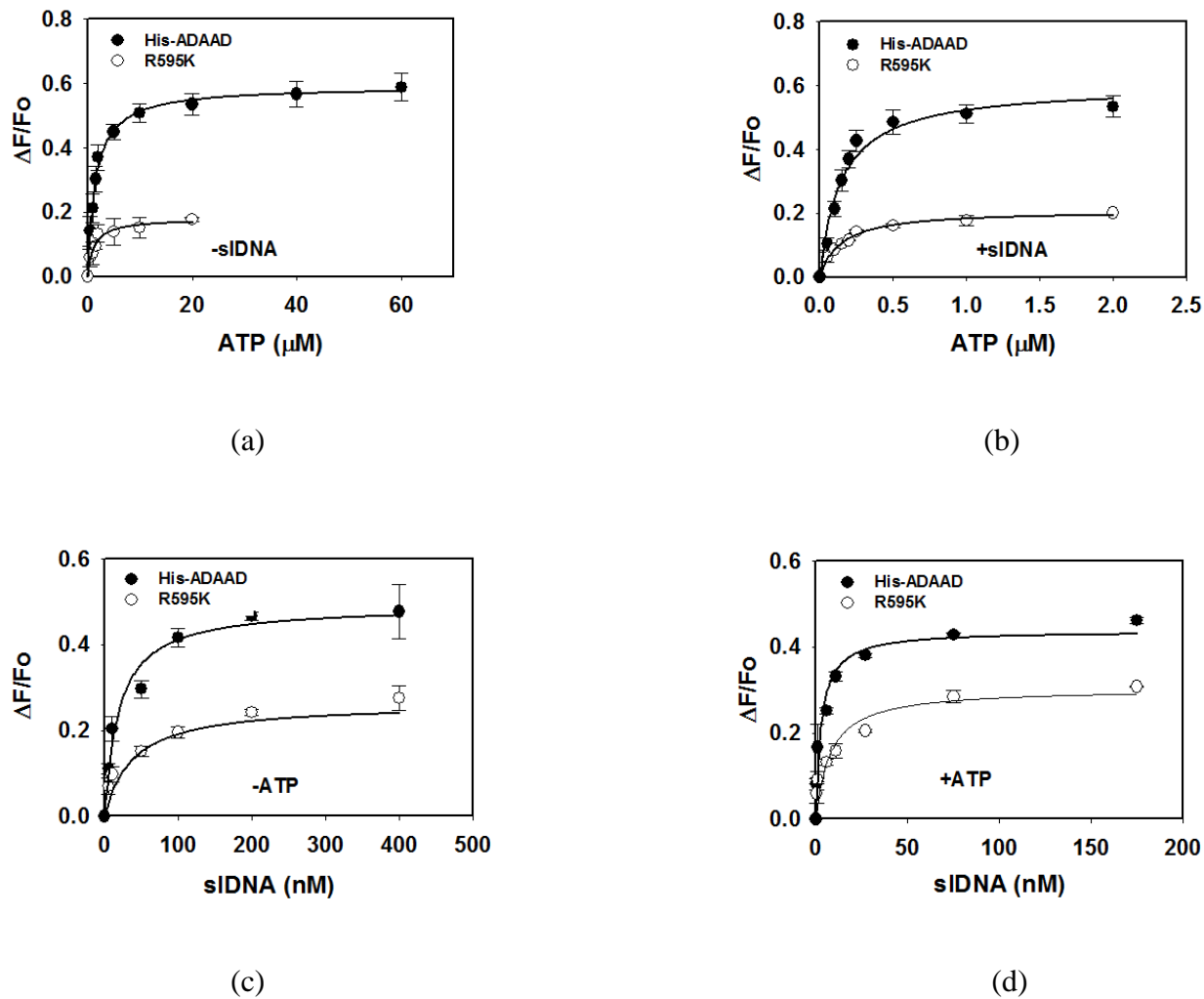
Figure: 3.8: Binding plots for motif VI mutant R595K and wild-type ADAAD protein. (a) ATP binding in the absence of stem-loop DNA, (b) ATP binding in the presence of stem-loop DNA, (c) stem-loop DNA binding in the absence of ATP and, (d) stem-loop DNA binding in the

presence of ATP. Each analysis represents the average of two independent experiments, whose standard deviations are shown in the figures. The protein concentration used in these analyses was 0.5 µM. All the data were fitted using one-site saturation model.


(a)   (b)
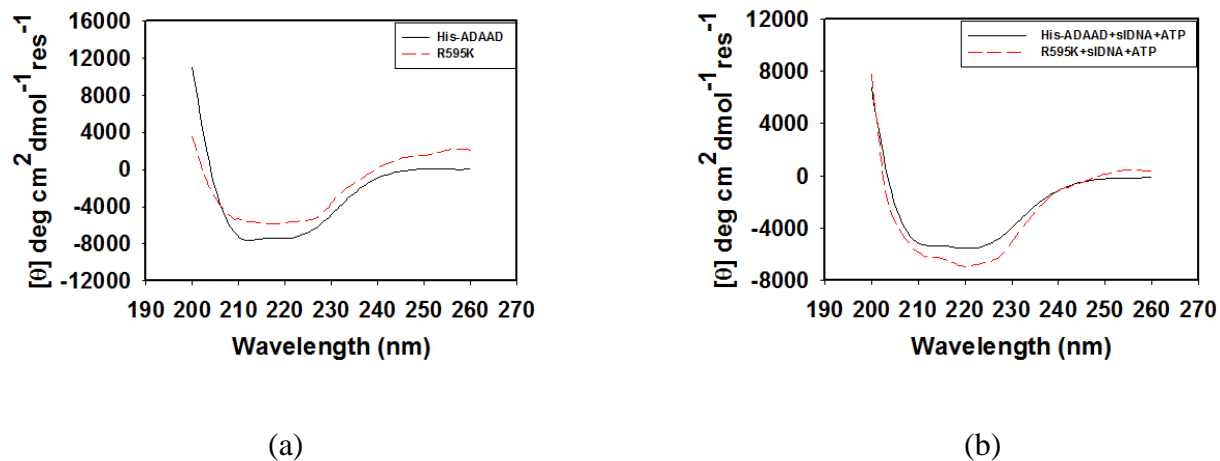
Figure: 3.9: Global conformational change recorded by CD spectroscopy in motif VI mutant R595K as compared to the wild-type protein for (a) protein alone and (b) after incubation with stem-loop DNA in the presence of ATP. The protein concentration used in CD analyses was 0.1 mg/ml.
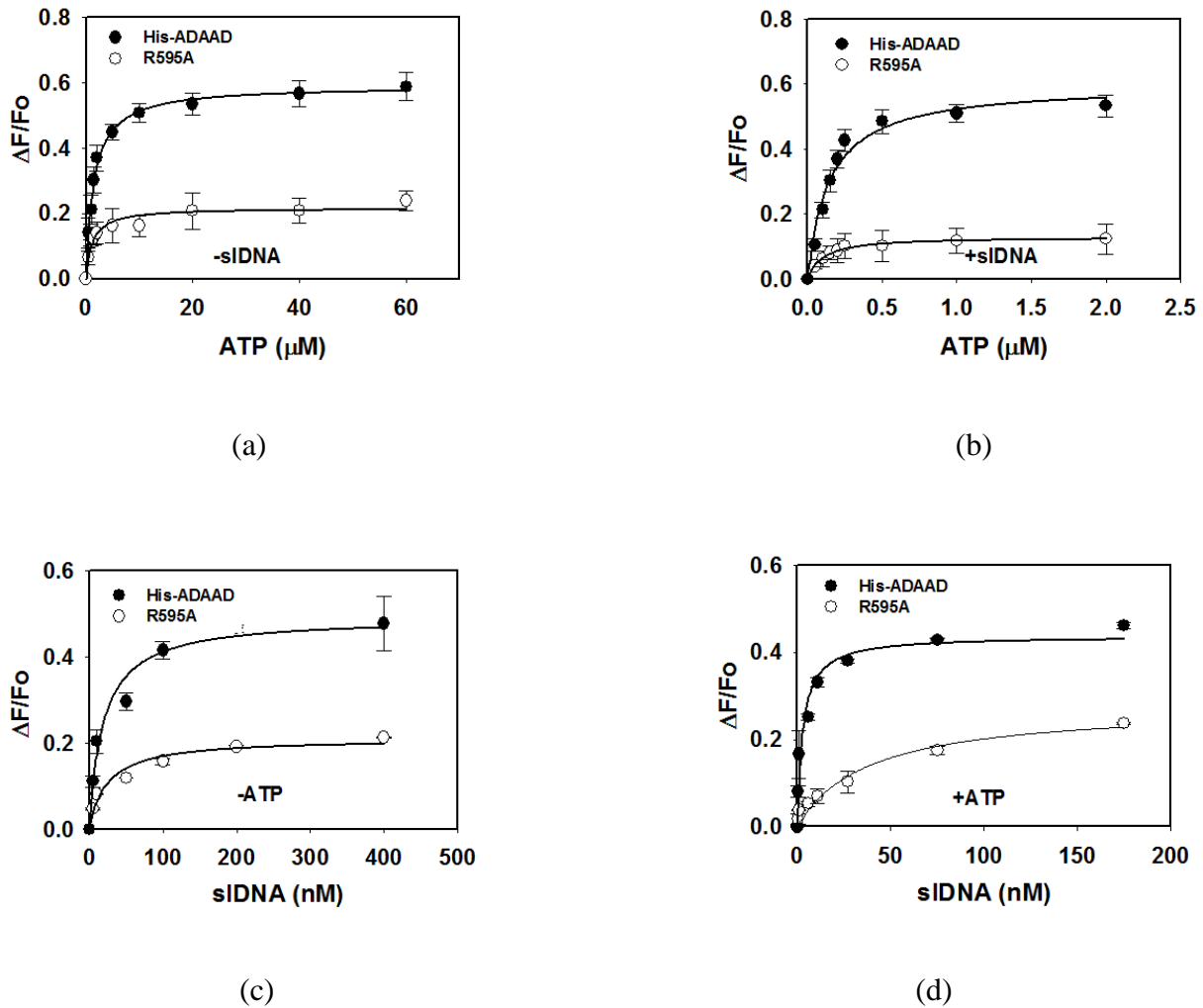
Figure 3.10: Comparison of binding plots for motif VI mutant R595A and wild-type ADAAD after the addition of following ligands: (a) ATP binding in the absence of stem-loop DNA, (b) ATP binding in the presence of stem-loop DNA, (c) stem-loop DNA binding in the absence of ATP and, (d) stem-loop DNA binding in the presence of ATP. Each analysis represents the average of two independent experiments, whose standard deviations are shown in the figures. The protein concentration used in these analyses was 0.5 µM. All the data were fitted using one-site saturation model.

(a)   (b)
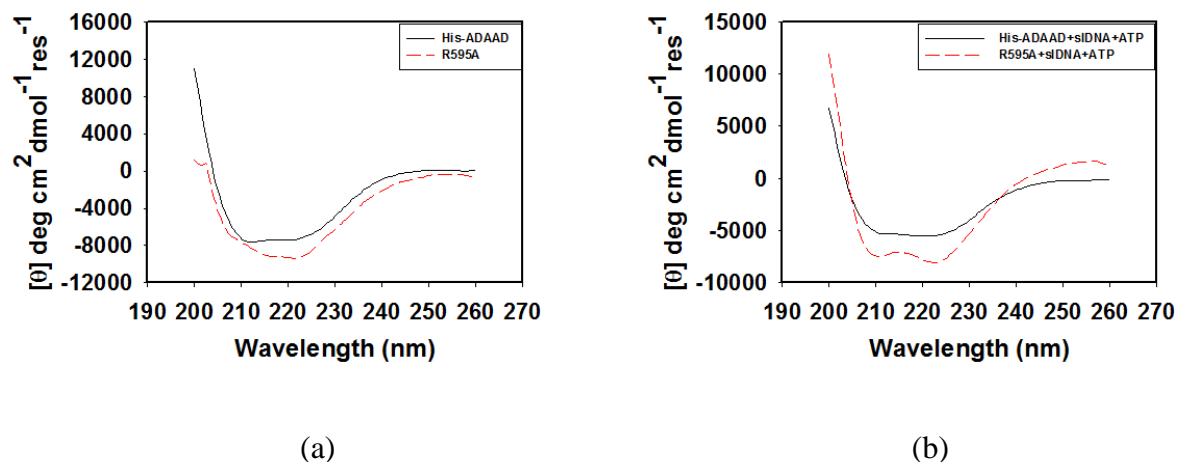
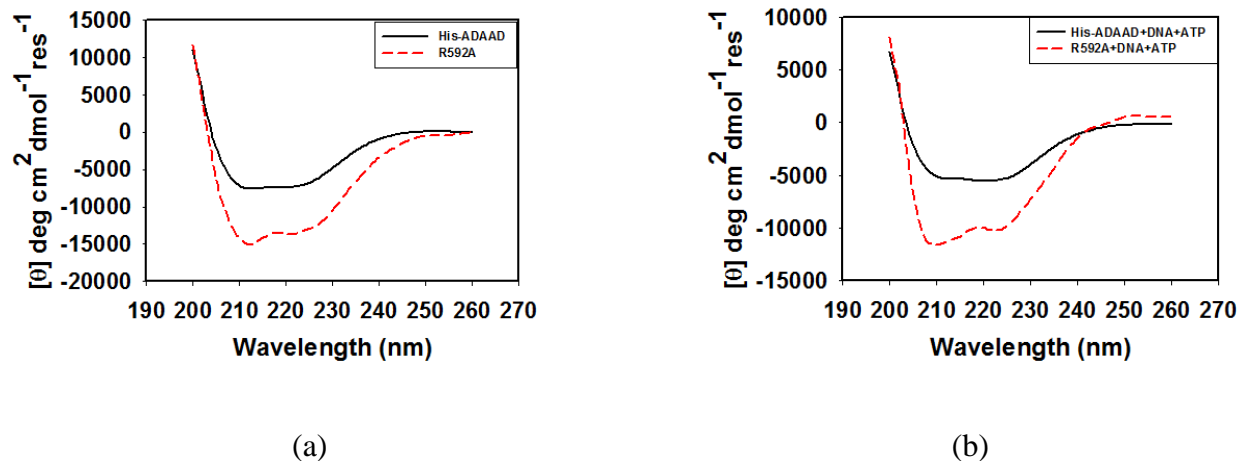**Figure 3.11**: Global conformational change recorded by CD spectroscopy in motif VI mutant R595A as compared to the wild-type ADAAD protein. Secondary structure parameters were compared for (a) protein alone and, (b) protein incubation with stem-loop DNA in presence of ATP. The protein concentration used in CD analyses was 0.1 mg/ml.

Table 3.1: Ligand-protein interactions for motif VI mutants were calculated by fluorescence spectroscopy and fitted as one site saturation hyperbolic curve.

	ADAAD	R592A	H594A	R595K	R595A
ATP (X 10^{-6}M)	(1.6 ± 0.5)	(2.3 ± 0.3)	(1.6 ± 0.2)	(1.2 ± 0.2)	(1.2 ± 0.2)
ATP(+slDNA) (X 10^{-6}M)	(0.14 ± 0.03)	(0.24 ± 0.02)	(0.13 ± 0.01)	(0.11 ± 0.01)	(0.1 ± 0.01)
slDNA (X 10^{-9}M)	(19.9 ± 4.9)	(31.5 ± 7)	(10.9 ± 2.6)	(25.1 ± 7.4)	(24.0 ± 0.63)
slDNA(+ATP) (X 10^{-9}M)	(3.4 ± 0.2)	(1.45 ± 0.7)	(6.6 ± 0.9)	(7.9 ± 2.1)	(32.9 ± 8.9)

Table 3.2: Dichroweb analysis (using K2D3 software) to calculate percentage of α helix and β sheet present in wild-type ADAAD and motif VI mutants, both in the absence of ligands and presence of stem-loop DNA after incubating with ATP.

	ADAAD	R592A	H594A	R595K	R595A
PROTEIN	35% α helix ,16% β sheet	56% α helix ,10% β sheet	34% α helix ,16% β sheet	27% α helix ,21% β sheet	31% α helix ,11% β sheet
PROTEIN+ slDNA+ATP	26% α helix ,25% β sheet	41% α helix ,19% β sheet	50% α helix ,19% β sheet	30% α helix ,21% β sheet	53% α helix ,25% β sheet

Table 3.3: Calculation of kinetic parameters for H594A and its comparison with wild-type ADAAD.

Protein	K_M (nM)	V_{max} (μmol NADH oxidized/min)	k_{cat} (min⁻¹)	k_{cat}/K_M (nM⁻¹ min⁻¹)
ADAAD	0.88± 0.22	2.18± 0.03	14.56±0.2	17.26± 1
H594A	36.04± 4.54	0.54± 0.2	21.5± 8.5	0.59± 0.2

DISCUSSION

Helicase families SF1 and SF2 possess seven conserved helicase motifs necessary for ATP hydrolysis (Fairman-Williams et al., 2010; Gorbalenya and Koonin, 1993). Both classes of families possess the arginine rich-motif VI. The function of motif VI has been characterized in ATP binding and hydrolysis in many SF1 and SF2 helicases but not in SWI2/SNF2 proteins. Hence, I initiated the biochemical characterization of motif VI of ADAAD in ligand binding and ATP hydrolysis. My studies have shown the difference between the biochemical activities of motif VI of SWI2/SNF2 proteins and DEAD box containing RNA helicases.

Various roles for the motif VI in SF1 helicases have been reported. PcrA and Rep helicases has shown the importance of arginine residue of motif VI in an interaction with ATP (Korolev et al., 1997; Velankar et al., 1999). However, eIF4A has shown the importance of arginine residue primarily in RNA interaction and ATP catalysis (Pause et al., 1993). I present in this chapter the first study on the significance of motif VI of ADAAD for ATP and stem-loop DNA interaction, conformational integrity and understanding the molecular mechanism of ATP hydrolysis. The motif VI in SWI2/SNF2 proteins including ADAAD is made up of three positively charged amino acid residues, two arginine and one histidine that can potentially make contacts with the negatively charged phosphate backbone of DNA. Therefore, in my first objective, I have focused on the biochemical role of positively charged residues of motif VI in stem-loop DNA binding and ATP hydrolysis. Mutations introduced in these residues (H594A, R592A, R595K and R595A) results in loss in ATPase activity. This loss in ATPase activity appears to be predominantly due to a decrease in the affinity for stem-loop DNA in the presence of saturated

amount of ATP. The defect in stem-loop DNA binding in presence of ATP was only two-fold in case of R595K, however, this defect was about 10-fold for R595A mutation.

Further, R595 appears to be important for stem-loop DNA binding only in the presence of ATP thus, reinforcing the importance of positive charge/basic nature of arginine in making a contact with stem-loop DNA. In harmony with our results, the crystal structure of a SF1 helicase, HCV NS3 helicase have also shown the significance of positive charge guanidium groups of arginine residues of motif VI in binding the negatively charged phosphate backbone of DNA (Kim et al., 1998). Taken together, our data implies that lack of stem-loop DNA binding in presence of ATP is a contributing factor for the loss in ATPase activity in case of the mutant proteins.

The importance of conserved motif VI residues in coupling ssDNA interaction and ATP hydrolysis activity to perform DNA unwinding function has been shown experimentally in the SF1 family member UvrD protein (Hall et al., 1998). This coupling was proposed to be directed not via direct interaction with nucleotide but by inducing a conformational change in the protein (Hall et al., 1998). The CD spectroscopy performed using the wild-type ADAAD and the mutant allowed me to probe whether the loss in ATPase activity was also because of a conformational change in the mutant protein. There was no significant change observed in the case of histidine mutant H594A in the absence of the ligand but a conformation change as compared to the wild-type protein was observed in presence of stem-loop DNA and ATP. In contrast, the secondary structure of R592A, R595A, and R595Kwas altered as compared to the wild-type protein even in the absence of the ligands indicating the importance of these residues in maintenance of the secondary structure of the protein.

All these results led me to conclude that both the arginine residues have a strong role, whereas the histidine residue has a moderate role in maintaining ATPase activity. The arginine residues of motif VI are important for maintaining the conformational integrity of the protein and arginine 595 for the contact with the stem-loop DNA in the presence of ATP. In summary, my studies lead me to suggest that motif VI also plays a role in coupling the stem-loop DNA binding to the ATPase activity.

4: Study of inter-domain and intra-domain cross-talk in ADAAD.

INTRODUCTION

As stated in the previous chapters, the SNF2 proteins are known to possess conserved helicase motifs within their catalytic core (Dürr et al., 2006; Flaus et al., 2006; Gorbalenya and Koonin, 1993). The catalytic core consists of two RecA-like domains with motifs I, Ia, II and III being present in RecA-like domain 1A while the RecA-like domain 2A consists of motifs IV, V and VI (Dürr et al., 2006). Studies have shown that the residues present in these motifs mediate both inter-domain as well as intra-domain interaction necessary for ligand interaction and ATP hydrolysis. Various models have also been proposed to show how conformational changes in the protein after ligand interaction induces the two RecA-like domains to come together to form a closed complex that is competent for ATP hydrolysis and DNA unwinding (Dürr et al., 2005; Thomä et al., 2005).

A general method of communication between two RecA like domain involves interaction of arginine finger present in one RecA-like domain with the nucleotide interacting active site of the other RecA-like domain (Caruthers et al., 2000). In this context, a salt bridge formation between motif VI present in RecA-like domain 2A with motif II present in RecA-like domain1A has been shown to be important in SF2 helicase members like UvrB, HCV and eIF4A (Caruthers et al., 2000; Theis et al., 1999; Yao et al., 1997). In these RNA helicases the first amino acid of motif VI, glutamine in case of UvrB and HCV and histidine in case of eIF4A, forms a salt bridge with the second aspartate of motif II (Caruthers et al., 2000; Theis et al., 1999; Yao et al., 1997) (Figure: 4a). These interactions in SF2 helicase members are slightly different from those present in SF1 helicase proteins where the lysine amino acid of motif V from RecA-like domain 2A interacts with motif II of RecA-like domain 1A to form a salt bridge (Korolev et al., 1997;

Velankar et al., 1999). The crystal structure of eIF4A, a SF2 helicase, has also shown that motif V and VI connect the carboxyl terminal RecA-like domain 2A to the amino terminal RecA-like domain 1A, suggesting it as a mechanism for coupling DNA binding and ATP hydrolysis (Caruthers et al., 2000).

There are very few reports that suggest the importance of intra-domain interactions in helicases in coupling DNA binding to ATP hydrolysis. One such report came with the determination of crystal structures of DEAD-box proteins wherein the importance of phenylalanine of motif IV in forming van der Waal interactions with the first arginine of motif VI was shown (Banroques et al., 2008).

Despite these structural information, we still lack the data about the amino residues and their association with each other involved in inter-domain and intra-domain interactions in SWI2/SNF2 proteins. Thus, the main aim of this chapter is to explore the inter-domain and intra-domain interactions in ADAAD and to understand the importance of these interactions vis-á-vis ligand binding and ATP hydrolysis.

Hypothesis and objectives

In my first chapter, I have deciphered the role of motif VI in ligand binding and ATP hydrolysis. In this chapter, I will be exploring the role of this motif in mediating inter-domain as well as intra-domain communications with other motifs. Specifically, I decided to study the interaction between motif VI and motif Ia; motif VI and motif II; and motif VI and motif IV present in ADAAD(Figure: 4b). Of these, the first two interactions (between motif VI and motif Ia; motif VI and motif II) constitute inter-domain interactions as motif Ia and II are present in the RecA-like domain 1A while the third interaction constitutes intra-domain interaction as both motif VI and motif IV are present in RecA-like domain 2A (Figure 4.1 and 2).

The biochemical studies performed in SF2 helicase members like UvrB, HCV and eIF4A has shown a salt bridge formation between motif VI present in RecA-like domain 2A with motif II present in RecA-like domain1A. In these RNA helicases the first amino acid of motif VI, glutamine in case of UvrB and HCV and histidine in case of eIF4A, forms a salt bridge with the second aspartate of motif II. In contrast to DEAD box helicases, ADAAD, and indeed all SWI2/SNF2 proteins, possess a histidine residue in motif II (DEXH) and aspartate residue in motif VI (DRAHRIGQ) that were postulated to form a salt bridge with each other. To study this intra-domain interaction, I mutated D591 present in motif VI to histidine creating D591H mutant and also mutated H329 present in motif II in the background of D591H mutation to create H329D/D591H mutant.

In the crystal structures of SF2 helicases, the conserved phenylalanine of motif IV has been shown to stack on the conserved arginine of motif IV. The stacking interaction in case of SWI2/SNF2 proteins has not been reported in the available crystal structure. Therefore, to probe

whether this interaction exists in ADAAD, F507 (motif IV) and R592 (motif VI) was mutated to create a double mutant. The residues F507 and R592 were chosen as they correspond to the residues present in other SF2 helicases.

To explore these interactions, I designed the following experiments:

- Site-directed mutagenesis to construct double mutation in the background of pre-constructed single mutant protein of motif VI.

- Protein purification of the site-directed double mutants- H329D/D591H, W277A/R595H and F507A/R592A.

- NADH-coupled ATPase assay to determine the impact of these double mutations on ATP hydrolysis.

- Analysis of ligand interaction using fluorescence spectroscopy.

- Analysis of protein conformation in the absence and presence of ligands using CD spectroscopy.

- Molecular simulations to study the RMSD and SASA values for single mutant W277A, R595H and double mutant protein W277A/R595H.
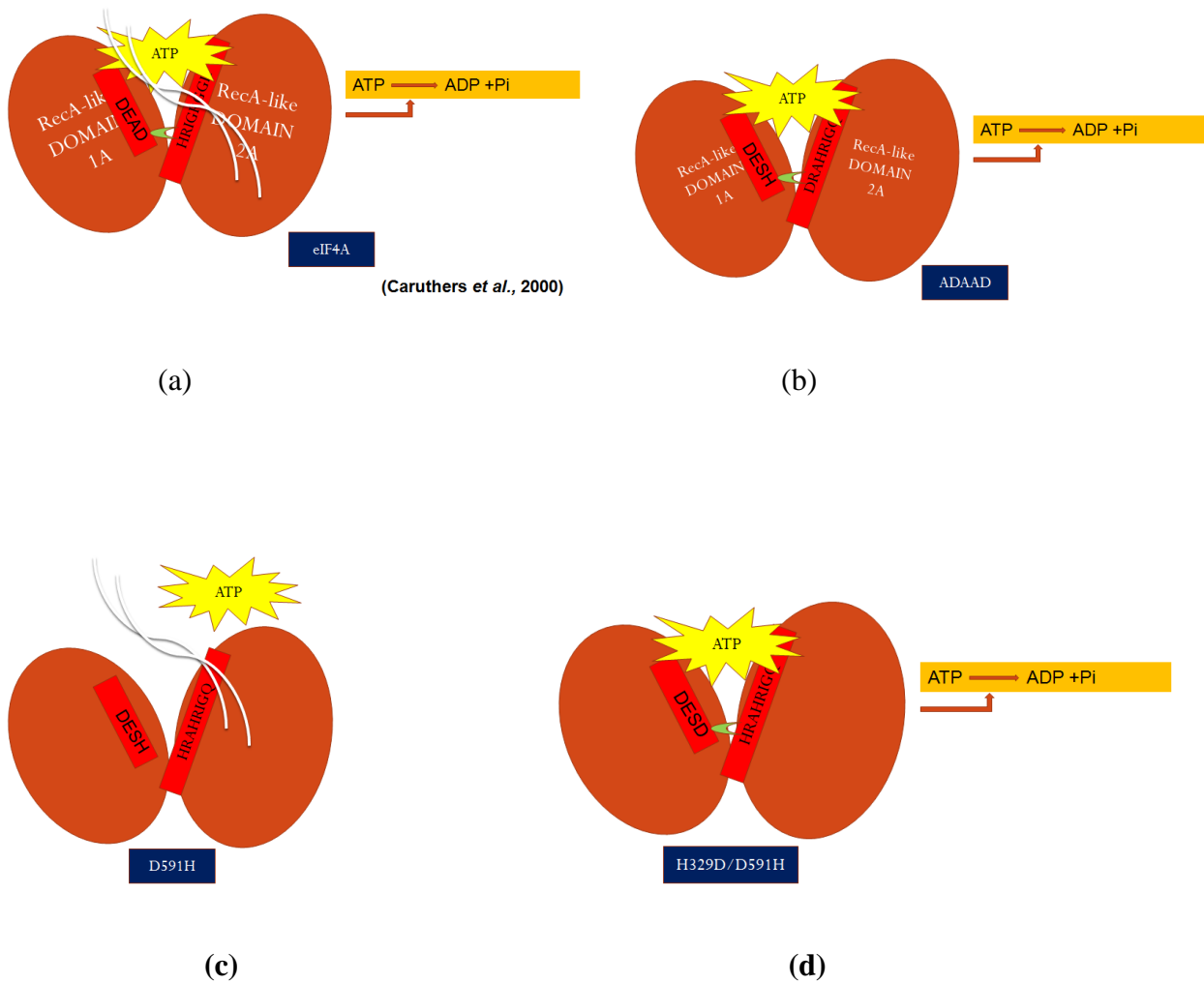
Figure 4a: Schematic representation of cartoons of eIF4A(a) and ADAAD(b). A hypothesis proposed for creating D591H mutant which is unable to hydrolyse ATP(c). A hypothesis proposed to show how ATPase activity should be restored in H329D/D591H mutant protein(d).

DOMAIN II

DOMAIN I

ATP
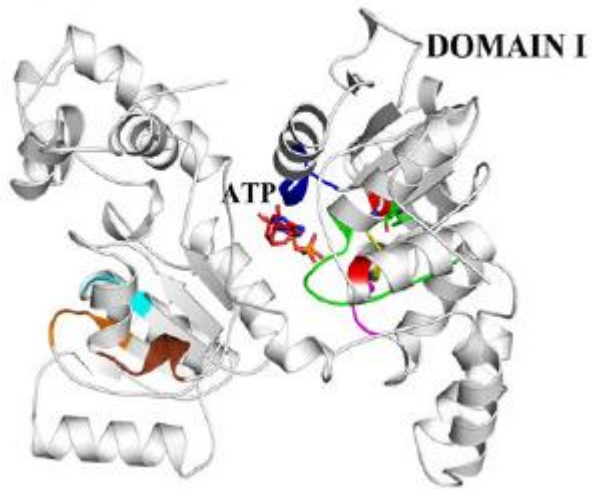
■ Motif-Q ■ Motif-I ■ Motif-Ia ■ Motif-II ■ Motif-III ■ Motif-IV ■ Motif-V ■ Motif-VI

Figure 4b: Molecular model of ADAAD based on the homology between ADAAD and Rad54.

The Highlighted color shows the conserved motif in the protein.

RESULTS

Aspartate 591 of motif VI is important for conformational integrity

As mentioned earlier, the crystal structure studies of eIF4A has shown a salt bridge formation between aspartate residue of motif II (DEAD) and histidine residue of motif VI (HRIGRXXR) (Caruthers et al., 2000). In contrast to DEAD box helicases, ADAAD, and indeed all SWI2/SNF2 proteins, possess a histidine residue in motif II (DEXH) and aspartate residue in motif VI (DRAHRIGQ) (Figure 4c). Thus, the position of histidine and aspartate are reversed in SWI2/SNF2 proteins as compared to other SF2 helicases. I hypothesized that the histidine of motif II might form a salt bridge with the aspartate residue of motif VI.

Conserved Motif	eIF4A	ADAAD
Motif II	DEAD	DESH
Motif VI	HRIGRXXR	DRAGRIGQ

Figure 4c: The sequence comparison between motif II and motif VI of eIF4A and ADAAD.

To study this intra-domain interaction, I first mutated D591 present in motif VI to histidine creating D591H mutant (Figure 4.1a). This mutant was unable to hydrolyze ATP (Figure 4.1b). The dissociation constant for D591H-ATP interaction, both, in the absence and presence of stem-loop DNA was approximately same as compared to wild-type ADAAD-ATP interaction, suggesting that D591 residue does not play important role in ATP binding (Figure 4.2a and b) (Table 4.1). The affinity for D591H-stem-loop DNA interaction in the absence of ATP was similar to that of wild-type ADAAD–stem-loop DNA interaction (Figure 4.2c) (Table 4.1).

However, the D591H-stem-loop DNA interaction in the presence of ATP was two-fold weaker as compared to the wild-type ADAAD-stem-loop DNA interaction in the presence of ATP (Figure 4.2d) (Table 4.1). CD spectroscopy revealed that the conformation of the mutant protein D591H (Figure 4.3a) both in the absence and presence of ligands was altered as compared to the wild-type ADAAD protein (Figure 4.3b). Furthermore, dichroweb analysis showed increase in alpha helical content of the D591H as compared to wild-type ADAAD both in the absence and presence of ligands (Table 4.2). From these results, I concluded that the D591 residue was necessary for maintaining the conformational integrity of the protein, disruption of which resulted in reduced affinity towards stem-loop DNA in presence of ATP and ultimately impaired ATPase activity.

(a)   (b)
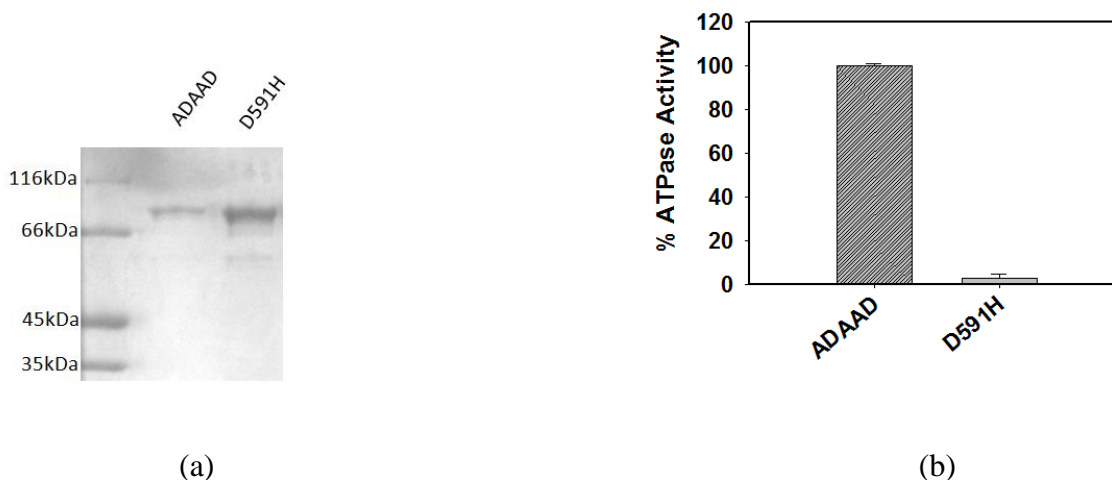
Figure: 4.1 (a) Purification of ADAAD and motif VI mutant D591H as shown by Coomassie Brilliant Blue staining. (b) The ATPase activity of D591H mutant in the presence of stem-loop DNA as compared to the wild-type ADAAD. Each analysis represents the average of two independent experiments, whose standard deviations are shown in the figures. The protein concentration used in these analyses was 0.1 µM.
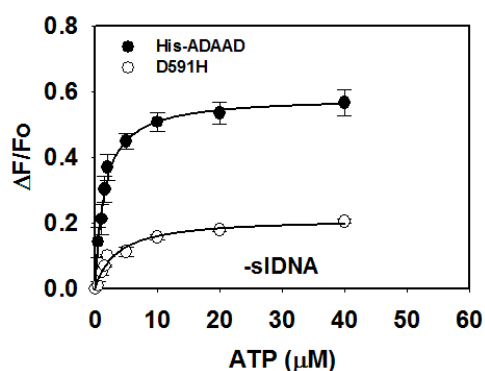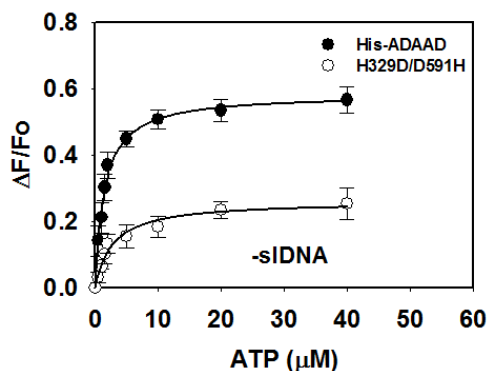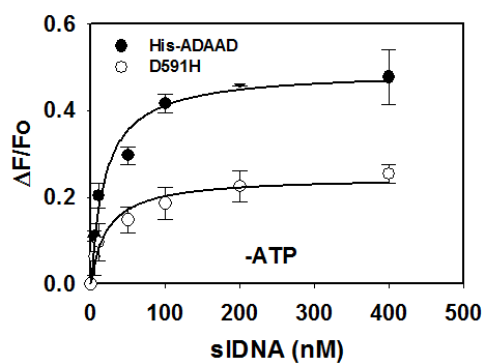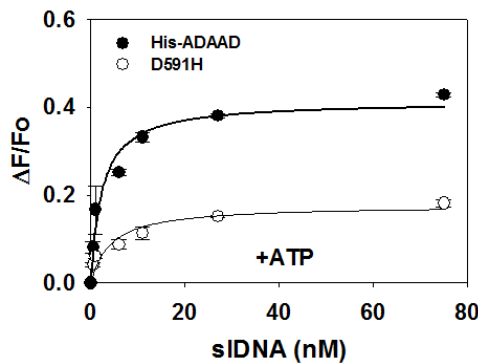
Figure: 4.2 Ligand-binding plots for D591H. Binding curves fitted to one-site saturation model for binding of D591H to (a) ATP in the absence of stem-loop DNA, (b) ATP in the presence of stem-loop DNA, (c) stem-loop DNA in the absence of ATP and, (d) stem-loop DNA in the presence of ATP. Each analysis represents the average of two independent experiments, whose standard deviations are shown in the figures. The protein concentration used in these analyses was 0.5 µM. All the data were fitted using one-site saturation model.
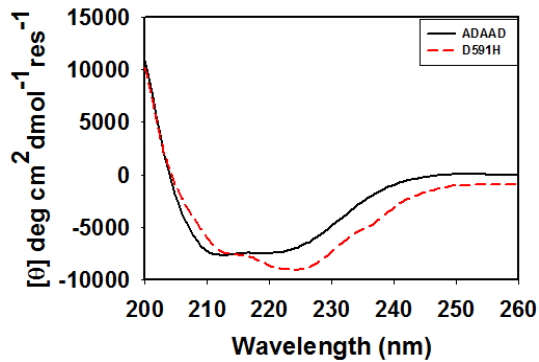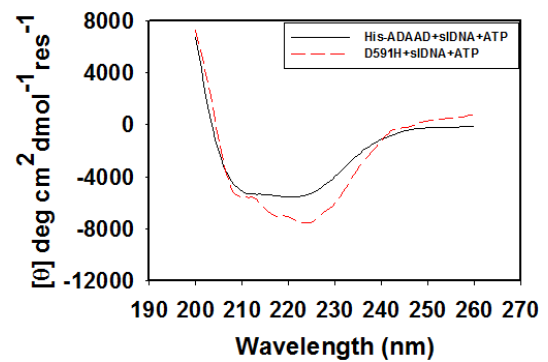
(a)   (b)

Figure: 4.3 Global conformational changes were recorded for motif VI mutant D591H and compared to wild-type ADAAD as following: (a) protein alone and, (b) protein after incubation with stem-loop DNA in the presence of ATP. The protein concentration used in CD analyses was 0.1 mg/ml.

H329D/D591H double mutant shows the possibilty of lack of salt bridge formation between the two RecA-like domains of ADAAD

Next, I created a double mutant where histidine of motif II was mutated to aspartate, H329D, in the background of D591H mutant. I hypothesized that if there was a salt bridge formation between these two residues of motif II and motif VI, a restoration of ATPase activity should be seen in double mutation H329D/D591H. The double mutant protein H329D/D591H was purified and checked for ATPase activity (Figure 4.4a). To our surprise and in contrast to our assumption, the ATPase activity was not restored in the double mutant H329D/D591H, suggesting a possibility of lack of salt bridge formation between motif II and motif VI in ADAAD (Figure

4.4b). Further, ATP binding both in the absence and presence of stem-loop DNA was found to be unchanged in the double mutant H329D/D591H as compared to wild-type ADAAD-ATP interaction (Figure 4.5a and b) (Table 4.1). In addition, stem-loop DNA binding affinity was also seen to be unchanged in absence of ATP (Figure 4.5c) (Table 4.1); however, stem-loop DNA binding affinity in the presence of ATP decreased three-fold as compared to the interaction of stem-loop DNA with wild-type ADAAD (Figure 4.5d) (Table 4.1). The CD spectroscopy analysis showed that the conformation of the protein was also altered both in the absence and presence of ligands as compared to both wild-type ADAAD and D591H (Figure 4.6a and b). Additionally, dichroweb analysis also showed a significant increase in alpha helical content of the protein as compared to wild-type ADAAD and D591H mutant, both in the absence and presence of ligands confirming that the conformation of the double mutant protein is altered (Table 4.2).

(a)   (b)
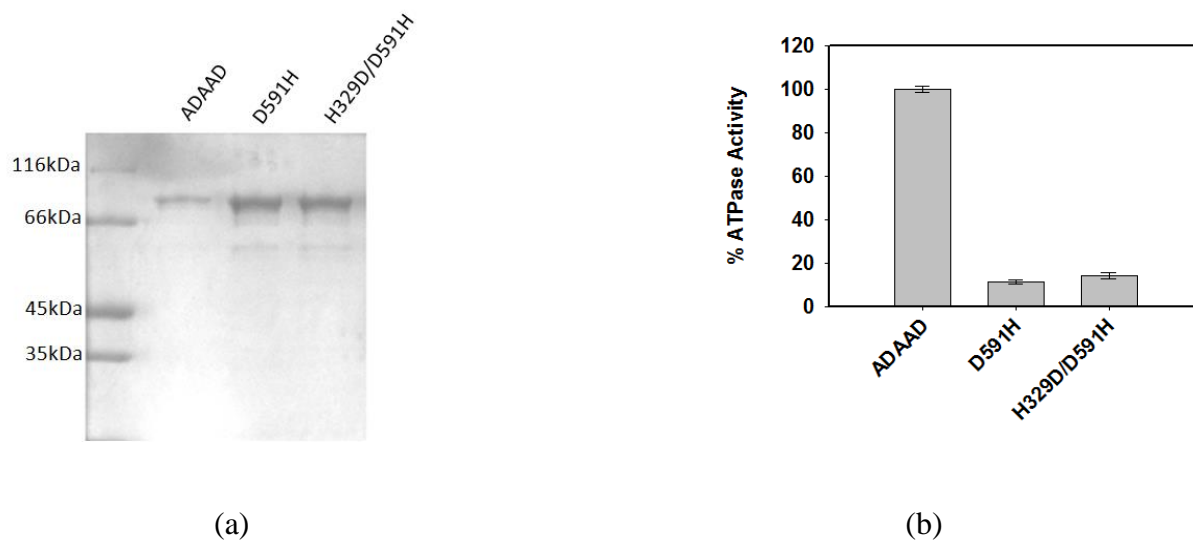
Figure: 4.4 (a) Purified double mutant protein H329D/D591H was resolved on SDS-10% polyacrylamide gel and stained by Coomassie Brilliant Blue staining. (b) Percentage ATPase

activity of H329D/D591H in the presence of stem-loop DNA was compared to the wild-type

ADAAD. Each analysis represents the average of two independent experiments, whose standard

deviations are shown in the figures. The protein concentration used in these analyses was 0.1

μM.

(a)

(b)

(c)

(d)
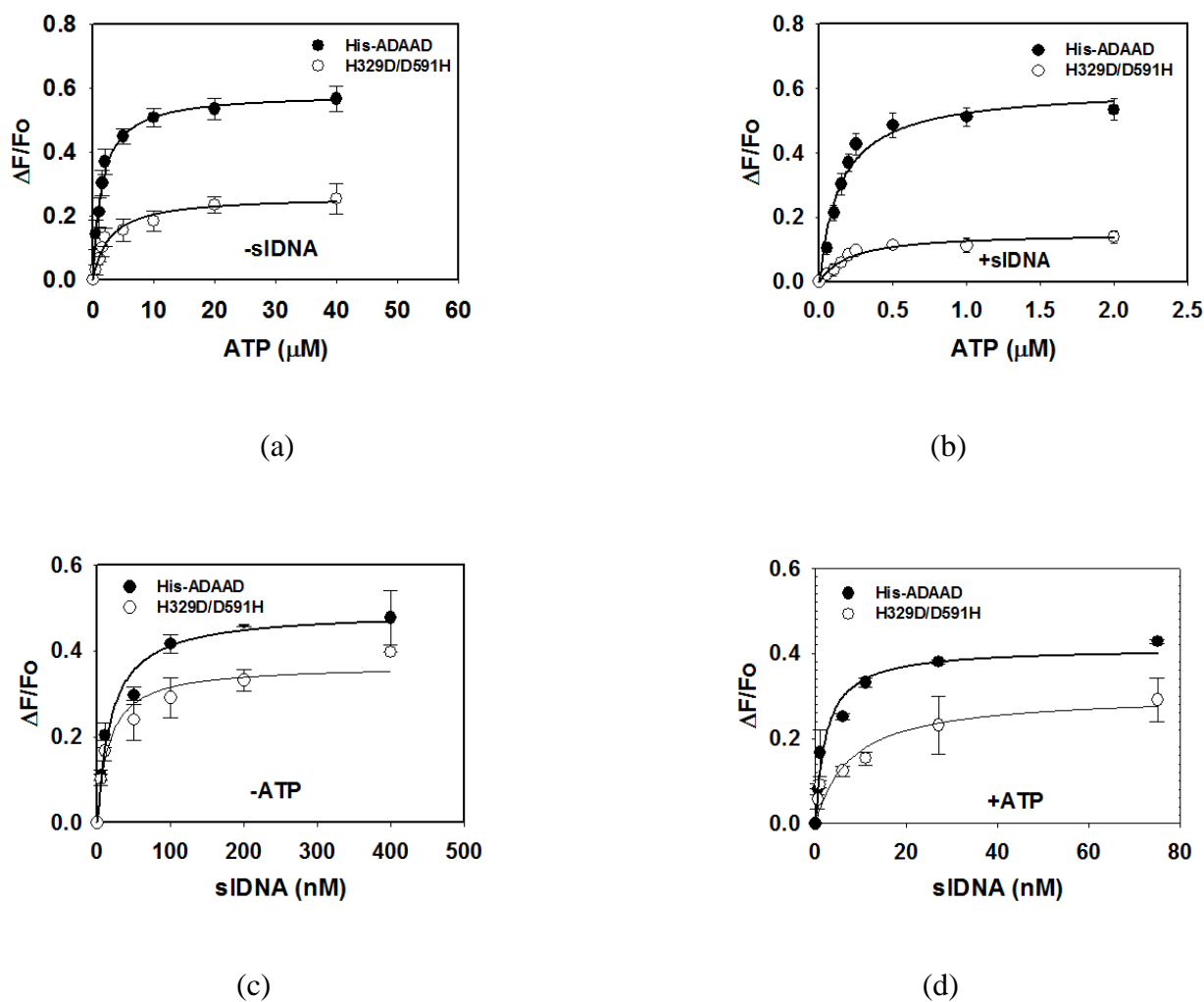
Figure: 4.5 Ligand-binding plots for H329D/D591H. The binding curves fitted to one-site

saturation model for the interaction of H329D/D591H with (a) ATP in the absence of DNA, (b)

ATP in the presence of stem-loop DNA, (c) stem-loop DNA in the absence of ATP and, (d)

stem-loop DNA in the presence of ATP. Each analysis represents the average of two independent

experiments, whose standard deviations are shown in the figures. The protein concentration used

in these analyses was 0.5 µM. All the data were fitted using one-site saturation model.

(a)   (b)
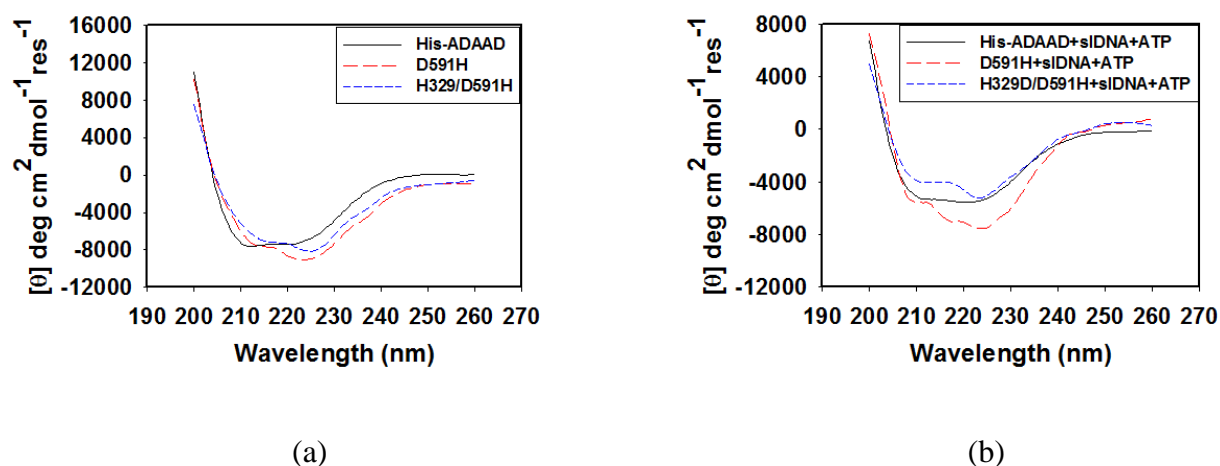
Figure: 4.6 Global conformational changes were plotted for: (a) H329D/D591H, D591H and

wild-type ADAAD in the absence of ligands and, (b) H329D/D591H, D591H and ADAAD after

incubation with stem-loop DNA in presence of ATP. The protein concentration used in CD

analyses was 0.1 mg/ml.

W277A/R595H double mutant shows an alteration in conformation and impaired stem-loop DNA binding in the presence of ATP

Dr. Meghna Gupta in our lab has shown that the conserved amino acid, W277, present in motif

Ia of RecA-like domain 1A plays a major role in stem-loop DNA binding in absence of ATP as

mutation of this residue to alanine impacts the interaction with stem-loop DNA by three-fold. This study was supported by molecular model of W277A that showed that the dynamic stability of the protein, as shown by RMSD values, decreased; possibly hindering the contact with stem-loop DNA (Figure 6.4a) (Gupta 2015). R595H, a conserved amino acid residue of motif VI present in RecA-like domain 2A, present on the opposite face of W277A, also showed a decrease in stem-loop DNA binding to four-fold in presence of ATP (Figure 6.2d) (Table 6.1). This defect in stem-loop DNA binding in R595H was also supported by molecular simulations like SASA and RMSD value calculations (Figure 4.10a and b). To decipher structural alterations in the mutant proteins as compared to the wild-type protein, the atomic coordinates of the wild-type ADAAD protein were obtained by homology modelling based on the crystal structure of a homologue protein SWI2/SNF2 ATPase core (Sulfolobus solfataricus) and the structures of mutants were generated using ROSETTABACKRUB point mutations server sampled by Monte Carlo simulated annealing using the Rosetta all-atom force field. The model showed that the two domains are separated by a cleft and connected via a loop. To study the effects of SMARCAL1 point mutations on ADAAD structure and dynamics, multiple molecular dynamics (MD) simulations of the wild-type as well as the mutants were performed with random seed for a period of 50 ns and the dynamic stability of the simulations was evaluated by calculating the root mean square deviation (RMSD) in comparison to the starting structure. It was observed that the RMSD of both the wild-type and the mutant proteins attained equilibrium during the 1 ns equilibrium dynamics. The average RMSD for the wild-type protein from both the runs was around 7A, indicating that the ATPase core was flexible, a pre-requisite for ligand binding (Fig. 3B). However, the average RMSD values for the mutants were 5.8A for R595H, indicating that the average overall fluctuations were slightly higher in the wild type ADAAD protein compared

to the R595H mutant proteins (Figure 4.10A). Next, the confirmation towards the conformational changes affecting the area accessibility for ligand binding was studied by solvent accessibility surface area (SASA), which give information about the biomolecular surface area that is accessible to solvent molecules. The SASA analysis for the wild-type protein was relatively higher than the R595H mutant protein throughout the simulation time period, hence clearly depicting the structural differences caused by the mutations(Figure 4.10B). As R595 is present on the face diagonally opposite to W277 and both show a defect in interaction with stem-loop DNA, I created a double mutant W277A/R595H to see whether there will be a synergistic defect in stem-loop DNA binding. I used the molecular simulation tool, RMSD calculation, to study the dynamic movement of W277A and R595H and found that RMSD value was altered in both the cases (Figure 4.10a). However, contrary to my hypothesis, molecular simulations showed that W277A/R595H had dynamic movement that was similar to the wild-type ADAAD (Figure 4.10a). This study suggested that stem-loop DNA binding should be restored in the W277A/R595H double mutant. I purified the double mutant W277A/R595H (Figure 4.7a) and found that the purified protein was unable to hydrolyze ATP (Figure 4.7b). Biochemical studies, using fluorescence spectroscopy, showed that the interaction of the protein with ATP both in the absence and presence of stem-loop DNA was similar to that of the wild-type ADAAD (Figure 4.8a and b) (Table 4.1). And as expected from our molecular simulation model, binding analysis confirmed that W277A/R595H binds to stem-loop DNA in the absence of ATP with the same affinity as the wild-type ADAAD (Figure 4.8c) (Table 4.1). However, the binding affinity decreased for the interaction of the double mutant W277A/R595H with stem-loop DNA in the presence of ATP as compared to ADAAD-stem-loop DNA interaction, suggesting this could be a reason for inactive protein (Figure 4.8d) (Table 4.1). The CD analysis showed that the

conformation of the protein was altered as compared to the wild-type ADAAD protein and single mutant R595H both in the absence (Figure 4.9a) and presence of ligands (Figure 4.9b). Furthermore, dichroweb analysis has shown the increase in alpha helical structure in double mutant protein W277A/R595H, both in the absence and presence of ligands, as compared to wild-type ADAAD protein and R595H (Table: 4.2 and 6.2).

(a)   (b)
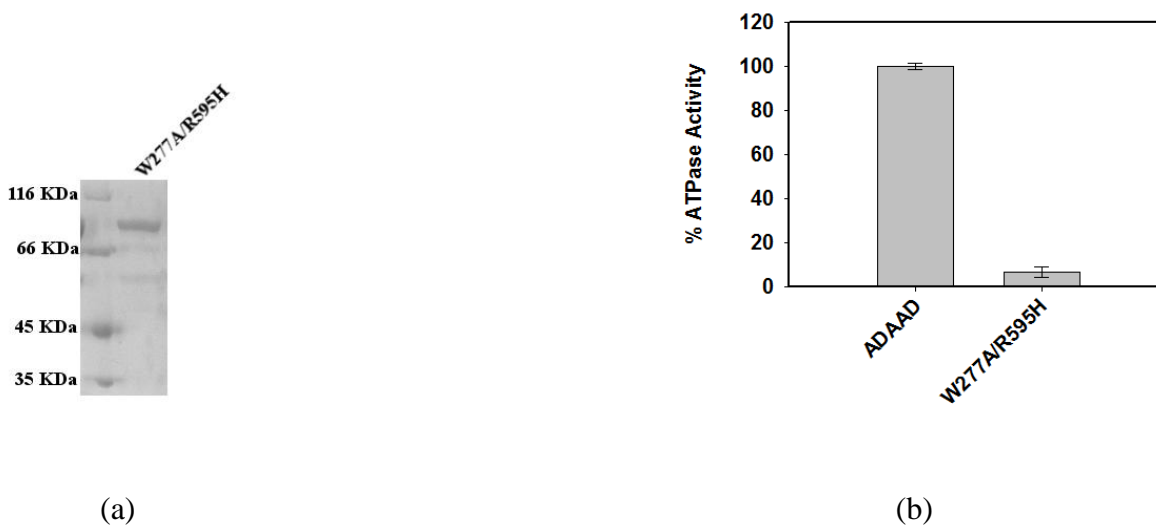
Figure: 4.7 (a) Purification of the double mutant protein W277A/R595H as analyzed on 10% SDS- polyacrylamide gel and stained using Coomassie brilliant blue. (b) ATPase activity of the double-mutant W277A/R595H in the presence of stem-loop DNA was compared to the wild-type ADAAD. Each analysis represents the average of two independent experiments, whose standard deviations are shown in the figures. The protein concentration used in these analyses was 0.1 μM.

(a)

(b)

(c)

(d)
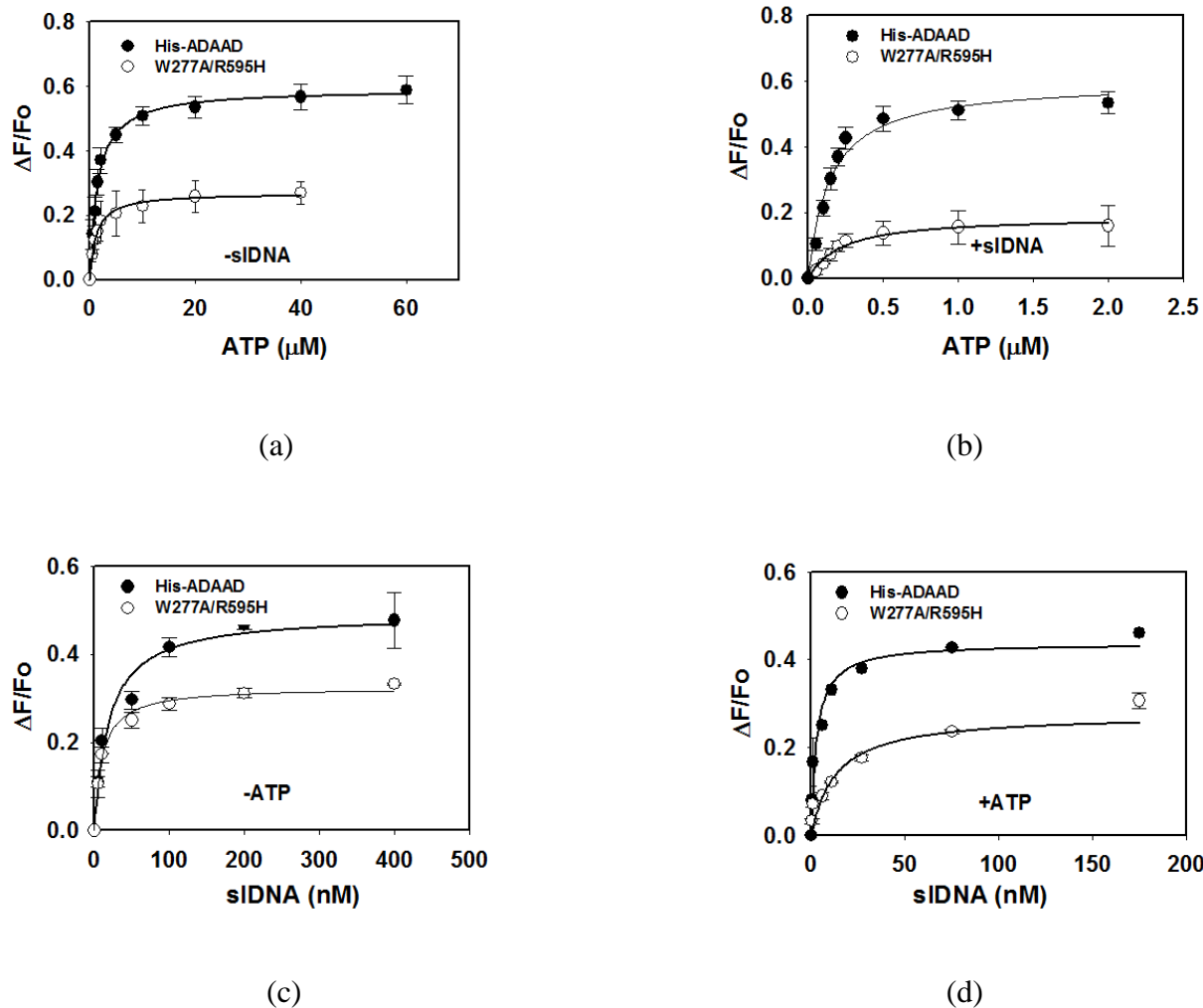
Figure: 4.8 Ligand-binding plots for W277A/R595H. The binding data was fitted to one-site saturation model for the interaction of W277A/R595H with (a) ATP in the absence of stem-loop DNA, (b) ATP in the presence of stem-loop DNA, (c) stem-loop DNA in the absence of ATP and, (d) stem-loop DNA in the presence of ATP. Each analysis represents the average ± standard deviation of two independent experiments. The protein concentration used in these analyses was 0.5 µM. All the data were fitted using one-site saturation model.
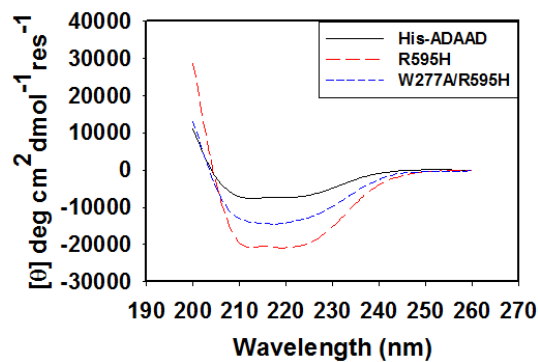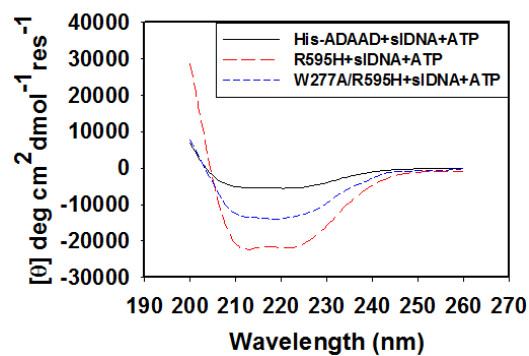
(a)   (b)

Figure: 4.9 Global conformational changes were plotted for: (a) W277A/R595H, R595H and

ADAAD in the absence of ligands and, (b) W277A/R595H, R595H and ADAAD after

incubation with stem-loop DNA in presence of ATP. The protein concentration used in CD

analyses was 0.1 mg/ml.
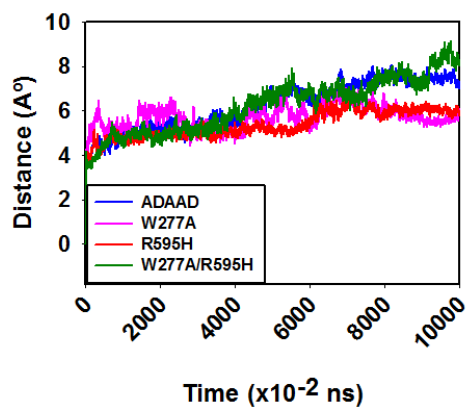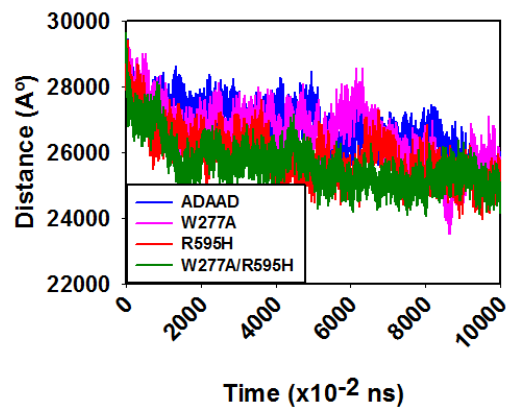
(a)   (b)

Figure: 4.10 Molecular simulation studies for ADAAD, W277A, R595H and W277A/R595H. (a) RMSD analysis of ADAAD, W277A, R595H, and W277A/R595H. (b) SASA analysis of ADAAD, W277A, R595H, and W277A/R595H.

F507A/R592A mutant completely alters the conformation of ADAAD

I created a double mutation in motif IV in the background of motif VI mutant R592A to study the synergistic effect of these two mutations, both present in a RecA-like domain 2A, on ligand binding and ATPase activity as previous studies have indicated an interaction between these two motifs (Banroques et al., 2008). The mutant protein was purified and was found to be incapable of hydrolyzing ATP (Figure 4.11a and b). Binding studies showed that the double mutant F507A/R592A was able to bind to stem-loop DNA in the absence of ATP with same affinity as the wild-type ADAAD protein (Figure 4.12c) (Table 4.1). It was also able to bind to ATP both in the absence and presence of stem-loop DNA with affinity similar to that of the wild-type ADAAD protein (Figure 4.12a and b) (Table 4.1). However, binding affinity for interaction with stem-loop DNA in the presence of ATP was weaker by three-fold (Figure 4.12d) (Table 4.1). CD analysis showed that the conformation of the double mutant F507A/R595H was completely altered and contained 0% alpha helical structure in the absence of ligands as compared to wild-type ADAAD protein and R592A mutant (Table 4.2) (Figure 4.13a). The structure remained altered even in the presence of stem-loop DNA and ATP suggesting the importance of these two residues in maintaining the secondary structure and ATPase function of ADAAD (Figure 4.13b).
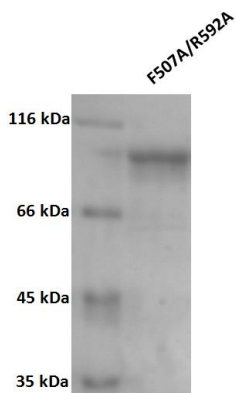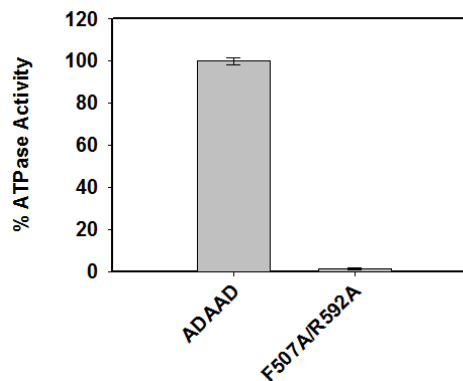
(a)   (b)

Figure 4.11: (a) Purification of the double mutant F507A/R592A. (b) The ATPase activity was calculated for the double mutant F507A/R592A in the presence of stem-loop DNA and compared to wild-type ADAAD. Each analysis represents the average ± standard deviation of two independent experiments. The protein concentration used in these analyses was 0.1 µM.
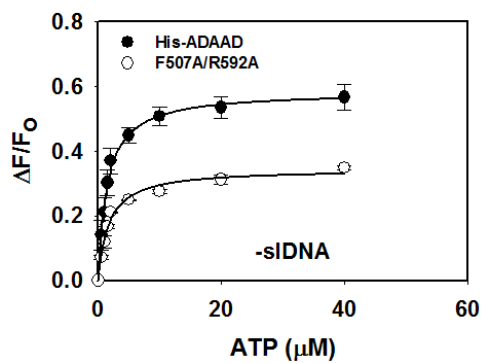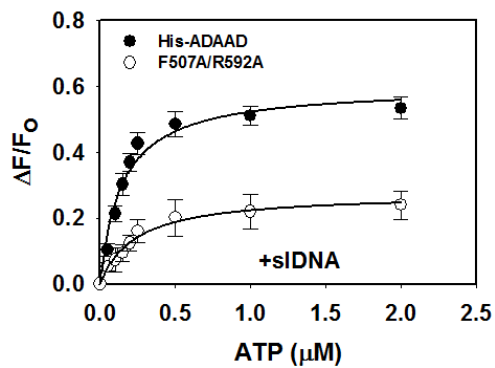

(a)   (b)
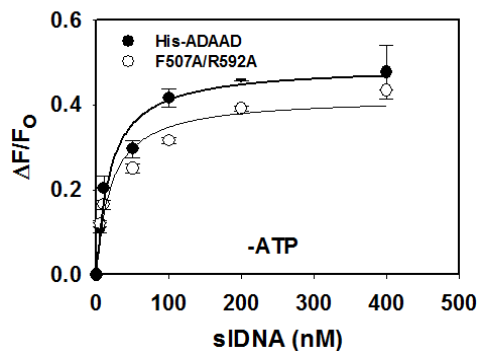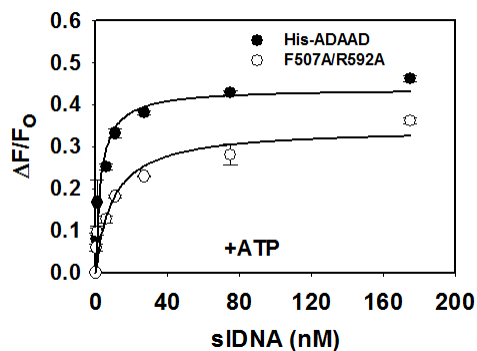
(c)   (d)

Figure 4.12: Ligand-binding plots for F507A/R592A. The binding data was fitted to one-site saturation model for the double mutant F507A/R592A for interaction with (a) ATP in the absence of stem-loop DNA, (b) ATP in the presence of stem-loop DNA, (c) stem-loop DNA in the absence of ATP and, (d) stem-loop DNA in the presence of ATP. Each analysis represents the average ± standard deviation of two independent experiments. The protein concentration used in these analyses was 0.5 µM. All the data were fitted using one-site saturation model.
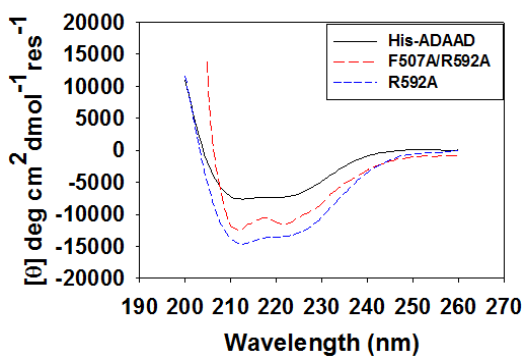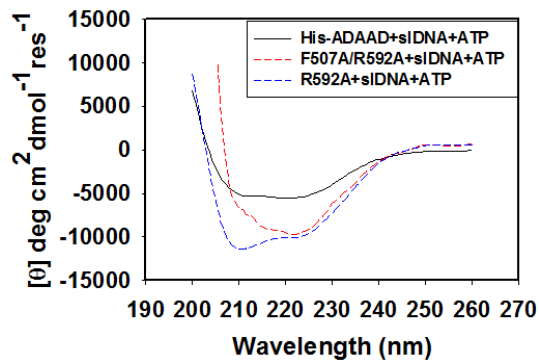
(a)   (b)

Figure 4.13: Global conformational changes were plotted for the double mutant F507A/R592A, R592A and wild-type ADAAD (a) in the absence of ligands and, (b) in the presence of stem-loop DNA after saturating the protein with ATP. The protein concentration used in CD analyses was 0.1 mg/ml.

Table 4.1: The K_d values calculated for ligand-protein interactions by fluorescence spectroscopy and fitted to one-site saturation hyperbolic curve.

	ADAAD	D591H	H329D/ D591H	W277A/ R595H	F507A/ R592A
ATP (X 10^{-6}M)	(1.6 ± 0.5)	(3.2 ± 0.3)	(2.8 ± 0.1)	(1.3 ± 0.4)	(1.7 ± 0.1)
ATP(+slDNA) (X 10^{-6}M)	(0.14 ± 0.03)	(0.23 ± 0.07)	(0.21 ± 0.06)	(0.21 ± 0.09)	(0.24 ± 0.04)
slDNA (X 10^{-9}M)	(19.9 ± 4.9)	(21.3 ± 7)	(16.1 ± 5.7)	(10.8 ± 3.9)	(27.8 ± 3.6)
slDNA(+ATP) (X 10^{-9}M)	(3.4 ± 0.2)	(6.9 ± 1.4)	(9.9 ± 2.2)	(16.8 ± 1.2)	(9.41 ± 0.93)

Table 4.2: Dichroweb analysis (using K2D3 software) to calculate percentage of α helix and β sheet present in wild-type ADAAD and mutants both in the absence of ligands and presence of stem-loop DNA after incubated with ATP.

	ADAAD	D591H	H329D/ D591H	W277A/ R595H	F507A/ R592A
PROTEIN	35% α helix ,16% β sheet	33% α helix ,17% β sheet	29% α helix ,22% β sheet	55% α helix ,11% β sheet	0% α helix ,100% β sheet
PROTEIN+ slDNA +ATP	26% α helix ,25% β sheet	29% α helix ,23% β sheet	13% α helix ,39% β sheet	46% α helix ,23% β sheet	0% α helix ,100% β sheet

DISCUSSION

Although the role of ADAAD in DNA-dependent ATP hydrolysis has been widely studied, the

mechanism by which ADAAD couples stem-loop DNA binding to ATP hydrolysis still remains

a mystery (Nongkhlaw et al., 2009). To understand the mechanism of ATP hydrolysis, I analyzed

the importance of amino acid residues postulated to form inter-domain and intra-domain

interaction in ADAAD. In my earlier chapter, I have shown that motif VI is important for

maintaining the conformational integrity of the protein and stem-loop DNA binding in presence

of ATP. The crystal structure of Rad54 has shown the importance of motif VI in mediating an

inter-lobe communication between the two RecA-like domains (Thomä et al., 2005). Motif VI, in

Rad54, is found to make a contact with sulfate ion which is bound via gamma phosphate binding

site of ATP present in motif I (Thomä et al., 2005). The crystal structure of eIF4A has also

shown the importance of motif VI in mediating a salt bridge formation with motif II (Caruthers

et al., 2000). In contrast, such a salt bridge was not observed in ADAAD as the double mutant

H329/D591H did not show any residual ATPase activity. The reason could be conformation of

the protein such that the RecA-like domain 2 is flipped out 180° with respect to RecA-like

domain 1 and such orientation in the presence of the ligands does not permit any interaction

between the two domains. This type of behavior has been reported for NS3 protein from

Hepatitis C virus where similar double mutant was unable to hydrolyze ATP.

The intra-lobe communication in Ded1 protein have been postulated to involve in cation-

π interaction between the phenylalanine residue of motif IV and arginine residue of motif VI .

In contrast, our studies suggest that the arginine of motif VI plays a more important role as the

double mutant shows the characteristics of the arginine mutant rather than the phenylalanine mutant in maintaining the ATPase activity and conformation of the protein.

Keeping all these reports in mind, I mutated the amino acid in motif VI of RecA-like domain 2A postulated to impair the network of connections with other amino acid present in RecA-like domain 1A or RecA-like domain 2A. I followed this mutation by creating double mutation to better understand the role of synergistic effect of both of these amino acid residues in ligand interaction and ATP hydrolysis. Thus, I performed a detailed dissection of motif VI (DRAGRIGQ) of ADAAD to better comprehend the role of D591, R592 and R595 in forming inter-domain and intra-domain interactions.

Based on eIF4A crystal structure studies, the side chain of histidine residue of motif II of ADAAD was predicted to form an interaction with aspartate residue of motif VI. The mutant D591H, where D591 present in motif VI was replaced by histidine, was unable to hydrolyze ATP and contrary to my hypothesis, the double mutant H329D/D591H was also unable to hydrolyze ATP. Binding studies showed that the interaction with ATP in the absence and presence of stem-loop DNA was unchanged in both D591H and H329D/D591H, signifying that D591 present in motif VI, does not play a role in ATP binding. The interaction of stem-loop DNA in the absence of ATP was also not altered in case of both the mutants; however, the interaction with stem-loop DNA in the presence of ATP was two-fold weaker in case of D591H and three-fold weaker in the case of H329D/D591H, suggesting that two domains work synergistically to bind stem-loop DNA. The weaker DNA interaction in case of double mutant H329D/D591H was also supported by alteration in conformation of the protein as compared to the wild-type ADAAD protein. Although the single mutant D591H also showed a change in protein conformation, the double-mutant H329/D591H showed a dramatically altered

conformation and this was also supported by dichroweb analysis. This study suggests that salt bridge is probably not formed in case of ADAAD between motif VI and motif II; however, the weakened interaction with stem-loop DNA in case of the double mutant H329D/D591H as compared to D591H mutant signifies the importance of coordination of both RecA-like domains of ADAAD in the interaction with stem-loop DNA in the presence of ATP.

To elucidate the inter-domain interaction between motif Ia and motif VI, a double mutant W277A/R595H was created. In a previous study reported by Dr. Meghna Gupta in the laboratory showed that W277A binds stem-loop DNA with a two-fold lower affinity in the absence of ATP (Gupta 2015). As R595 of motif VI is also required for stem-loop DNA binding, I postulated a greater interference in stem-loop DNA binding in the case of double mutant W277A/R595H. Molecular modeling and simulation performed with W277A and R595H showed that the RMSD value was different as compared to the wild-type ADAAD suggesting a defect in dynamic stability of the mutants. However, to my surprise, this defect in RMSD value was restored back to the level of wild-type in the double mutant W277A/R595H, suggesting stem-loop DNA binding in case of W277A/R595H would be like that of the wild-type ADAAD protein. This result was supported by stem-loop DNA binding studies where the affinity for stem-loop DNA in the absence of ATP in case of the double mutant W277A/R595H was similar to that of the wild-type ADAAD protein. However, the interaction with stem-loop DNA in the presence of ATP was five-fold weaker in case of the double mutant W277A/R595H in contrast to the three-fold weak binding observed in case of R595H as compared to wild-type ADAAD. This data again suggests that both the RecA-like domains work in a coordinated manner to bind stem-loop DNA in the presence of ATP. Further, like H329D/D591H, the double mutant W277A/R595H also shows a dramatic alteration in the conformation of protein with respect to wild-type ADAAD.

In the DEAD box subfamily member Ded1, the phenylalanine amino acid residue present in motif IV has been shown to communicate with first arginine of motif VI of the protein via van der Waal stacking interactions (Banroques et al., 2008). It has been proposed that this communication is needed for the interaction with the ligands. I studied these intra-domain interactions in ADAAD by creating a double mutant protein where phenylalanine (F507) of motif IV and first arginine (R592) of motif VI both were changed to alanine. Both the single mutants-F507A (work done by Mr. Vijendra Arya in the laboratory and hence, not reported in this thesis) and R592A and the double mutant F507A/R592A were unable to hydrolyze ATP. Further, the CD structure showed that the conformation in case of the double mutant F507A/R592A was totally disrupted with almost 0% alpha helical structure suggesting the importance of both motif V and motif VI, acting together, in maintaining the conformation of the protein. However, the interaction of the double mutant F507A/R592A with stem-loop DNA in both absence and presence of ATP was observed to be similar to that of R592-stem-loop DNA interaction suggesting that these two amino acids do not behave synergistically for the interaction with stem-loop DNA. Thus, the two amino acids of motif IV and motif VI are required only for maintaining the conformation of the protein.

Taken all these studies together, we can say that a communication between both the RecA-like domains is crucial for stem-loop DNA binding in the presence of ATP. Further, our results also suggest that an intra-domain interaction within RecA-like domain 2A is significant for maintaining the conformational integrity of the protein.

5: Elucidating the mechanism of

c-*MYC* transcription regulation by SMARCAL1

INTRODUCTION

Cells have evolved numerous mechanisms to manipulate and tightly package DNA into chromatin. The compact chromatin structure provides an efficient mechanism to regulate transcription of a cohort of genes inside cell. Two classes of protein complexes have been identified which coordinate to fulfill the need of transcriptional regulation: chromatin remodelers and histone modifiers (Strahl and Allis, 2000; Sudarsanam and Winston, 2000). Histone modifying enzymes recognize and covalently mark the residues in histone tails by phosphorylation, acetylation, methylation, sumoylation and ubiquitination and hence, maintain the remarkable plasticity required for transcriptional activation and repression of various genes (Strahl and Allis, 2000). Chromatin remodelers perform a different action for transcription regulation as they harness the energy released from ATP hydrolysis to unwrap, disposition or evict the nucleosomes from their positions, and consequently recruit the transcriptional activators/repressors at these sites (Sudarsanam and Winston, 2000).

The SWI/SNF family of ATP-dependent chromatin remodeler proteins were first identified in *Saccharomyces cerevisae* after the mutation introduced in the gene encoding SWI/SNF protein led to altered expression of various genes (Neigeborn and Carlson, 1984; Peterson and Herskowitz, 1992; Stern et al., 1984; Winston and Carlson, 1992). Mutations in SWI and SNF genes results in the development of pleiotropic defects, suggestive of a function of SWI/SNF proteins in gene expression regulation (Kokavec et al., 2008). Sudarsanam *et al*. showed that the SWI/SNF proteins mediate transcriptional regulation of around 6% of genes in *Saccharomyces cerevisae* (Sudarsanam and Winston, 2000). The SWI/SNF proteins individually control the expression of many genes at the promoter level rather than full chromosomal domains, indicating

121

their increased specificity for transcription regulation (Sudarsanam and Winston, 2000). DNA-binding transcription activators recruit SWI/SNF proteins to the specific promoter sequence. For instance, Swi5, a transcriptional activator, recruits SWI/SNF proteins at the HO promoter site in *Saccharomyces cerevisae* (Cosma et al., 1999; Mitra et al., 2006). In addition, activation domains of proteins like Gcn4, Hap4 and VP16 have also shown association with the SWI/SNF proteins to recruit them at the specific promoter sites (Natarajan et al., 1999). Consistent with the results obtained in *Saccharomyces cerevisae*, mammalian SWI/SNF protein complexes have also been shown to be targeted to the promoter by transcription activators, suggesting it to be a conserved feature across all eukaryotes (Fryer and Archer, 1998; Lee et al., 1999). Another method of SWI/SNF targeting is through their interaction with general transcription factors in both yeast and humans (Cho et al., 1998; Wilson et al., 1996). The association of the SWI/SNF proteins with transcriptional activators augments the related gene's transcription in two possible ways:

- The SWI/SNF proteins help in remodeling the nucleosome at sites where activators bind with weak affinity, resulting in tighter affinity of activator proteins towards those sites (Burns and Peterson, 1997).

- Association of the SWI/SNF proteins with specific activators helps in recruitment of additional activators/transcription factors. For instance, SWI/SNF proteins help in additional recruitment of TATA binding protein (TBP) at the TATA box containing nucleosomal site (Gregory et al., 1999; Ryan et al., 1998).

Some genes require continuous SWI/SNF assistance for gene expression. For instance, HO transcription requires continuous presence of SWI/SNF proteins in conjunction with other

activators possessing acetyl transferase activity, such as SAGA (Spt-Ada-Gcn5 acetyltransferase) complex (Cosma et al., 1999; Krebs et al., 1999). In fact, SWI/SNF proteins are also required at HO promoter to inhibit the repression effect of Sin3 and Rpd3 deacetylase protein complexes (Krebs et al., 1999; Sternberg et al., 1987).

The function of SWI/SNF proteins is not limited to transcriptional activation. Mutational studies have shown that the SWI/SNF proteins also repress transcription of various genes (Murphy et al., 1999; Trouche et al., 1997). Two main mechanisms have been established for transcription repression by SWI/SNF proteins:

- SWI/SNF proteins remodel the nucleosomes such that heterochromatin-type structure is formed leading to transcriptional repression by favoring the inactive nucleosomal stage at promoters (Schnitzler et al., 1998; Sudarsanam and Winston, 2000).

- The SWI/SNF proteins repress transcription with their special ability of histone modification, independent of their nucleosomal remodeling function, at the promoter sites. For example, SWI/SNF proteins are identified with histone deacetylase activity to perform centromeric silencing, DNA methylation and transcription inhibition in *Drosophila* (Bird and Wolffe, 1999; Brown et al., 1997; Kehle et al., 1998; Knoepfler and Eisenman, 1999).

Substantial evidence suggests that chromatin remodelers and histone modifiers like, histone deacetylases might work in a coordinated manner for the transcription inhibition of many genes (Narlikar et al., 2002). The ATP-dependent chromatin remodelers have also been assigned for the regulation of transcription factors like c-*myc* and c-*kit* by altering the DNA structure at promoter regions (Baradaran-Heravi et al., 2012).

The DNA is present in B-form almost all of the time and accounts for most of the biophysical processes undergoing in the cell (Cluzel et al., 1996). However, some sequences favor formation of structures like G-quadraplexes, Z-DNA, cruciform DNA and DNA hairpins (Mirkin, 2008; Zhao et al., 2010). The in vitro formation of G quadruplexes has been confirmed by mutation introduced in these G quadruplexes replacing gunanine to alanine that increased the c-*myc* transcription three fold. The in vivo formation of G quadruplexes has also been confirmed by Tmpy4 that stabilizes the G-quadraplexes and hence decreased the transcription of c-*myc* at a significant level (Siddiqui-Jain et al., 2002). Hairpin structures of DNA are formed at the metabolic sites like DNA replication, transcription and DNA repair where single-stranded DNA emerges and folds back on itself (Leach, 1994; Siddiqui-Jain et al., 2002; Voineagu et al., 2008). These hairpin forms can be further stabilized and modulated by proteins involved in these processes. SMARCAL1, for example, has been postulated to stabilize/modulate these secondary structures at promoter sites, thereby regulating gene expression (Baradaran-Heravi et al., 2012).

c-MYC is a transcription factor involved in various biological processes like cell proliferation, growth and apoptosis with high expression level found in tumor cells (Aasland et al., 1995; Evan and Littlewood, 1993; Evan et al., 1992; Marcu et al., 1992). C-*myc* expression is shown to have an inverse relationship with cellular differentiation (Evan et al., 1992). Transcriptional regulation of c-*myc* is a complex phenomenon involving many transcription start sites from various promoters (Levens, 2008). C-*myc* is transcribed mainly from promoters, P1 and P2 and was one of the first genes reported to possess poised RNA polymerase II (Levens, 2008). Around 85-90% of c-*myc* transcriptional activation is controlled by nuclease hypersensitivity element (NHE)III$_1$, localized upstream of the P1 promoter of c-*myc* (González and Hurley, 2010). The c-*myc* promoter contains G-quadraplex structures that are shown to be present in-vivo and play a

critical role in transcriptional regulation (Siddiqui-Jain et al., 2002; Yang and Hurley, 2006). A single nucleotide transition from G to A in the G-quadruplex forming sequence increases c-*myc* transcription (Grand et al., 2004). In contrast, stabilization of the G-quadraplex secondary structure by a small molecule TMPyP4 leads to transcription repression (Qin et al., 2007). These results provide a basis for transcription regulation at c-*myc* promoter sites, with the help of secondary structures like G-quadraplexes.

The role of ATP-dependent chromatin remodeling proteins in c-*myc* transcription has also been studied. BRG1, a chromatin remodeler, has been found to unmask the nucleosomal region at FUSE region, upstream site of P2 promoter of c-*myc* leading to transcription activation (Chi et al., 2003; Liu et al., 2006). The incorporation of histone variant H2A.Z has also been reported to inhibit the transcription of c-*myc* and this process might require an ortholog of yeast chromatin remodeling protein complex SWR1 (Farris et al., 2005).

SMARCAL1 has been postulated to alter the promoter region of genes like c-*myc* and c-*kit* in order to regulate their transcription process (Baradaran-Heravi et al., 2012). SMARCAL1 has also been reported to alter the expression of genes involved in development of zebra fish like *gata1* and *beta-E1 globin* (Huang et al., 2010). Downregulation of *smarcal1* causes cell cycle arrest at G0/G1 phase suggesting that SMARCAL1 is important for cell cycle regulation (Figure 5a) (Huang et al., 2010). This data has been supported by decreased expression of cyclinE2, G1/S phase transition protein, and increased expression of CDK inhibitor A (p21), an inhibitor of G1/S phase transition, in SMARCAL1 downregulated cells (Figure 5a) (Huang et al., 2010). Moreover, downregulation of *smarcal1* leads to altered gene expression profile in mice suggesting its role as a transcriptional regulator (Baradaran-Heravi et al., 2012). All these data,

125

hence, indicates the function of SMARCAL1 in transcription regulation of cell- cycle associated genes.
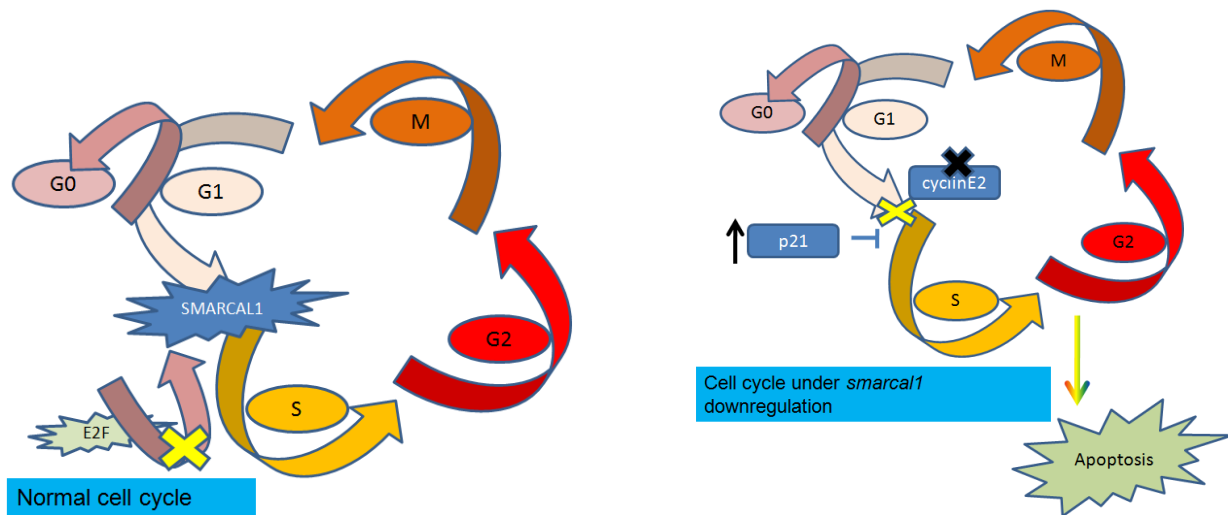
Figure: 5a Model depicting the cell cycle regulation under normal conditions and in *smarcal1* downregulation. (a) During normal cell cycle, SMARCAL1 plays role in G1/S phase transition. (b) Expression of certain cell cycle regulatory genes (indicated in blue) have been shown to be altered during *smarcal1* downregulation. Altered gene expression profile leads to G0/G1 cell cycle arrest, resulting in apoptosis (Huang et al., 2010).

Hypothesis and objectives:

SMARCAL1 has been shown to hydrolyze ATP preferentially in the presence of double-strand to single-strand transition regions like stem-loop DNA (Muthuswami et al., 2000). Baradaran-Heravi have hypothesized that SMARCAL1 might be regulating transcription by altering the conformation of promoter regions present in genes such as c-*myc* (Baradaran-Heravi et al., 2012). Therefore, in this chapter I have investigated whether SMARCAL1 can alter the structure of c-*myc* promoter and thereby, regulate its transcription.

To test the hypothesis, I have designed the following experiments:

- Protein occupancy assay to determine the presence of SMARCAL1 and other proteins on the c-*myc* promoter.

- Using QGRS mapper and Mfold software to identify the secondary structures in c-*myc* promoter.

- Purification of ADAAD followed by ATPase assays to determine whether the c-*myc* promoter DNA is an effector of the protein.

- Analysis of conformational modulation in c-*myc* promoter in the absence and presence of ADAAD protein using CD spectroscopy.

- Analysis of conformational modulation in c-*myc* promoter DNA by the ATPase dead mutant, K241A, using CD spectroscopy.

- Calculation of dissociation constants for c-*myc* promoter DNA binding by both ADAAD and K241A using fluorescence spectroscopy.

RESULTS

The occupancy of SMARCAL1 increases on the c-*myc* promoter during serum starvation

Previous studies in lab have shown that downregulation of *smarcal1*, using shRNA, leads to decreased expression of both *BRG1* and c-*myc,* suggesting a direct/indirect role of SMARCAL1 in transcriptional regulation of these genes (Sharma 2016). Further, Dr. Sharma showed that SMARCAL1 directly regulates c-*myc* transcription by serum starving the cells. Serum insufficiency leads to reduction in basal activity of the cells leading to withdrawal of cells from proliferation phase, thus arresting them in quiescent G0 phase (Khammanit et al., 2008). The cell undergoing G0 arrest shows decreased expression of c-*myc* (Bennett et al., 1994; Hoffman and Liebermann, 2008). Therefore, this experimental system could be used to study the occupancy of SMARCAL1, BRG1 and RNAPII on c-*myc* promoter through ChIP assay in unstarved and serum starved cells. The ChIP experiment revealed that the occupancy of SMARCAL1 and RNAPII increased on the *c-myc* promoter upon serum starvation. However, when serum starved cells were again supplemented with serum, SMARCAL1 occupancy decreased on c-*myc* promoter suggesting that the protein negatively regulates the expression of *c-myc* (Sharma 2016).

Based on the above information I first performed protein occupancy assay in *smarcal1* downregulated cells to gain insight of protein occupancy at the c-*myc* promoter when SMARCAL1 is absent. Therefore, as explained in materials and methods, the chromatin isolated from the control and *smarcal1* downregulated cells was digested using Micrococcal nuclease (Mnase) enzyme that digests DNA that is not tightly bound by proteins. *smarcal1* downregulated cells were constructed using shRNA against *smarcal1*. The protein bound DNA, which is of size approximately 150 bp, was extracted and purified. Then, this purified DNA was amplified using

the specific overlapping primers, generated against c-*myc* promoter, by quantitative real-time RT-PCR. The amplification of the c-*myc* promoter sites was compared between the control and the *smarcal1* downregulated cells. The highest protein occupancy was found in the region amplified by primer pairs I and II in control HeLa cells (Figure 5.1b). However, in the *smarcal1* downregulated cells this region showed decreased protein occupancy (Figure 5.1b). In contrast to the above results, the DNA regions amplified by primer pairs III, IV, V and VI showed no protein occupancy in both control and *smarcal1* downregulated cells (Figure 5.1b). Comparing the results from this assay with the ChIP experiment led us to conclude that BRG1 and RNAPII was present on c-*myc* promoter in control cells but in the *smarcal1* downregulated cells this region was completely free of proteins. This is not surprising as *brg1* is also downregulated in *smarcal1* downregulated cells (Haokip et al. 2016).

Next, I performed the same protein occupancy assay in unstarved cells, serum starved cells and cells released from serum starvation condition. Once again I found protein occupancy in the region amplified by primer pairs I and II during unstarved conditions. The occupancy in these regions decreased in cells starved of serum. Addition of serum back to serum starved cell caused no change in protein occupancy on primer pair I but showed increased occupancy on primer pair II. The primer pairs III, IV, V and VI were entirely free of proteins in unstarved condition, but the protein occupancy increased in primer pair region IV and V during serum starvation. Re-introduction of serum showed a further increase in protein occupancy in the primer pair regions III, IV, V and VI (Figure 5a). Once again comparing this data with ChIP result enabled us to determine the location of SMARCAL1, BRG1 and RNAPII on c-*myc* promoter in unstarved, serum starved, and released conditions. BRG1 and RNAPII were found to be present on -653 to -506 bp region upstream of P2 promoter of c-*myc* during unstarved condition. Further, under this

condition, SMARCAL1 showed no occupancy on the promoter in this region. On serum starvation, SMARCAL1 gets recruited on to the c-*myc* promoter and occupies the region from -653 to -299 bp. Reintroduction of serum in serum starved cells showed reappearance of BRG1 and RNAPII at -653 to -299 position while SMARCAL1 was present only at -523 to -299 bp region.
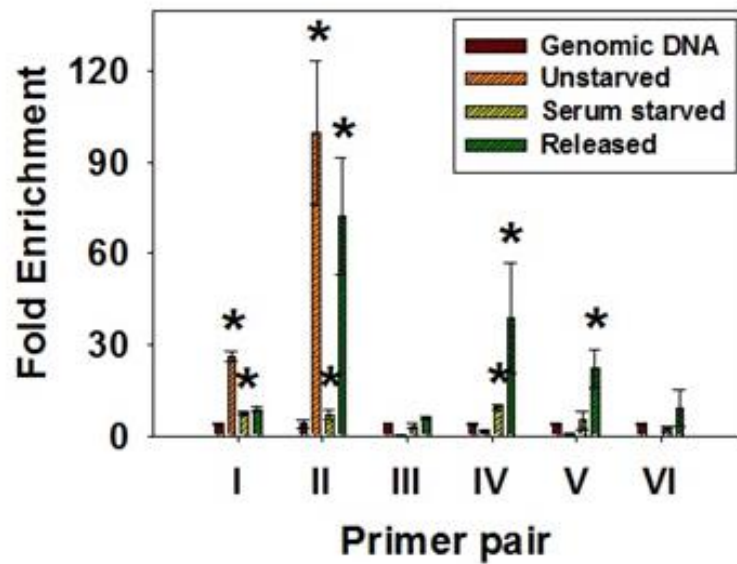
(a)
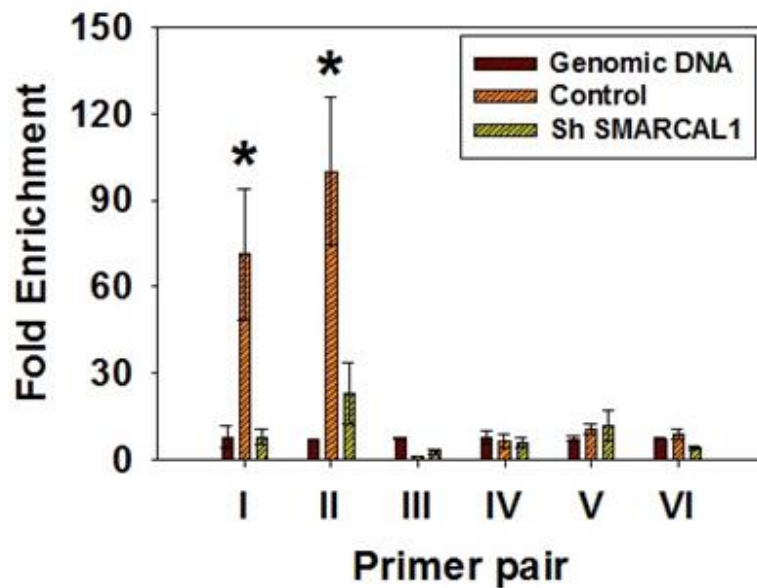
(b)

Figure: 5.1 Quantitative real-time RT-PCR analysis of protein occupancy on the c-*myc* promoter (primer I to primer VI). (a) Protein occupancy was compared in serum starvation experiment during unstarved, starved and pre-released conditions. (b) Occupancy of proteins was compared in control and *smarcal1* downregulated cells.

The c-*myc* promoter acts as an effector for ATPase activity of ADAAD.

Earlier studies conducted in our lab have shown the ability of ADAAD to hydrolyze ATP in presence of DNA molecules containing double-strand to single-strand transition regions such as stem-loop DNA (Figure 5.2d). Therefore, I was inquisitive to know whether the c-*myc* promoter region can also act as an effector molecule for ADAAD. Based on the ChIP analysis, Dr. Tapan Sharma identified that SMARCAL1 binds to three specific regions on the c-*myc* promoter-primer

pair regions A, B and C during serum starvation. I used Mfold software

(http://unafold.rna.albany.edu/?q=mfold/DNA-Folding-Form) to check the secondary structure

formation in all these primer pair regions and found that B region of c-*myc* promoter DNA

(herewith termed as Myc_B$_{159}$) showed the highest probability to fold into stem-loops (Figure

5.2a, 5.2b and 5.2c) (Sharma et al., 2015). As c-*myc* promoter is also known to contain many G-

quadraplex structures, I wanted to ask whether the promoter region Myc_B$_{159}$ also contains these

secondary structures. The QGRS software (Kikin et al., 2006)

(http://bioinformatics.ramapo.edu/QGRS/analyze.php) predicted a G-quadruplex in the

Myc_B$_{159}$ region. Consequently, I syntheized a 40 bp DNA from Myc_B$_{159}$ region termed as

G$_E$C$_E$ that was capable of forming putative G-qudruplex as well as stem-loop structures.

Next, I performed ATPase assay with G$_E$C$_E$ DNA and found that it could effect ~ 50% of the

ATPase activity as compared to the control stem-loop DNA (Figure 5.3b). To check whether

single-stranded DNA is sufficient to act as effector for ATPase activity of ADAAD, I used G$_E$

DNA that is one single-stranded strand of G$_E$C$_E$ DNA. G$_E$ DNA was found to be a poor effector

of ATPase activity as compared to G$_E$C$_E$ DNA (Figure 5.3a). Also, the well-known CT element

of c-*myc* promoter, termed as G$_P$ DNA, that forms an unstable hairpin kind of structure in Mfold

program was used as a negative control in this experiment and indeed it was a poor effector as

compared to G$_E$C$_E$ DNA (Figure 5.3a). As we wanted to know whether ATPase activity of c-*myc*

promoter region was due to the formation of stem-loop like structure, I heated the DNA and

cooled it rapidly. As the formation of stem-loop is an intramolecular event, fast cooling after heat

denaturation promotes this secondary structure formation. The ATPase activity of ADAAD with

G$_E$C$_E$ region was found to be higher when the DNA was heated and cooled rapidly as compared

to the slow-cooled DNA suggesting the formation of stem-loop DNA structure in c-*myc*

promoter region (Figure 5.3b). Finally, I calculated the K_M values of ADAAD with G_EC_E to be

11.8 ± 2 nM, about 10-fold higher than that for ADAAD-stem-loop DNA (Figure 5.3c).

(a)   (b)   (c)   (d)
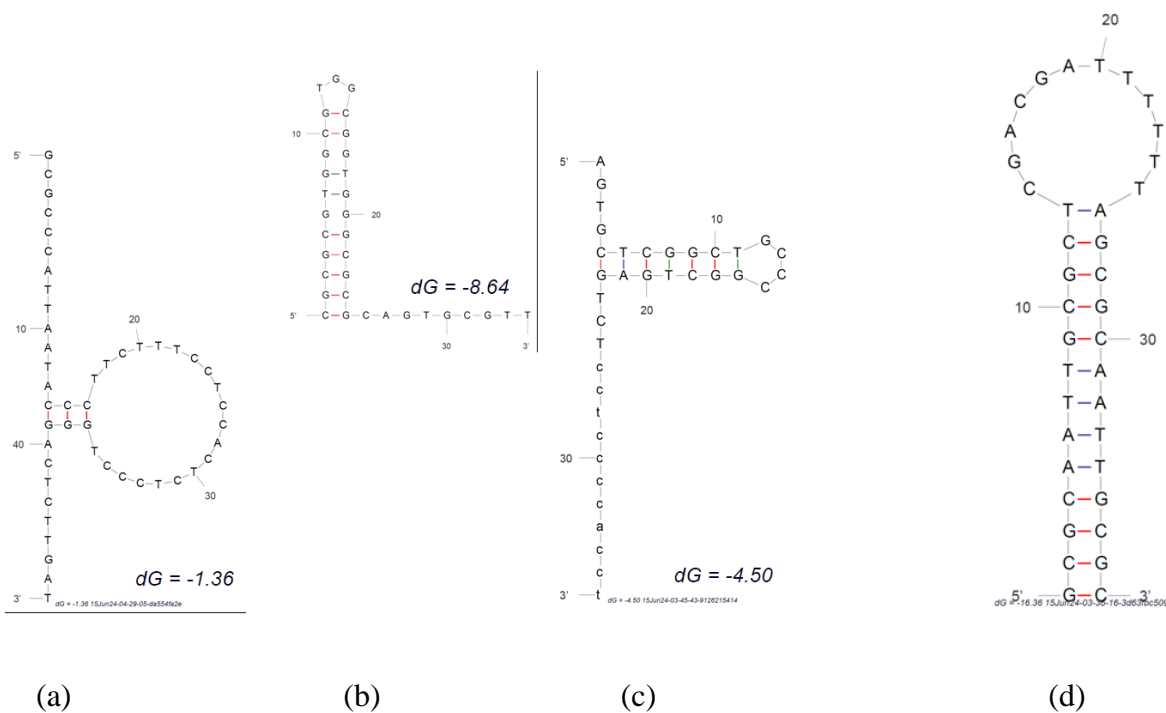
Figure: 5.2 (a), (b), (c) Mfold analysis was performed to identify the secondary structure

formation in all the regions of primer A, B and C respectively. Gibbs free energy for the

formation of each of the secondary structure is also shown. Here (b) denotes the G_EC_E region of

primer B and, (d) represents the structure of the control stem-loop DNA used in our studies.
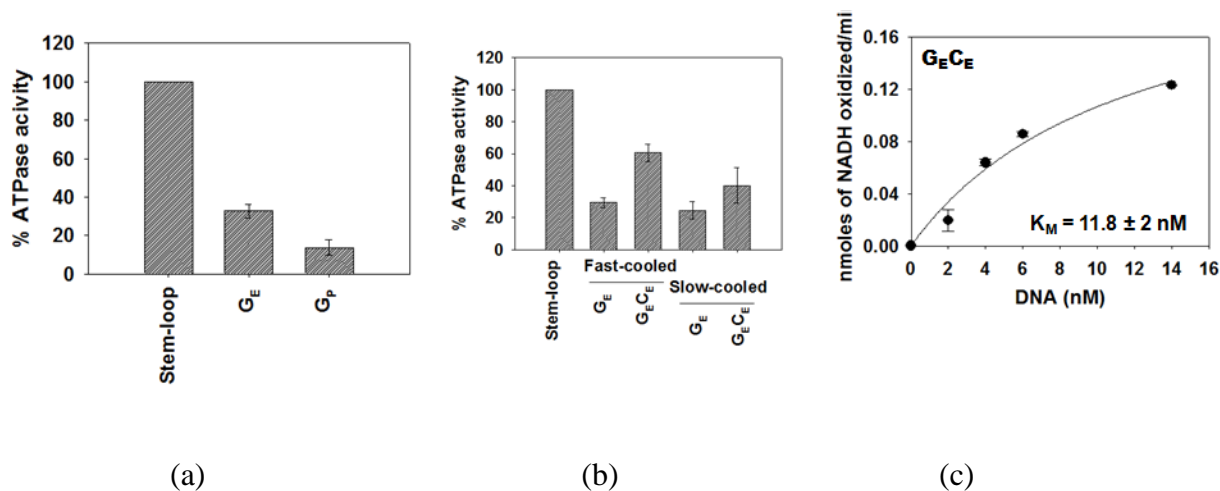
Figure: 5.3 (a) Comparison of ATPase activity elicited with G_E, G_EC_E and stem-loop DNA. (b) The ATPase activity was also compared for fast heat-cooled and slow heat-cooled G_EC_E DNA with stem-loop DNA. (c) Michaelis-Menten constant calculation for G_EC_E–ADAAD interaction.

ADAAD and K241A mutant binds to G_EC_E with affinity similar to that of stem-loop DNA

ADAAD binds to stem-loop DNA with a dissociation constant of 3.4 ± 0.2 nM in the presence of saturating concentration of ATP. Therefore, I wanted to determine whether ADAAD binds to G_EC_E DNA also with same affinity, and hence I performed fluorescence binding experiment with G_EC_E DNA. The dissociation constant for G_EC_E DNA binding with ADAAD in the presence of ATP was found to be 8.3 ± 3 nM (Figure 5.4a), only two-fold less affinity as compared to ADAAD-stem-loop DNA interaction ($K_d = 3.4 \pm 0.2$ nM). Further, to know that this binding of G_EC_E DNA with ADAAD was ATPase dependent, I repeated the experiment with K241A which is an ATPase inactive mutant of ADAAD (Gupta 2015). Lysine 241 of ADAAD plays a crucial

134

role in maintaing the ATP hydrolysis activity of ADAAD. ATPase assays have shown that substitution of this lysine with alanine (K241A) results in loss of ATPase activity of ADAAD. However, the interaction of the mutant protein K241A with the ATP in presence of G_EC_E DNA was unimpaired (K_d= 0.19 ± 0.01 µM) with respect to wild-type ADAAD-ATP interaction in presence of G_EC_E DNA (K_d= 0.22 ± 0.04 µM). Moreover, the interaction of the mutant protein with the G_EC_E DNA in presence of ATP was similar to ADAAD-G_EC_E DNA and thus, two-fold less affinity as (K_d= 7.3 ± 2.4 nM) compared to wild-type ADAAD-stem-loop DNA interaction (K_d= 3.4 ± 0.2 nM). Also, my studies with G_EC_E DNA showed that the interaction between DNA and K241A mutant protein (K_d= 7.3 ± 2.4 nM) was similar to the wild-type ADAAD-G_EC_E DNA interaction (K_d= 8.77 ± 3.2 nM) in the presence of ATP. All these results suggest that the ATPase activity of ADAAD was not necessary to bind with the c-*myc* promoter DNA and is approximately same as the stem-loop binding (Figure 5.4a and b) (Table 5.1).
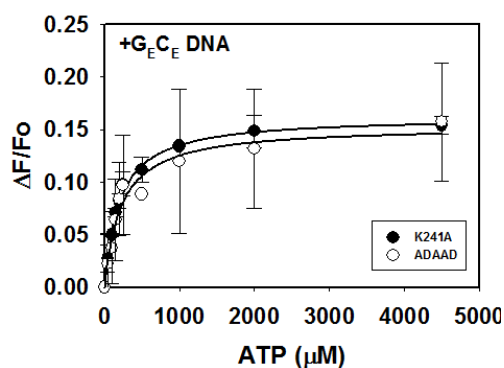
(a)
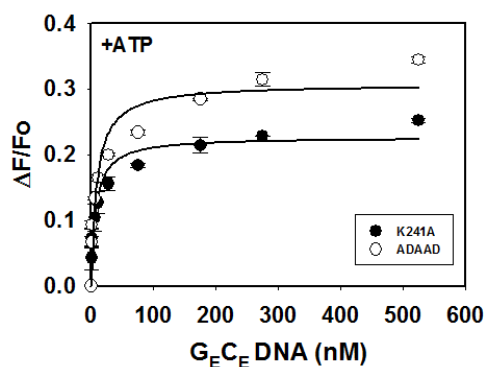
(b)

Figure: 5.4 (a) K_d values for ATP binding were calculated both for ADAAD and K241A in the presence of G_EC_E DNA(2µM). (b) The K_d values were calculated for G_EC_E DNA binding in the presence of saturated concentration of ATP(20µM).

The conformational change induced in c-*myc* promoter DNA by ADAAD is similar to the conformational change induced in stem-loop DNA

Next, I used CD spectroscopy to study the conformational changes induced in G_EC_E DNA by ADAAD in the presence of ATP. As predominant G-quadraplex structure formed on c-*myc* promoter gets stabilized in the presence of potassium ions, I performed CD studies both in the absence and presence of potassium ions. Classic G-quadraplex has been shown to possess a 260 nm positive peak and 240 nm negative peak; however, G_EC_E DNA was found to possess two positive peaks-one at 210 nm and other at 258 nm-indicating that ADAAD does not induce a G-quadraplex structure in G_EC_E DNA. As Mfold analysis as well as ATPase activity suggested the formation of stem-loop like structure in the G_EC_E region, I, therefore, studied the conformational change induced in the stem-loop DNA by ADAAD and ATP. The stem-loop DNA also showed two positive peaks-one at 215 nm and other at 250 nm in the presence of ADAAD and ATP, indicating the the G_EC_E DNA in the absence of potassium ions possibly adopts a stem-loop kind of structure (Figure 5.5a). Next, the structure of G_EC_E DNA as well as stem-loop DNA in the presence of potassium ions was studied using CD. Once again the CD spectra of G_EC_E DNA and stem-loop DNA, induced by ADAAD, were similar though, the spectra in the absence and presence of K^+ ions were different (Figure 5.5b).

From these studies I propose that the G_EC_E region of the c-*myc* promoter forms a stem-loop kind of secondary structure in the presence of ADAAD, and this stem-loop structure formation possibly helps in repressing the transcription of c-*myc*.

(a)

(b)
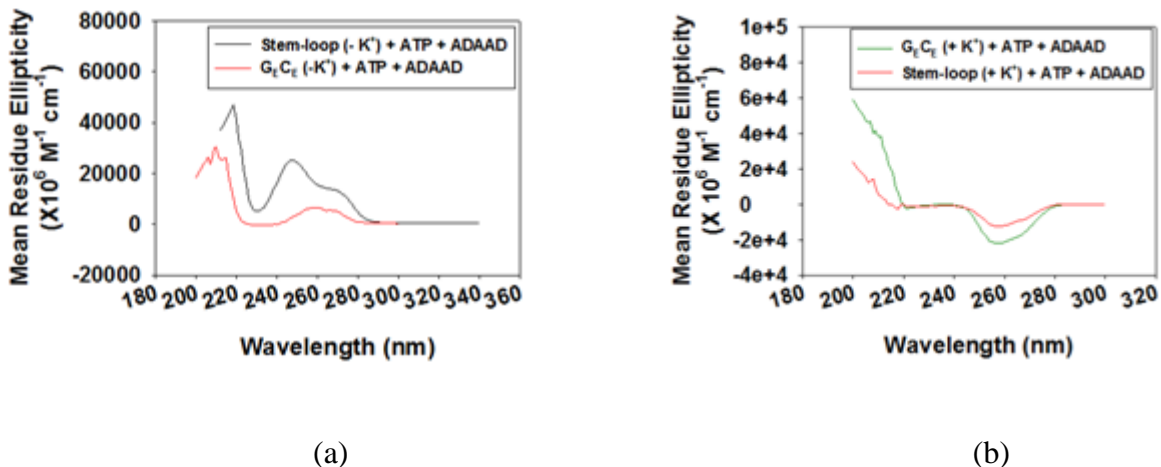
Figure: 5.5 Comparison of conformations of G_EC_E DNA and stem-loop DNA by ADAAD (a) in the absence of potassium ions and, (b) in the presence of potassium ions.

The conformational change induced in G_EC_E DNA by ADAAD is ATPase-dependent

Next, to understand whether the ATP hydrolysis activity of ADAAD is necessary to induce the conformational change in c-*myc* promoter, I executed CD studies of c-*myc* promoter G_EC_E DNA using the ATPase dead mutant of ADAAD. I used the K241A mutant to study the effect of ATPase activity of ADAAD in inducing a conformational change in G_EC_E DNA. I performed these experiments with GST-tagged ADAAD and GST-tagged mutant K241A. As stated earlier, I observed two positive peaks at 210 nm and 258 nm after the incubation of fast heat cooled

137

G_EC_E DNA with ADAAD and ATP (Figure 5.6a). However, when G_EC_E DNA was incubated with K241A I got a postive peak at 210 nm and a negative peak at 258 nm (Figure 5.6b). Moreover, I found that the CD spectra obtained after incubation with G_EC_E DNA was similar to the spectra obtained after stem-loop DNA incubation with K241A mutant. Incubation with EDTA, chelator of magnesium ions, which disrupts the ATPase activity of ADAAD, resulted in the formation of negative peak both at 210 nm and 260 nm (Figure 5.6c).
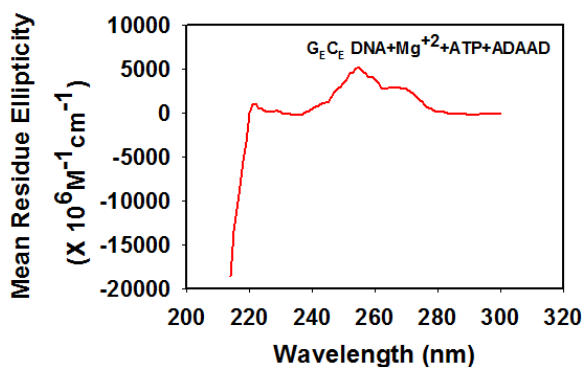
(a)
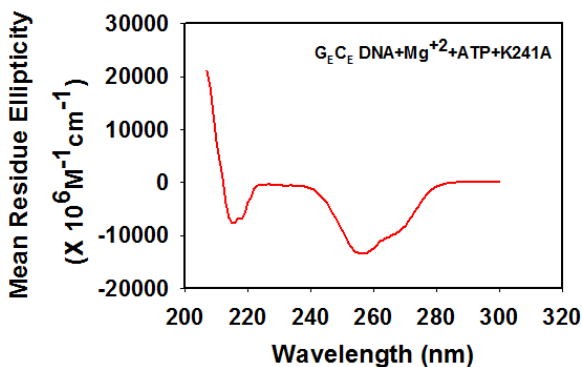
(b)
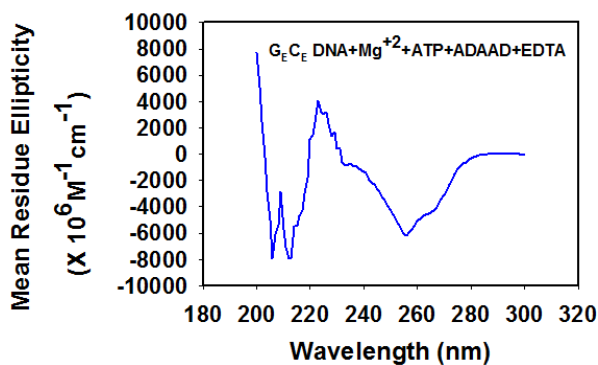
(c)

Figure: 5.6 (a) CD spectroscopy depicting the conformational change in G_EC_E DNA due to the presence of GST-ADAAD. (b) Conformational change in G_EC_E DNA after incubation with K241A mutant protein. (c) Conformational change in G_EC_E DNA after incubation with EDTA followed by the addition of GST-ADAAD.

Table 5.1: The K_d values were calculated for ligand-protein (wild-type ADAAD and K241A mutant) interactions by fluorescence spectroscopy and fitted to one-site saturation hyperbolic curve.

	ADAAD	K241A
ATP(+ G_EC_E DNA) (X 10^{-6}M)	(0.22 ± 0.04)	(0.19 ± 0.01)
G_EC_E DNA(+ATP) (X 10^{-9}M)	(8.3 ± 3)	(7.3 ± 2.4)

DISCUSSION

In this chapter I have shown the binding of SMARCAL1 to the c-*myc* promoter DNA that results in c-*myc* transcription inhibition.

Although downregulation of *smarcal1* has been shown to alter the expression of cell-cycle associated genes, no direct evidence has been reported indicating transcription regulation of these genes by SMARCAL1 (Huang et al., 2010). However, Baradaran-Heravi have postulated that SMARCAL1 may play a role in transcriptional regulation of oncogenes like c-*myc* and c-*kit* by introducing a conformational change in the promoter region (Baradaran-Heravi et al., 2012). Analysis of c-*myc* expression both in *smarcal1* downregulation and during serum starvation showed a decrease in its transcription, suggesting a direct or indirect role of SMARCAL1 in c-*myc* transcription regulation. Experiments that I have presented in this chapter confirm that SMARCAL1 directly regulates the expression of c-*myc*. However, the mode of reduction in c-*myc* expression was found to be completely different in these two conditions.

In our lab, SMARCAL1 and BRG1 have been shown to mutually co-regulate each other (Haokip et al. 2016). Further, BRG1 has already been reported in previous reports to positively regulate the transcription of c-*myc* (Chi et al., 2003). Thus, downregulation of *smarcal1* causes reduced expression of BRG1 and this possibly results in diminished level of c-*myc* expression in these cells.

In contrast, during serum starvation, the protein occupancy as well as ChIP assay has shown increased occupancy of SMARCAL1 on c-*myc* promoter and at this stage, BRG1 and RNAPII are absent. Reintroduction of the serum results in increased c-*myc* expression and at this point the

140

occupancy of BRG1 and RNAPII occupancy increases while that of SMARCAL1 decreases on the c-*myc* promoter. Although we have observed that SMARCAL1 associates with c-*myc* promoter, the mechanism by which SMARCAL1 gets recruited to the promoter site is still unknown. An association of SMARCAL1 with RPA during DNA damage repair raises the possibility that SMARCAL1 might be recruited via an interaction with the single-strand binding protein, RPA (Ciccia et al., 2009). Conjointly, these results prove that BRG1 act as a positive regulator while SMARCAL1 act as a negative regulator of c-*myc* transcription.

The importance of double-strand to single-strand transition regions of DNA to act as effector molecules during various DNA processing activities of SMARCAL1 like telomere maintaineance and DNA damage repair has already been reported (Bansbach et al., 2009; Ciccia et al., 2009; Cox et al., 2016; Poole et al., 2015). An analysis of the c-*myc* promoter using QGRS mapper and Mfold software identified both G-quadruplex structure as well as stem-loop structures within the 159 bp Myc_B_{159} DNA region. The sequence that was able to form both G-quadruplex and stem-loop DNA was termed as G_EC_E DNA. Analysis showed that fast cooling of G_EC_E DNA after heat denaturation is as good an effector of ADAAD as the control stem-loop DNA. To our surprise, slow cooling of the G_EC_E DNA after heat denaturation was not as effective as the fast cooling suggesting that the formation of secondary structures due to intramolecular interactions was an important criteria for a DNA molecule to be an effector of ADAAD. These differences in ATPase activity suggests that presence of secondary structures at c-*myc* promoter are necessary for ADAAD to regulate the c-*myc* transcription. Analysis of the kinetic parameters showed that the K_M for ADAAD-G_EC_E interaction was 11.8 ± 2 nM, which is about 10-fold higher than that for ADAAD-stem-loop interaction. Dr. Tapan Sharma in the laborotary determined the K_M for ADAAD-Myc_B159 region and found it to be 3-fold lower

than that for ADAAD-G_EC_E interaction sugesting that there might be some other regions also that are necessary for ADAAD-Myc_B_{159} interaction (Sharma 2016).

The SWI/SNF proteins have been shown to interact with the nucleosomes and DNA with high affinty in a structure-specific manner (Darst et al., 2001; Dürr et al., 2005; Nongkhlaw et al., 2009; Zofall et al., 2004). Further, these proteins have been reported to modulate the DNA structure and their topology for performing their nucleosome remodeling fiunction (Darst et al., 2001; Zofall et al., 2004). The CD analysis showed that like other SWI/SNF proteins, ADAAD also can modulate the conformation of G_EC_E DNA. This conformation was found to be similar to that of stem-loop DNA leading us to conclude that binding of SMARCAL1 to the c-*myc* promoter induces the formation of a secondary structure that is similar to stem-loop. From the above observations, it can be postulated that ADAAD induces a stem-loop like conformational change in c-*myc* promoter DNA necessary for ATP hydrolysis and transcription inhibition.

The ATP hydrolysis function of chromatin remodeling proteins like SWI/SNF and NuRF is essential for the alteration of nucleosomal structure (Côté et al., 1998; Workman and Kingston, 1998). The data presented in this chapter show that the ATPase activity of ADAAD, like that of SWI/SNF and NuRF, is crucial for regulating c-*myc* transcription. K241A, an ATPase dead mutant, was unable to induce the requisite conformational change in the G_EC_E DNA suggesting the importance of ATPase activity of ADAAD in modulating the formation of stem-loop like secondary structure in the G_EC_E region of c-*myc* promoter for transcription inhibition. However, the binding data showed that both ADAAD and K241A mutant have similar affinty for the G_EC_E DNA indicating that binding is not dependent on the ATPase activity of ADAAD. This data

suggests that the ATPase activity is not necessary for interaction with the c-*myc* promoter DNA , but is important for the conformation modulation during transcription regulation by ADAAD.

Based on all the above results, I have designed a model to explain the mechanism of transcription regulation of c-*myc* by SMARCAL1 (Figure 5b). In normal unstarved cells, the c-*myc* promoter is occupied by activator proteins BRG1, RNAPII and H3 histones resulting in normal expression of c-*myc*. On serum starvation, SMARCAL1 along with RPA and H3 histones gets recruited to the c-*myc* promoter and using its ATPase activty, modulates structure of the Myc_159 region by possibly inducing the formation of a stem-loop like structure thus, resulting in transcriptional inhibition of c-*myc* (Figure 5b). Reintroduction of serum results in removal of SMARCAL1 from the promoter and recruitment of BRG1 and RNAPII so that normal c-*myc* expression can be restored (Figure 5b).
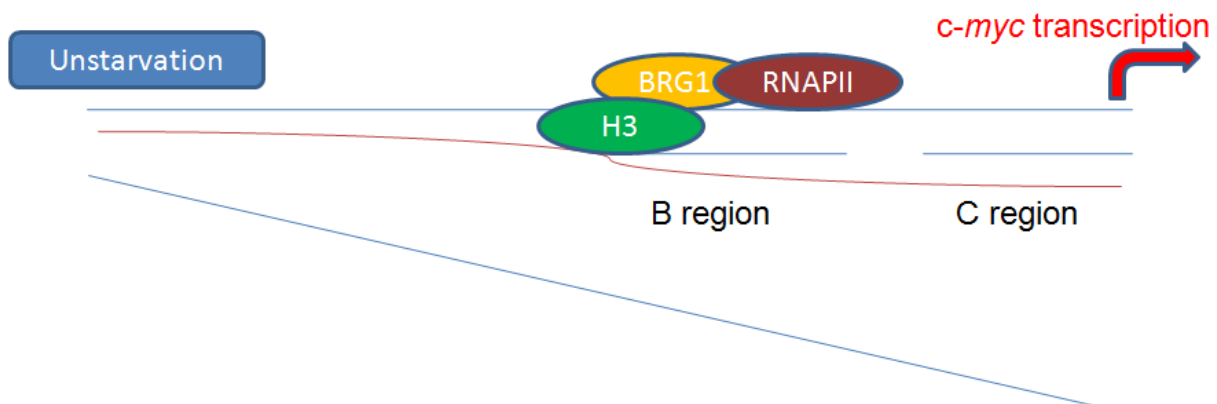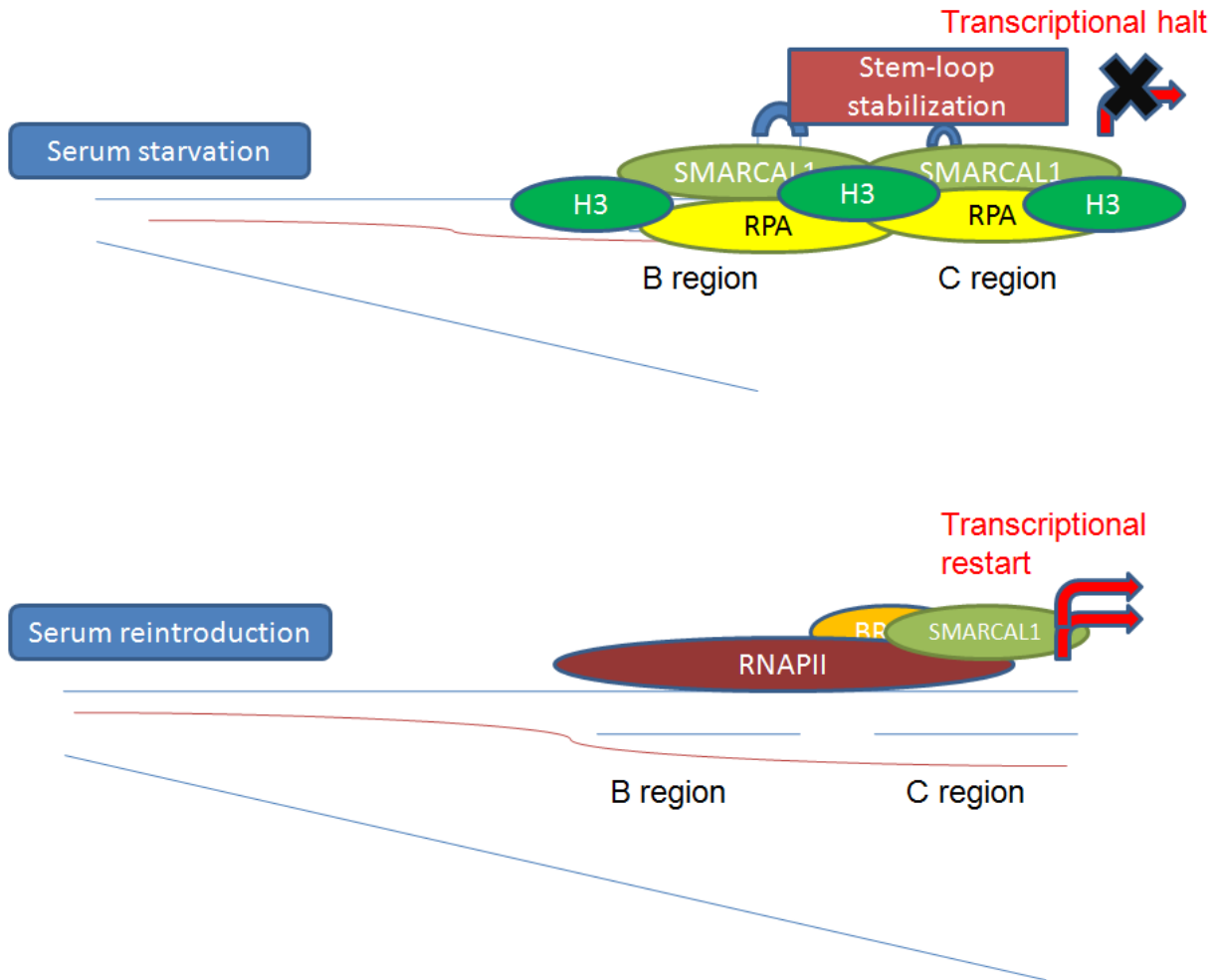
Figure: 5b Model depicting the c-*myc* transcription regulation during serum starvation assay. (a) During unstarved conditions, HeLa cells shows increased expression of c-*myc* with BRG1, RNAPII and H3 presence on the promoter region. (b) After serum starvation, SMARCAL1 along with RPA and H3 gets recruited on the c-*myc* promoter where it stabilizes the stem-loop structures and causes decrease in c-*myc* transcription. (c) Serum reintroduction again recruits BRG1 and RNAPII on the promoter resulting in increased c-*myc* transcription.

6: Effect of R820H, a mutation present in SIOD patients, on ligand binding, ATP hydrolysis and transcriptional mechanism of SMARCAL1

INTRODUCTION

Chromatin remodelers and disease manifestation

The SWI/SNF chromatin remodelers regulate a wide variety of mRNA transcript molecules and hence, mutation in even a single *swi/snf* gene causes altered phenotypes resulting in disease manifestation (Kokavec et al., 2008). Thus, mutations in chromatin remodelers are widely associated with developmental disorders. For example, mutations in CHD7 is strongly associated with CHARGE syndrome, mutations in SNF2H resultes in Williams syndrome, mutations in BRG1 is associated with rhabdoid tumors, and mutations in SMARCAL1 is linked to Schimke immuno-osseous dysplasia (SIOD) (Gibbons, 2008; Kokavec et al., 2008). In the following sections, I will be discussing the role of the chromatin remodelers in disease manifestation.

SNF2H and Williams Syndrome

SNF2H is member of the ISWI family with 87% homology to the related ISWI ATPase member, SNF2L (Aihara et al., 1998). Although SNF2H is expressed in almost all haematopoitic cells, the co-expression of SNF2H and SNF2L is seen only in testes, liver and brain (Lazzaro and Picketts, 2001). The two proteins form distinct complexes-ACF1 involving interaction of SNF2H with Acf1; and WICH complex involving SNF2L and WSTF interaction (Bozhenok et al., 2002; Fyodorov and Kadonaga, 2002). Further, it has been shown that SNF2L cannot substitute for SNF2H indicating distinct functions for the two proteins (Stopka and Skoultchi, 2003). SNF2H null mutants show apoptosis and cell proliferation arrest resulting in Williams syndrome (Meyer-Lindenberg et al., 2006). This syndrome is characterized by growth and mental retardation, learning diabilities, short-term memory, cardiac abnormalities and high empathy for other people

(Meyer-Lindenberg et al., 2006). Williams syndrome transcription factor(WSTF) was found to interact with SNF2H, forming WICH complex, to restrict the nucleosomal assembly activity of SNF2H changing the chromatin fluidity and thereby, causing Williams syndrome (Kokavec et al., 2008). The exact role of WICH complex in Williams syndrome needs to be further studied (Kokavec et al., 2008).

CHD7 and CHARGE Syndrome

CHD7, another chromatin remodeler, is widely expressed in brain, skeletal muscle and kidney (Bosman et al., 2005). Approximately ten CHD genes have been reported till now and the CHD proteins have been shown to possess two chromodomains, a DNA binding domain, and the characteristic SNF2 helicase domain (Delmas et al., 1993; Woodage et al., 1997). Haploinsufficiency of CHD7 causes CHARGE syndrome that is characterized by growth retardation and developmental defects, heart defects, ear anomalies, coloboma etc. (Jongmans et al., 2008; Vissers et al., 2004). As CHD7 is shown to regulate various functions like transcription and chromatin structure maintaineance, it is still mysterious which function of CHD7 is abrogated in CHARGE syndrome (Feng et al., 2013; Ho and Crabtree, 2010; Murawska and Brehm, 2011). Further, CHD7 gene variations also shown to cause idiopathic scoliosis type 3 in children (White et al., 2005).

BRG1 and Cancer

Significant findings that connect the role of SWI/SNF proteins with tumorigenesis came with the mutational and gene inactivation studies of BRG1 (Bultman et al., 2000; Wong et al., 2000). In many human cancerous cell lines like prostate, pancreas, breast, neuroendocrine, uterinecervix

and lung, both alleles of BRG1 are found to be mutated (Teng et al., 2001; Wong et al., 2000). Readdition of BRG1 to BRG1 deficient cell lines resulted in induction of normal G1 cell cycle arrest; however, this was mediated by co-expresssion of pRb gene (Wong et al., 2000). This study signifies that BRG1 and pRb work together to control the G1/S transition, probably by E2F transcription factor regulation (Wong et al., 2000). BRG1 has also been found to interact with BRCA1 in response to p53-mediated transcription (Bochar et al., 2000). Additionally, BRG1 is also found to be present at regulatory regions of various genes; for example ZNF185 and CYP3A4, in tumor samples (Medina et al., 2005). All these studies suggest that BRG1 regulate cell cycle progression, and a wide variey of genes and thus, mutation in this SWI/SNF protein can be a basis for tumor generation in various cancerous cell lines (Hendricks et al., 2004). Recently, BRG1 has been found to be mutated in Coffin-Siris syndrome (CSS) (Tsurusaki et al., 2012). Coffin-Siris syndrome is characterized by intelletual disabilities, abnormal facial features, growth retardation and hypoplastic nail of finger/toe (Tsurusaki et al., 2012). Other then BRG1, chromatin remodelers like-SMARCB1, SMARCA2(BRM), SMARCE1, ARID1A and ARID1B have also been found to be mutated in CSS (Tsurusaki et al., 2012).

SMARCAL1 and SIOD

SIOD is a disease characterized by disproportionate growth failure caused by spondyloepiphyseal dysplasia, renal failure caused by glomerulosclerosis, T-cell immunodeficiency and bone marrow failure (Boerkoel et al., 2000; Boerkoel et al., 2002). SIOD disease frequently leads to various other diseases like hypothyroidism, cerebral ischemia, dental

anomalies, autoimmune enteropathy, pulmonary hypertension, and various opportunistic infections (Boerkoel et al., 2000; Boerkoel et al., 2002; Elizondo et al., 2006; Kilic et al., 2005)

Based on the genome sequence of patients and positional candidate approach, SMARCAL1 was found to be the most likely gene causing SIOD, an autosomal recessive disorder (Boerkoel et al., 2002; Carroll et al., 2013; Deguchi et al., 2008; Elizondo et al., 2009; Santangelo et al., 2014). Analysis of patients with severe disease revealed the presence of frameshift or nonsense mutations in both the alleles of SMARCAL1, whereas patients with milder disease were identified to have a missense mutation in a single allele (Boerkoel et al., 2002). Downregulation of SMARCAL1 leads to increased p53-independent cell death suggesting that defect in SMARCAL1 might be causing developemental defects in SIOD patients (Boerkoel et al., 2002). However, the exact role of SMARCAL1 mutations in the development of SIOD pathogenesis needs to be elucidated.

Most of the biallelic mutations causing SIOD have been mapped to the conserved helicase motifs of SMARCAL1 (Boerkoel et al., 2002). Mutations in conserved arginine residue (R644W, R586W, R645C, R820H and R764Q) among *C. elegans*, *Homo sapiens*, *Mus musculus* and *Drosophila melanogaster* have been reported to be the most prevalent missense mutation (Boerkoel et al., 2002). Comparison with other SNF2 proteins showed that mutations A468P and I548N lie within the ATP binding domain (Boerkoel et al., 2002). The other mutations-S579L, T705I, R764Q and R820H identfied in SIOD patients were found to reside within a domain necessary for DNA binding, NTP hydrolysis and coupling between DNA binding and ATPase activity (Boerkoel et al., 2002). R644W, K647Q and K647T mutations were found to lie within the region necessary for targeting of protein to the nucleus (Boerkoel et al., 2002). In our lab, Dr.

Meghna Gupta has analyzed the biochemical function of three mutations-A245P, I325N and S355L present in SIOD patients (Gupta et al., 2015). These mutations result in weaker affinity for stem-loop DNA in the presence of ATP and hence, impaired conformational intgrity of the protein leading to defective ATPase activity of the protein (Gupta et al., 2015). These mutations also result in increased replication stress in the cells (Gupta et al., 2015). All these studies suggest that defect in ATPase activity and DNA damage repair might be causing the development of SIOD disease. However, the role of other mutations present in SIOD patients have not been studied in detail.

Hypothesis and objectives

I have presented experimental data in the previous chapter showing the importance of R595 in DNA binding. The corresponding residue of R595 in SMARCAL1 is R820 which is often mutated to histidine in SIOD patients (Boerkoel et al., 2002). Therefore, I hypothesized that R820H mutation in SMARCAL1 would impair conformational stability of the protein, DNA binding and therefore, ATP hydrolysis. Further, I hypothesized as ATP hydrolysis is needed both for transcription as well as DNA repair, therefore, R820H mutation would impair both DNA damage repair as well as transcriptional regulation.

To explore these studies, I designed the following experiments:

- To create R595H mutation in ADAAD and R820H mutation in SMARCAL1.

- To study the ATP binding in the absence and presence of stem-loop DNA in R595H mutant.

- To study the stem-loop DNA binding in the absence and presence of ATP in R595H mutant.

- To measure the change in conformational integrity induced by R595H mutant.

- To decipher the transcriptional regulation of *c-myc* and *dgcr8* by R595H mutant protein.

- To study the role of R820H mutation in DNA damage repair and its comparison with K464A, an ATPase dead protein.

RESULTS

Motif VI mutant protein R595H, present in SIOD patients, induces a conformation alteration and hence defective ATP hydrolysis

My studies conducted with R595 residue present in motif VI, thus far, have shown that this amino acid is important for interaction with stem-loop DNA in the presence of ATP. To carry this study forward, I wanted to study the biochemical function of R595H mutation present in motif VI that has been identified in SIOD patients. As with other R595 mutants, purified R595H (Figure 6.1a) protein too was found to be ATPase inactive (Figure 6.1b). As expected, binding studies showed that the ATP interaction in the absence and presence of stem-loop DNA was similar to that of the wild-type ADAAD protein (Figure 6.2a and b) (Table 6.1). Additionally, binding analysis confirmed that stem-loop DNA binding in the absence of ATP was similar to wild-type ADAAD–stem-loop DNA interaction (Figure 6.2c) (Table 6.1). However, I observed a 3-fold weaker affinity for stem-loop DNA in the presence of ATP in case of R595H mutant as compared to the wild-type ADAAD protein (Figure 6.1d) (Table 6.1). Secondary structure studies conducted by CD showed that the conformation was altered in the mutant protein as compared to the wild-type ADAAD protein both in the absence and presence of ligands (Figure 6.3a and b). Dichroweb analysis also confirmed that R595H protein has a distorted protein conformation both in the absence and presence of ligands as compared to wild-type ADAAD protein (Table 6.2).
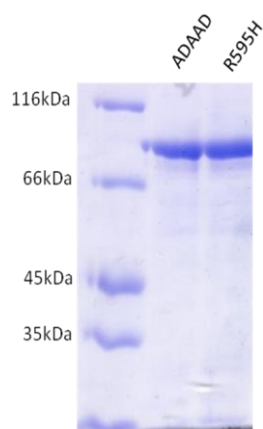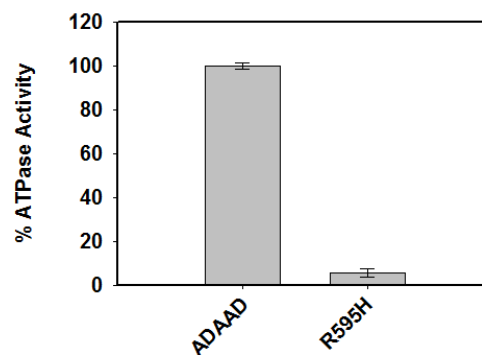
(a)   (b)

Figure: 6.1 (a) Purification of the motif VI mutant protein R595H, present in SIOD patients, on

10% SDS-polyacrylamide gel and its staining using Coomassie brilliant blue. (b) ATPase

activity of the mutant protein R595H in the presence of stem-loop DNA was compared to the

wild-type ADAAD protein. Each analysis represents the average ± standard deviation of two

independent experiments. The protein concentration used in these analyses was 0.1 µM.
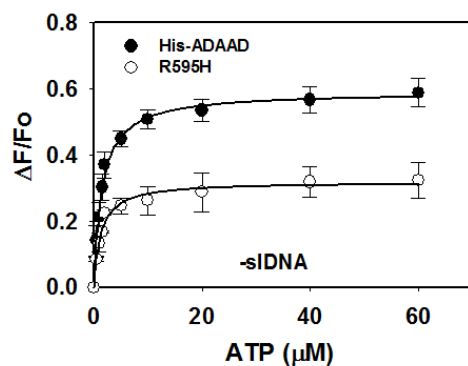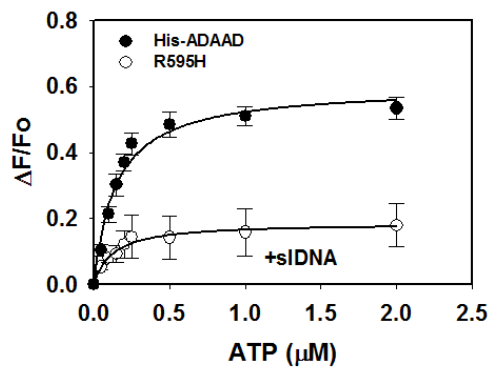
(a)   (b)
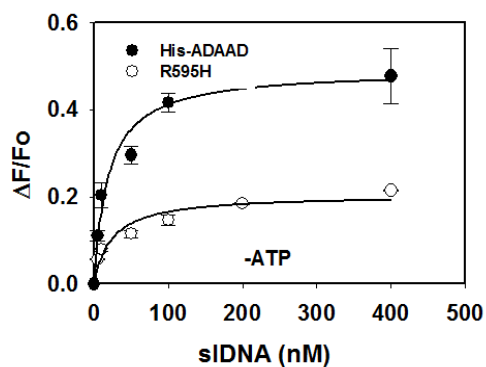
(c)
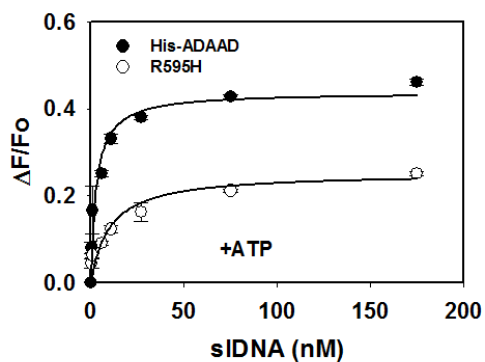
(d)

Figure: 6.2 Comparison of ligand binding parameters for R595H with wild-type ADAAD protein. The binding curves fitted to one-site saturation model for the interaction of R595H with (a) ATP in the absence of stem-loop DNA, (b) ATP in the presence of stem-loop DNA, (c) stem-loop DNA in the absence of ATP and, (d) stem-loop DNA in the presence of ATP. Each analysis represents the average ± standard deviation of two independent experiments. The protein concentration used in these analyses was 0.5 µM.
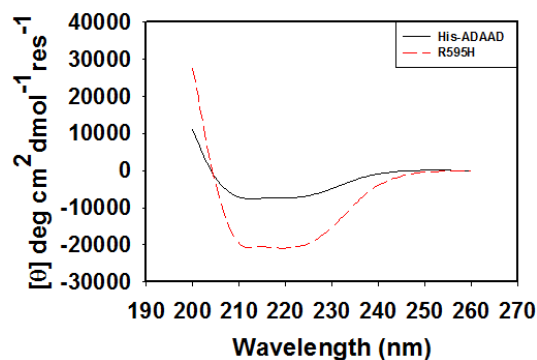
(a)
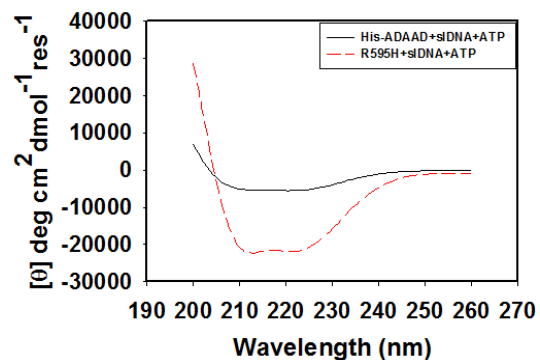
(b)

Figure: 6.3 Circular dichroism spectra recorded in mutant protein R595H and compared to wild-type ADAAD protein. Comparison of CD structures for (a) protein alone, (b) protein incubation with stem-loop DNA in presence of ATP. The protein concentration used in CD analyses was 0.1 mg/ml.

Molecular simulation study predicts a loss in dynamic stability in case of R595H

Due to the unavailability of crystal structure of ADAAD, molecular simulation studies were used for studying the stability and molecular movement of ADAAD (courtesy: Mr. Mohit Mazumdar). The dynamic stability and the overall movement, compared to the starting structure, during the molecular simulation runs were evaluated by calculating root mean square deviation (RMSD). The RMSD values showed that the wild-type ADAAD protein was stable during the entire run of 100 ns of molecular simulations (Figure 6.4a). However, R595H showed a significant decrease in RMSD values confirming the fluctuations in the mutant protein (Figure 6.4a). The solvent accessible surface area (SASA) calculation showed that overall accessible surface area decreased

for wild-type ADAAD during the molecular simulation run (Figure 6.4b). This decrease was higher in case of R595H mutant predicting it as a reason for unavailability of surface area to bind stem-loop DNA in the presence of ATP (Figure 6.4b).

(a)   (b)
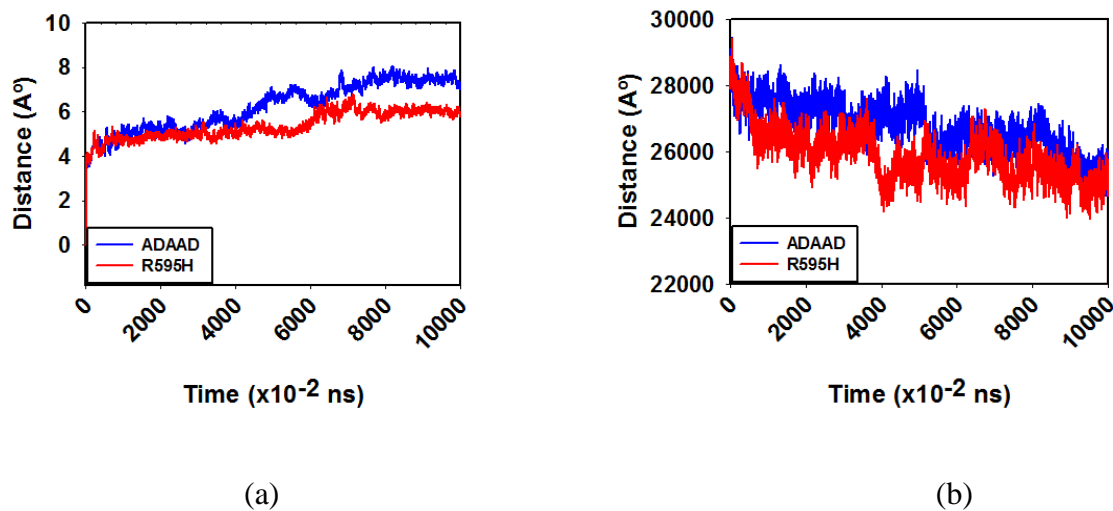
Figure: 6.4 Molecular simulation studies for ADAAD and R595H. (a) RMSD analysis of ADAAD and R595H. (b) SASA analysis of ADAAD and R595H.

R595H is unable to induce the conformational change in *c-myc* and *dgcr8* promoter

Till now, I have studied the role of R595H in ligand binding and ATP hydrolysis using ADAAD as the model system. Due to the presence of R595 residue in motif VI and the link of R595H mutation with SIOD disease, now I wanted to study the effect of this mutation on the transcriptional mechanism. I used circular dichroism to study the conformational changes in the promoter DNA induced by R595H as compared to the wild-type protein. For these experiments, I used two promoters: c-*myc* and *dgcr8*, both of which are negatively regulated by SMARCAL1

(courtesy: Ketki Patne) (Sharma 2016). Conformational changes induced in G_EC_E c-*myc* DNA by ADAAD were specified by two positive peaks, at 210 nm and at 259 nm, similar to the conformation induced in stem-loop DNA (Figure 6.5a). To address the question, whether R595H could induce similar conformational changes in G_EC_E DNA, G_EC_E c-*myc* DNA was incubated with R595H mutant protein, and conformational changes were recorded through CD spectroscopy. G_EC_E c-*myc* DNA showed a positive peak at 210 nm but a negative peak at 260 nm, suggesting that conformation alteration induced by R595H mutation disrupts the transcriptional inhibition process of c-*myc* catalyzed by SMARCAL1 (Figure 6.5b). This change in conformation of c-*myc* was similar to K241A suggesting ATPase activity of ADAAD is crucial for regulating the transcription of c-*myc*. Recently, it has been shown in our lab that SMARCAL1 inhibits the transcription of *dgcr8* in A549 cells by inducing a positive peak at 210 nm and a negative peak at 260 nm (Figure 6.6a) (courtesy: Ketki Patne). To understand the changes induced in *dgcr8* promoter DNA by R595H, I incubated *dgcr8* DNA with R595H and found that it resulted in a positive peak at 210 nm and 260 nm. The altered secondary structure of c-*myc* and *dgcr8* promoter induced by the R595H mutant protein suggests that R595H is not able to regulate the transcription of these genes (Figure 6.6b).
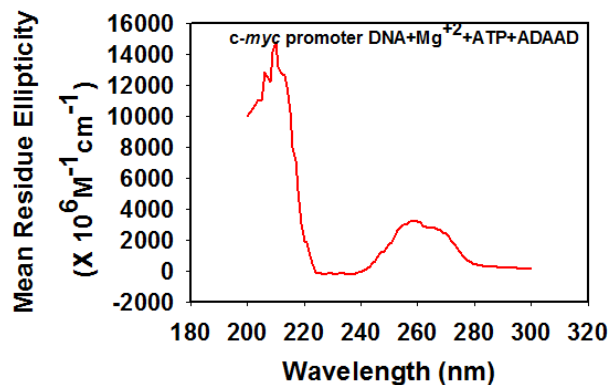
(a)
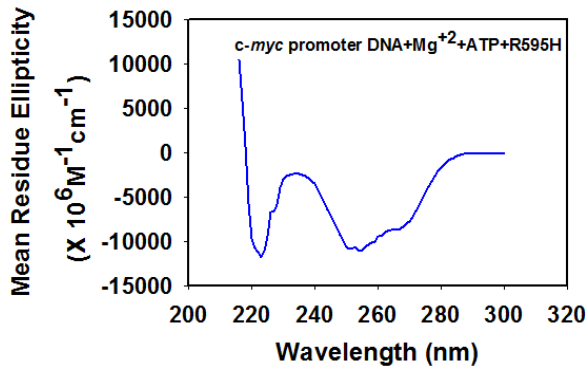
(b)

Figure: 6.5 (a) C-*myc* promoter DNA conformation in the presence of ATP and Mg^{2+} induced by wild-type ADAAD. (b) Change in the conformation of c-*myc* promoter DNA after its incubation with R595H.
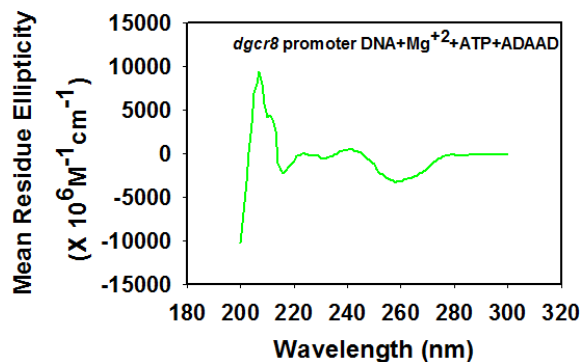
(a)
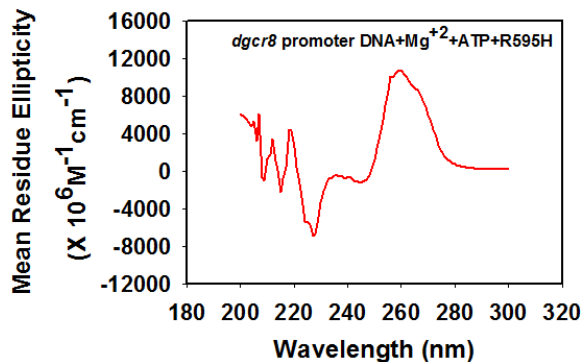
(b)

Figure: 6.6 (a) *dgcr8* promoter DNA conformation in the presence of ATP and Mg^{+2} induced by wild-type ADAAD. (b) Change in conformation of *dgcr8* promoter DNA after its incubation with R595H.

The R820H mutant co-localizes with γH2AX

As studied in various reports, SMARCAL1 plays crucial role in DNA damage repair. In my study, I overexpressed the SMARCAL1 in HeLa cells and studied its locaization. SMARCAL1 was found to be present predominantly in nucleus during the normal conditions (Figure 6.7a). However, when DNA damage was induced by the treatment of 2 µM doxorubicin, SMARCAL1 showed formation of foci and colocalization with γH2AX (Figure 6.7b). H2AX is a variant of H2A histone, a component of histone octamer. During DNA damage, H2AX is phosphorylated by ATM and ATR kinases forming gamma-H2AX. γH2AX foci formation act as a first step in recruiting repair proteins like SMARCAL1 at the DNA damage sites. This colocalization of SMARCAL1 and γH2AX depicts the interaction between these proteins during DNA damage. It has been previously shown in our laboratory that K464A, the ATPase dead mutant of SMARCAL1, cannnot form foci and co-localize with γH2AX on DNA damage (Figure 6.7a and b) (Gupta 2015). Therefore, I used K464A as a negative control and transfected HeLa cells with either K464A expressing construct or R820H expressing construct. Immunocytochemistry experiments showed that R820H like K464A localizes both in the cytoplam and nucleus (Figure 6.7a). However, unlike K464A, R820H was able to form foci and co-localize with γH2AX on induction of DNA damage by doxorubicin treatment (Figure 6.7b).

Untreated	Hoechst	GFP SMARCAL1	γH2AX	Merged
Lap Zeo vector 0.015 ± 0.012				
Wild type SMARCAL1 0.065 ± 0.023				
K464A 0.001 ± 0.016				
R820H 0.003 ± 0.015				
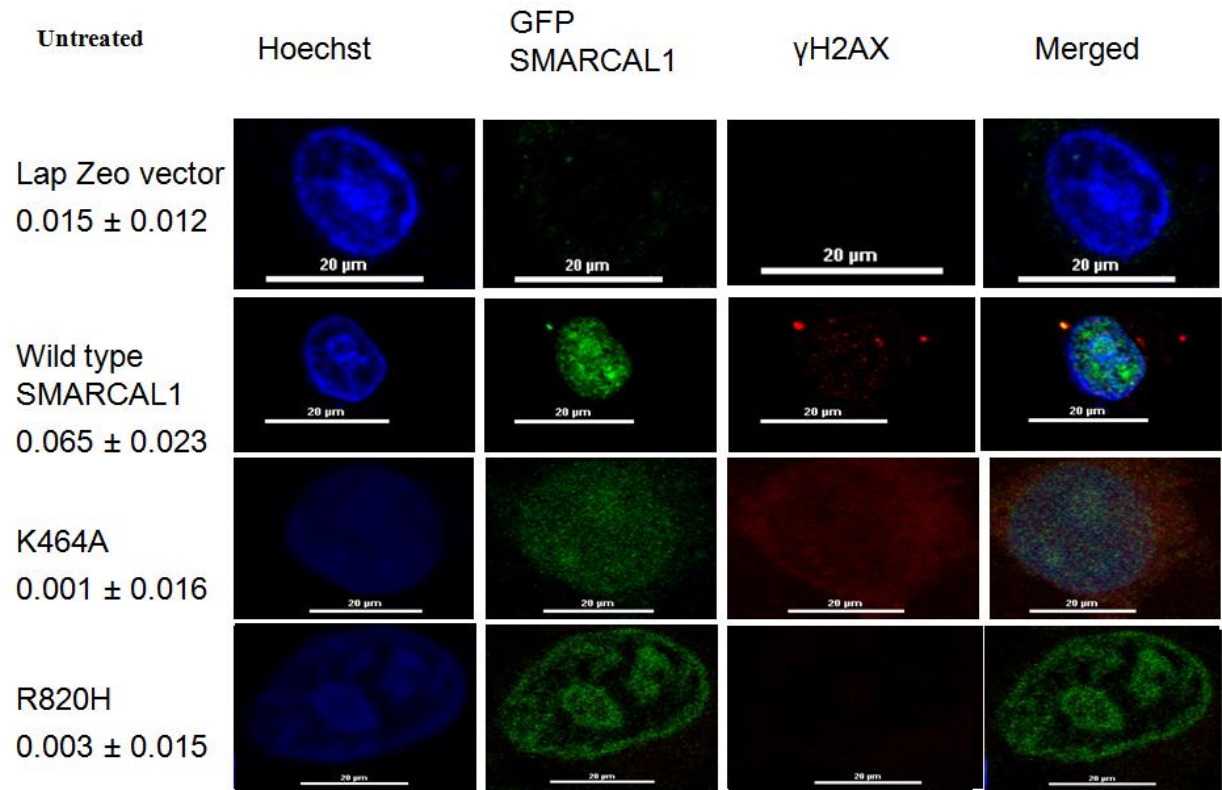
(a)
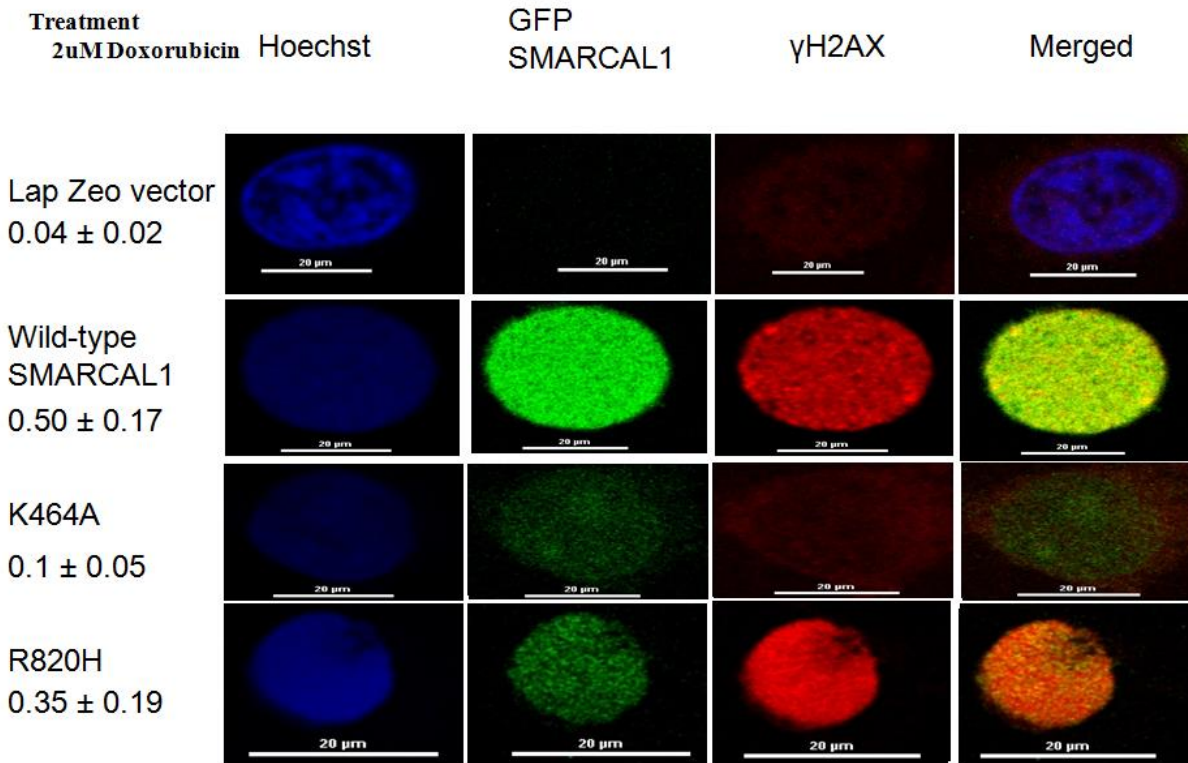
(b)

Figure: 6.7 K464A do not co-localize with γH2AX but SMARCAL1 and R820H co-localize with γH2AX. (a) Localization of wild-type SMARCAL1 and mutant proteins (K464A and R820H) in the absence of doxorubicin and, (b) Localization of wild-type SMARCAL1 and mutant proteins (K464A and R820H) in the presence of 2 μM doxorubicin. Pearson's coefficient for each sample is indicated in the figure (Scale- 20μm).

161

Table: 6.1 Ligand-protein (ADAAD and R595H) interactions were calculated and fitted as one-site saturation hyperbolic curve.

	ADAAD	R595H
ATP (X 10^{-6}M)	(1.6 ± 0.5)	(1.3 ± 0.2)
ATP(+slDNA) (X 10^{-6}M)	(0.14 ± 0.03)	(0.11 ± 0.04)
slDNA (X 10^{-9}M)	(19.9 ± 4.9)	(24.3 ± 6.1)
slDNA(+ATP) (X 10^{-9}M)	(3.4 ± 0.2)	(10.12 ± 0.33)

Table: 6.2 Dichroweb analyses (using K2D3 software) to calculate percentage of α helix and β sheet present in wild-type ADAAD and R595H mutant both in the absence of ligands and presence of stem-loop DNA after incubated with ATP.

	ADAAD	R595H
PROTEIN	35% α helix ,16% β sheet	61% α helix ,18% β sheet
PROTEIN+ slDNA+ ATP	26% α helix ,25% β sheet	65% α helix ,15% β sheet

DISCUSSION

Mutation in most of the helicase motifs of SMARCAL1 is associated with multisystem disorder SIOD which is characterized by disruption in ATPase activitiy (Boerkoel et al., 2002). Previously, we have studied the A245P, I325N and, S355L mutations that are present in SIOD patients and shown that the interaction of stem-loop DNA with these mutants in the presence of ATP impaired as compared to the wild-type ADAAD (Gupta et al., 2015). Further, it was shown that the conformation of the mutant proteins was altered as compared to the wild-type ADAAD protein both in the absence and presence of ligands indicating that the altered conformation might result in impaired stem-loop DNA binding and therefore, result in loss of ATPase activity (Gupta et al., 2015). These three mutations are present in RecA-like domain 1A; however, mutations mapping to the RecA-like domain 2A remain in SIOD patients has not been explored till now. As R820H mutation found in SIOD patients maps to motif VI, therefore, I decided to characterize this mutant with respect to its biochemical and physiological function. The R820H mutation in SMARCAL1 corresponds to R595H in ADAAD. Biochemical studies showed that R595H mutation leads to a loss in ATPase activity similar to other motif VI mutants. The R595H mutation was also found to alter the conformational integrity of protein resulting in weaker stem-loop DNA binding in the presence of ATP.

Molecular simulations performed with R595H showed a distorted dynamic stability as compared to the wild-type ADAAD protein. Finally, the R595H mutant was unable to induce conformational alteration in both c-*myc* and *dgcr8* promoter region necessary for their transcriptional regulation. This finding suggests that induction of disruptive transcriptional

regulation mechanism through the expression of mutant protein R595H might be a reason for the development of disease SIOD.

SMARCAL1 has shown to play a crucial role in DNA repair, cell-cycle regulation and genome maintaineance (Bansbach et al., 2009; Ciccia et al., 2009; Couch et al., 2013; Huang et al., 2010). The protein has been shown to co-localize with γH2AX on induction of DNA damage as, for example, by doxorubicin treatment (Driscoll and Cimprich, 2009; Gupta et al., 2015; Yuan et al., 2009). I found that the wild-type SMARCAL1 was present in both nucleus and cytoplasm in normal cells but forms foci on induction of DNA damage. Further, it co-localizes with γH2AX after the induction of DNA damage. These studies correlates well with the reports which suggest the role of SMARCAL1 in DNA damage repair after replication stress. The ATPase dead mutant protein, K464A, was not able to colocalize with γH2AX, suggesting the importance of ATPase activity of SMARCAL1 for DNA damage response. However, to my surprise, the R820H mutant was able to co-localize with γH2AX. As observed earlier, DNA damage induced by doxorubicin leads to upregulation of SMARCAL1 and BRG1. SMARCAL1 and BRG1 together, increase the expression of *dgcr8* gene and hence, induce the formation of non-coding RNA and 53BP1 foci (courtesy: Ketki Patne). As we observed that R595H was unable to induce the same conformational change in *dgcr8* promoter as ADAAD, we hypothesize that even after co-localization with γH2AX, R595H mutant protein is not able to mediate DNA damage repair.

Thus, these results help in better understanding of the biochemical and functional alteration induced in SIOD patients by the mutations and provide a plausible reason for the development of the disease and cancer predisposition.

SUMMARY

SMARCAL1 is a distant member of SNF2 family of ATP-dependent chromatin remodelers that are known to utilize the energy released from ATP hydrolysis in repositioning/evicting nucleosomes and thus, modulate various functional activities like DNA replication, DNA repair, transcription and cell-cycle regulation. The ATPase core of SMARCAL1, like other SF1 and SF2 members of helicases, consists of RecA-like domain 1A and 2A that spans motif I, Ia, II and III in the domain 1A and motif IV, V and VI in the domain 2A. Though very few studies have reported the function of these motifs in ligand binding and ATP hydrolysis, experiments conducted in our lab using ADAAD (N-terminal truncated proteolytic fragment of bovine homologue of SMARCAL1), have shown the importance of motif Q and I in ATP hydrolysis and not in ATP binding. The crystal structure of RNA helicase member, eIF4A, has suggested the role of motif VI in RNA binding and coupling ATPase activity to its helicase function. Based on this observation, I explored the importance of motif VI of ADAAD in coupling ATPase activity to stem-loop DNA binding and maintaining the conformational integrity of ADAAD.

I analysed the biochemical role of motif VI by constructing site-directed mutations in the positively and negatively charged conserved amino acid residues (D591, R592, H594 and R595) of motif VI. All motif VI mutations (D591H, R592A, H594A, R595K and R595A) displayed loss of ATPase activity suggesting the importance of motif VI in maintaining the functional aspects of ADAAD protein. Circular dichroism experiment indicated R592 and R595, are essential for maintaining the conformational integrity of the protein. The altered conformation of the mutant protein led to defective stem-loop DNA binding in presence of ATP. In accordance with various helicase reports, the positive charge of R595 of motif VI was found to be more important for conformational integrity and stem-loop DNA binding in presence of ATP.

Various reports have shown the communication between two RecA-like subdomains 1A and 2A suggesting that the domains might work together to create a functional protein. Therefore, I studied inter-domain interactions by creating double mutations in ADAAD where amino acid residue from both the RecA-like domains was mutated. The biochemical studies conducted with these double mutants showed that both the RecA-like subdomains 1A and 2A, indeed, act synergistically to interact with stem-loop DNA in presence of ATP. For example, studies with the double mutant protein W277A/R595H, where W277A is present in RecA-domain 1 and R595H is present in RecA-domain 2, showed that the interaction with stem-loop DNA binding in the presence of ATP was more defective as compared to the single mutants (W277A and R595H). Similarly, I found a more prominent defect in stem-loop DNA binding in presence of ATP in the double mutant H329D/D591H where histidine 329 residue of motif II was mutated to aspartate and aspartate 591 of motif VI was mutated to histidine. Further, although salt bridge formation has been observed between motif II and motif VI in many other RNA helicases like eIF4A, no such salt bridge formation was found in ADAAD. In line with these findings, understanding the synergistic effect towards stem-loop DNA binding in presence of ATP between two motifs of RecA-like domain 2A that might form intra-domain interaction also becomes important. Hence, I created a double mutant where phenylalanine 507 of motif IV and arginine 592 of motif VI was mutated to alanine. The double mutant F507A/R592A showed a completely altered conformation as compared to single mutants with 0% alpha-helical content. Alteration in conformation of double mutation suggests that intra-domain interactions within RecA-like domain 2A are important to maintain the secondary structure of the protein and hence, ATP hydrolysis.

Biallelic mutations in SMARCAL1 cause SIOD, a multisystem disorder involving renal failure, T-cell immunodeficiency and cancer predisposition. Considering the nature of this disease in which function of many human organs is impaired, SMARCAL1 is postulated to have multiple functions in the cell. SMARCAL1 downregulation causes a decrease in expression of cyclinE2 and increase in expression of CDK inhibitor A (p21) suggesting a role of SMARCAL1 in cell cycle progression and transcription regulation (Huang et al., 2010). Alteration in the expression of c-*myc*, a transcription factor involved in cellular proliferation and differentiation, has been reported in *smarcal1* downregulated cells and during serum starvation indicated a role of SMARCAL1 in c-*myc* transcription regulation (Sharma 2016). Although SMARCAL1 is found to mediate c-*myc* regulation at both the places, the mode of regulation is different. SMARCAL1 and BRG1 form a positive feedback loop where both the ATP-dependent chromatin remodelers regulate each other's expression in a positive manner. SMARCAL1 downregulation, thus, resulted in decreased BRG1 expression which ultimately inhibited the c-*myc* transcription. During serum starvation experiment, serum insufficiency causes decreased c-*myc* expression and hence, the presence of SMARCAL1 and BRG1 was investigated on c-*myc* promoter. Chromatin immunoprecipitation and protein occupancy assay confirmed the localization of SMARCAL1 on c-*myc* promoter during serum starvation while BRG1 was found to be absent on c-*myc* promoter, confirming a role of SMARCAL1 in transcriptional inhibition of c-*myc*. Though SMARCAL1 function in c-*myc* transcriptional regulation was confirmed, the mechanism by which SMARCAL1 regulates its expression was still a mystery. The fact that SMARCAL1 stabilizes a stem-loop kind of structure at double-strand to single-strand transition regions made me curious to study whether SMARCAL1 stabilizes stem-loop like secondary structures at c-*myc* promoter. In line with this hypothesis, I synthesized a small DNA region of c-*myc* promoter, $G_E C_E$ DNA,

where SMARCAL1 was found to be localized during serum starvation. The synthesis of this small G_EC_E DNA region was based on the Mfold and QGRS mapper analysis which showed the presence of G-quadraplex and stem-loop regions in this c-*myc* promoter region. ATPase assays showed that intramolecular heating of G_EC_E DNA resulted in higher ATPase activity as compared to intermolecular heating, suggesting that secondary structure formation in c-*myc* promoter is necessary for its transcriptional regulation by SMRACAL1. Additionally, fluorescence spectroscopy and circular dichroism confirmed that SMARCAL1 binds to G_EC_E DNA and changes its conformation in an exactly similar manner as stem-loop DNA, indicating that SMARCAL1 stabilizes a stem-loop like secondary structure at c-*myc* promoter and therefore, resulted in its transcriptional inhibition.

I further showed that the ATPase activity is important for transcriptional regulation as the ATPase dead mutant of ADAAD, K241A, was not able to induce this conformation in G_EC_E DNA in the presence of ATP even though fluorescence spectroscopy confirmed that the mutant was able to bind to the DNA. The mechanism of transcription inhibition of c-*myc* by ADAAD/SMARCAL1 is first such report of SMARCAL1 to regulate c-*myc* transcription by secondary structure stabilization.

Considering mutations in SMARCAL1 cause SIOD, I hypothesized that the motif VI associated mutation R820H (R595H in ADAAD), that is reported in SIOD patients, changes the conformation of the protein and stem-loop DNA binding affinity resulting in defective cellular functions of SMARCAL1. Therefore, I explored the biochemical role of motif VI associated SIOD mutation R595H. Fluorescence studies showed defect in stem-loop DNA binding in presence of ATP while circular dichroism analysis indicated conformational alteration in the

mutant protein R595H. Loss in conformational integrity of mutant protein R595H might be a reason for defect in ligand binding and ATPase activity of the protein. Additionally, molecular simulation studies like RMSD value showed a defect in dynamic stability of R595H, confirming the distortion in its secondary structure. As SMARCAL1 is reported to play role in transcriptional inhibition of c-*myc* and *dgcr8*, I also checked the transcriptional regulation of these genes by SIOD mutant R595H. Circular dichroism studies revealed that R595H was unable to induce the similar conformational alteration in both c-*myc* and *dgcr8* promoter region as ADAAD, suggesting that motif VI associated mutation R820H, present in SIOD patients, is not able to compensate the transcriptional role played by SMARCAL1.

Future Perspective

In my thesis work, I have studied the biochemical properties of motif VI of SMARCAL1. However, the role of other motifs of RecA-like domain 2A remains to be investigated. The biochemical information regarding all these motifs will help in unraveling the ATP and DNA binding sites of SMARCAL1. These studies will also help in understanding the ATP hydrolysis mechanism of SMARCAL1 in detail. Though the role of SIOD mutant present in motif VI has also been investigated in transcription regulation and DNA damage repair in my thesis, the roles of these mutations in causing T-cell immunodeficiency and apoptosis needs to be further investigated. The results from my thesis suggests an interaction between SMARCAL1, BRG1 and RPA at the c-*myc* promoter, yet the other proteins which might play a role in c-*myc* regulation like p21 and p53 also needs to be looked upon the c-*myc* promoter.

As I mentioned that inter-domain and intra-domain interactions in ADAAD is crucial for synergistic DNA binding and conformational integrity, respectively, the other motifs like motif III and motif V need to be further investigated for inter-domain interaction. For detailed study of the interaction of different motifs with ligands like stem-loop DNA and ATP, the crystal structure of ADAAD needs to be identified.

REFERENCES

Aasland, R., Gibson, T.J., and Stewart, A.F. (1995). The PHD finger: implications for chromatin-mediated transcriptional regulation. Trends in biochemical sciences *20*, 56-59.

Adra, C.N., Donato, J.-L., Badovinac, R., Syed, F., Kheraj, R., Cai, H., Moran, C., Kolker, M.T., Turner, H., and Weremowicz, S. (2000). SMARCAD1, a novel human helicase family-defining member associated with genetic instability: cloning, expression, and mapping to 4q22–q23, a band rich in breakpoints and deletion mutants involved in several human diseases. Genomics *69*, 162-173.

Ahmad, M., Ansari, A., Tarique, M., Satsangi, A.T., and Tuteja, R. (2012). Plasmodium falciparum UvrD helicase translocates in 3′ to 5′ direction, colocalizes with MLH and modulates its activity through physical interaction. PLoS One *7*, e49385.

Aihara, T., Miyoshi, Y., Koyama, K., Suzuki, M., Takahashi, E., Monden, M., and Nakamura, Y. (1998). Cloning and mapping of SMARCA5 encoding hSNF2H, a novel human homologue of Drosophila ISWI. Cytogenetic and Genome Research *81*, 191-193.

Alexeev, A., Mazin, A., and Kowalczykowski, S.C. (2003). Rad54 protein possesses chromatin-remodeling activity stimulated by the Rad51–ssDNA nucleoprotein filament. Nature Structural & Molecular Biology *10*, 182-186.

Alexiadis, V., and Kadonaga, J.T. (2002). Strand pairing by Rad54 and Rad51 is enhanced by chromatin. Genes & development *16*, 2767-2771.

Angus-Hill, M.L., Schlichter, A., Roberts, D., Erdjument-Bromage, H., Tempst, P., and Cairns, B.R. (2001). A Rsc3/Rsc30 zinc cluster dimer reveals novel roles for the chromatin remodeler RSC in gene expression and cell cycle control. Molecular cell *7*, 741-751.

Aravind, L., and Landsman, D. (1998). AT-hook motifs identified in a wide variety of DNA-binding proteins. Nucleic acids research *26*, 4413-4421.

Auble, D.T., Wang, D., Post, K.W., and Hahn, S. (1997). Molecular analysis of the SNF2/SWI2 protein family member MOT1, an ATP-driven enzyme that dissociates TATA-binding protein from DNA. Molecular and cellular biology *17*, 4842-4851.

Bachhawat, N., Ouyang, Q., and Henry, S.A. (1995). Functional Characterization of an Inositol-sensitive Upstream Activation Sequence in Yeast A cis-REGULATORY ELEMENT RESPONSIBLE FOR INOSITOL-CHOLINE MEDIATED REGULATION OF PHOSPHOLIPID BIOSYNTHESIS. Journal of Biological Chemistry *270*, 25087-25095.

Bakayev, V., Melnickov, A., Osicka, V., and Varshavsky, A. (1975). Studies an chromatin. II. Isolation and characterization of chromatin subunits. Nucleic acids research *2*, 1401-1420.

Banroques, J., Cordin, O., Doere, M., Linder, P., and Tanner, N.K. (2008). A conserved phenylalanine of motif IV in superfamily 2 helicases is required for cooperative, ATP-dependent binding of RNA substrates in DEAD-box proteins. Molecular and cellular biology *28*, 3359-3371.

Bansbach, C.E., Bétous, R., Lovejoy, C.A., Glick, G.G., and Cortez, D. (2009). The annealing helicase SMARCAL1 maintains genome integrity at stalled replication forks. Genes & development *23*, 2405-2414.

Bao, Y., Konesky, K., Park, Y.J., Rosu, S., Dyer, P.N., Rangasamy, D., Tremethick, D.J., Laybourn, P.J., and Luger, K. (2004). Nucleosomes containing the histone variant H2A. Bbd organize only 118 base pairs of DNA. The EMBO journal *23*, 3314-3324.

Baradaran-Heravi, A., Cho, K.S., Tolhuis, B., Sanyal, M., Morozova, O., Morimoto, M., Elizondo, L.I., Bridgewater, D., Lubieniecka, J., and Beirnes, K. (2012). Penetrance of biallelic

SMARCAL1 mutations is associated with environmental and genetic disturbances of gene expression. Human molecular genetics, dds083.

Barbera, A.J., Chodaparambil, J.V., Kelley-Clarke, B., Joukov, V., Walter, J.C., Luger, K., and Kaye, K.M. (2006). The nucleosomal surface as a docking station for Kaposi's sarcoma herpesvirus LANA. Science *311*, 856-861.

Barton, A.B., and Kaback, D.B. (1994). Molecular cloning of chromosome I DNA from Saccharomyces cerevisiae: analysis of the genes in the FUN38-MAK16-SPO7 region. Journal of bacteriology *176*, 1872-1880.

Becker, P.B., and Hörz, W. (2002). ATP-dependent nucleosome remodeling. Annual review of biochemistry *71*, 247-273.

Bednar, J., Horowitz, R.A., Grigoryev, S.A., Carruthers, L.M., Hansen, J.C., Koster, A.J., and Woodcock, C.L. (1998). Nucleosomes, linker DNA, and linker histone form a unique structural motif that directs the higher-order folding and compaction of chromatin. Proceedings of the National Academy of Sciences *95*, 14173-14178.

Beisel, C., Imhof, A., Greene, J., Kremmer, E., and Sauer, F. (2002). Histone methylation by the Drosophila epigenetic transcriptional regulator Ash1. Nature *419*, 857-862.

Bennett, M., Littlewood, T., Hancock, D., Evan, G., and Newby, A. (1994). Down-regulation of the c-myc proto-oncogene in inhibition of vascular smooth-muscle cell proliferation: a signal for growth arrest? Biochemical Journal *302*, 701-708.

Bird, A. (2002). DNA methylation patterns and epigenetic memory. Genes & development *16*, 6-21.

Bird, A.P., and Wolffe, A.P. (1999). Methylation-induced repression—belts, braces, and chromatin. Cell *99*, 451-454.

Bochar, D.A., Wang, L., Beniya, H., Kinev, A., Xue, Y., Lane, W.S., Wang, W., Kashanchi, F., and Shiekhattar, R. (2000). BRCA1 is associated with a human SWI/SNF-related complex: linking chromatin remodeling to breast cancer. Cell *102*, 257-265.

Boerkoel, C., O'neill, S., Andre, J., Benke, P., Bogdanović, R., Bulla, M., Burguet, A., Cockfield, S., Cordeiro, I., and Ehrich, J. (2000). Manifestations and treatment of Schimke immuno-osseous dysplasia: 14 new cases and a review of the literature. European journal of pediatrics *159*, 1-7.

Boerkoel, C.F., Takashima, H., John, J., Yan, J., Stankiewicz, P., Rosenbarker, L., Andre, J.-L., Bogdanovic, R., Burguet, A., and Cockfield, S. (2002). Mutant chromatin remodeling protein SMARCAL1 causes Schimke immuno-osseous dysplasia. Nature genetics *30*, 215-220.

Bosman, E.A., Penn, A.C., Ambrose, J.C., Kettleborough, R., Stemple, D.L., and Steel, K.P. (2005). Multiple mutations in mouse Chd7 provide models for CHARGE syndrome. Human molecular genetics *14*, 3463-3476.

Boyer, L.A., Latek, R.R., and Peterson, C.L. (2004). The SANT domain: a unique histone-tail-binding module? Nature reviews Molecular cell biology *5*, 158-163.

Bozhenok, L., Wade, P.A., and Varga-Weisz, P. (2002). WSTF–ISWI chromatin remodeling complex targets heterochromatic replication foci. The EMBO Journal *21*, 2231-2241.

Brosh, R., and Matson, S.W. (1995). Mutations in motif II of Escherichia coli DNA helicase II render the enzyme nonfunctional in both mismatch repair and excision repair with differential effects on the unwinding reaction. Journal of bacteriology *177*, 5612-5621.

Brown, K.E., Guest, S.S., Smale, S.T., Hahm, K., Merkenschlager, M., and Fisher, A.G. (1997). Association of transcriptionally silent genes with Ikaros complexes at centromeric heterochromatin. Cell *91*, 845-854.

Brownell, J.E., Zhou, J., Ranalli, T., Kobayashi, R., Edmondson, D.G., Roth, S.Y., and Allis, C.D. (1996). Tetrahymena histone acetyltransferase A: a homolog to yeast Gcn5p linking histone acetylation to gene activation. Cell *84*, 843-851.

Bultman, S., Gebuhr, T., Yee, D., La Mantia, C., Nicholson, J., Gilliam, A., Randazzo, F., Metzger, D., Chambon, P., and Crabtree, G. (2000). A Brg1 null mutation in the mouse reveals functional differences among mammalian SWI/SNF complexes. Molecular cell *6*, 1287-1295.

Burns, L.G., and Peterson, C.L. (1997). The yeast SWI-SNF complex facilitates binding of a transcriptional activator to nucleosomal sites in vivo. Molecular and cellular biology *17*, 4811-4819.

Cairns, B.R. (2007). Chromatin remodeling: insights and intrigue from single-molecule studies. Nature structural & molecular biology *14*, 989-996.

Cairns, B.R., Kim, Y.-J., Sayre, M.H., Laurent, B.C., and Kornberg, R.D. (1994). A multisubunit complex containing the SWI1/ADR6, SWI2/SNF2, SWI3, SNF5, and SNF6 gene products isolated from yeast. Proceedings of the National Academy of Sciences *91*, 1950-1954.

Cairns, B.R., Lorch, Y., Li, Y., Zhang, M., Lacomis, L., Erdjument-Bromage, H., Tempst, P., Du, J., Laurent, B., and Kornberg, R.D. (1996). RSC, an essential, abundant chromatin-remodeling complex. Cell *87*, 1249-1260.

Carroll, C., Badu-Nkansah, A., Hunley, T., Baradaran-Heravi, A., Cortez, D., and Frangoul, H. (2013). Schimke Immunoosseous Dysplasia associated with undifferentiated carcinoma and a novel SMARCAL1 mutation in a child. Pediatric blood & cancer *60*.

Caruthers, J.M., Johnson, E.R., and McKay, D.B. (2000). Crystal structure of yeast initiation factor 4A, a DEAD-box RNA helicase. Proceedings of the National Academy of Sciences *97*, 13080-13085.

Chan, K.L., Palmai-Pallag, T., Ying, S., and Hickson, I.D. (2009). Replication stress induces sister-chromatid bridging at fragile site loci in mitosis. Nature cell biology *11*, 753-760.

Chen, S., Davies, A.A., Sagan, D., and Ulrich, H.D. (2005). The RING finger ATPase Rad5p of Saccharomyces cerevisiae contributes to DNA double-strand break repair in a ubiquitin-independent manner. Nucleic acids research *33*, 5878-5886.

Cheung, P., Tanner, K.G., Cheung, W.L., Sassone-Corsi, P., Denu, J.M., and Allis, C.D. (2000). Synergistic coupling of histone H3 phosphorylation and acetylation in response to epidermal growth factor stimulation. Molecular cell *5*, 905-915.

Chi, T.H., Wan, M., Lee, P.P., Akashi, K., Metzger, D., Chambon, P., Wilson, C.B., and Crabtree, G.R. (2003). Sequential roles of Brg, the ATPase subunit of BAF chromatin remodeling complexes, in thymocyte development. Immunity *19*, 169-182.

Chicca, J.J., Auble, D.T., and Pugh, B.F. (1998). Cloning and biochemical characterization of TAF-172, a human homolog of yeast Mot1. Molecular and cellular biology *18*, 1701-1710.

Cho, H., Orphanides, G., Sun, X., Yang, X.-J., Ogryzko, V., Lees, E., Nakatani, Y., and Reinberg, D. (1998). A human RNA polymerase II complex containing factors that modify chromatin structure. Molecular and cellular biology *18*, 5355-5363.

Ciccia, A., Bredemeyer, A.L., Sowa, M.E., Terret, M.-E., Jallepalli, P.V., Harper, J.W., and Elledge, S.J. (2009). The SIOD disorder protein SMARCAL1 is an RPA-interacting protein involved in replication fork restart. Genes & development *23*, 2415-2425.

Citterio, E., Van Den Boom, V., Schnitzler, G., Kanaar, R., Bonte, E., Kingston, R.E., Hoeijmakers, J.H., and Vermeulen, W. (2000). ATP-dependent chromatin remodeling by the Cockayne syndrome B DNA repair-transcription-coupling factor. Molecular and cellular biology *20*, 7643-7653.

Clapier, C.R., Längst, G., Corona, D.F., Becker, P.B., and Nightingale, K.P. (2001). Critical role for the histone H4 N terminus in nucleosome remodeling by ISWI. Molecular and Cellular Biology *21*, 875-883.

Clark, M.W., Zhong, W.W., Keng, T., Storms, R.K., Barton, A., Kaback, D.B., and Bussey, H. (1992). Identification of a Saccharomyces cerevisiae homolog of the SNF2 transcrioptional regulator in the DNA sequence of an 8· 6 kb region in the LTE1-CYS1 interval on the left arm of chormosome I. Yeast *8*, 133-145.

Clements, A., Poux, A.N., Lo, W.-S., Pillus, L., Berger, S.L., and Marmorstein, R. (2003). Structural basis for histone and phosphohistone binding by the GCN5 histone acetyltransferase. Molecular cell *12*, 461-473.

Clewing, J.M., Fryssira, H., Goodman, D., Smithson, S.F., Sloan, E.A., Lou, S., Huang, Y., Choi, K., Lücke, T., and Alpay, H. (2007). Schimke immunoosseous dysplasia: suggestions of genetic diversity. Human mutation *28*, 273-283.

Cluzel, P., Lebrun, A., Heller, C., and Lavery, R. (1996). DNA: an extensible molecule. Science *271*, 792.

Coleman, M.A., Eisen, J.A., and Mohrenweiser, H.W. (2000). Cloning and characterization of HARP/SMARCAL1: a prokaryotic HepA-related SNF2 helicase protein from human and mouse. Genomics *65*, 274-282.

Cosma, M.P., Tanaka, T., and Nasmyth, K. (1999). Ordered recruitment of transcription and chromatin remodeling factors to a cell cycle–and developmentally regulated promoter. Cell *97*, 299-311.

Côté, J., Peterson, C.L., and Workman, J.L. (1998). Perturbation of nucleosome core structure by the SWI/SNF complex persists after its detachment, enhancing subsequent transcription factor binding. Proceedings of the National Academy of Sciences *95*, 4947-4952.

Couch, F.B., Bansbach, C.E., Driscoll, R., Luzwick, J.W., Glick, G.G., Bétous, R., Carroll, C.M., Jung, S.Y., Qin, J., and Cimprich, K.A. (2013). ATR phosphorylates SMARCAL1 to prevent replication fork collapse. Genes & development *27*, 1610-1623.

Cox, K.E., Maréchal, A., and Flynn, R.L. (2016). SMARCAL1 resolves replication stress at ALT telomeres. Cell reports *14*, 1032-1040.

Cox, M.M., Goodman, M.F., Kreuzer, K.N., Sherratt, D.J., Sandler, S.J., and Marians, K.J. (2000). The importance of repairing stalled replication forks. Nature *404*, 37-41.

Dang, W., and Bartholomew, B. (2007). Domain architecture of the catalytic subunit in the ISW2-nucleosome complex. Molecular and cellular biology *27*, 8306-8317.

Dang, W., Kagalwala, M.N., and Bartholomew, B. (2006). Regulation of ISW2 by concerted action of histone H4 tail and extranucleosomal DNA. Molecular and cellular biology *26*, 7388-7396.

Darst, R.P., Wang, D., and Auble, D.T. (2001). MOT1-catalyzed TBP–DNA disruption: uncoupling DNA conformational change and role of upstream DNA. The EMBO journal *20*, 2028-2040.

Dasgupta, A., Juedes, S.A., Sprouse, R.O., and Auble, D.T. (2005). Mot1-mediated control of transcription complex assembly and activity. The EMBO journal *24*, 1717-1729.

Davey, C.A., Sargent, D.F., Luger, K., Maeder, A.W., and Richmond, T.J. (2002). Solvent mediated interactions in the structure of the nucleosome core particle at 1.9 Å resolution. Journal of molecular biology *319*, 1097-1113.

Dechassa, M.L., Zhang, B., Horowitz-Scherer, R., Persinger, J., Woodcock, C.L., Peterson, C.L., and Bartholomew, B. (2008). Architecture of the SWI/SNF-nucleosome complex. Molecular and cellular biology *28*, 6010-6021.

Deguchi, K., Clewing, J.M., Elizondo, L.I., Hirano, R., Huang, C., Choi, K., Sloan, E.A., Lücke, T., Marwedel, K.M., and Powell, R.D. (2008). Neurologic phenotype of Schimke immuno-osseous dysplasia and neurodevelopmental expression of SMARCAL1. Journal of Neuropathology & Experimental Neurology *67*, 565-577.

DeLange, R.J., Fambrough, D.M., Smith, E.L., and Bonner, J. (1969). Calf and pea histone IV III. Complete amino acid sequence of pea seedling Histone IV; comparison with the homologous calf thymus histone. Journal of Biological Chemistry *244*, 5669-5679.

Delmas, V., Stokes, D.G., and Perry, R.P. (1993). A mammalian DNA-binding protein that contains a chromodomain and an SNF2/SWI2-like helicase domain. Proceedings of the National Academy of Sciences *90*, 2414-2418.

Dillon, S.C., Zhang, X., Trievel, R.C., and Cheng, X. (2005). The SET-domain protein superfamily: protein lysine methyltransferases. Genome biology *6*, 227.

Dirscherl, S.S., and Krebs, J.E. (2004). Functional diversity of ISWI complexes. Biochemistry and cell biology *82*, 482-489.

Dorigo, B., Schalch, T., Kulangara, A., Duda, S., Schroeder, R.R., and Richmond, T.J. (2004). Nucleosome arrays reveal the two-start organization of the chromatin fiber. Science *306*, 1571-1573.

Dou, Y., Bowen, J., Liu, Y., and Gorovsky, M.A. (2002). Phosphorylation and an ATP-dependent process increase the dynamic exchange of H1 in chromatin. The Journal of cell biology *158*, 1161-1170.

Doyen, C.M., Montel, F., Gautier, T., Menoni, H., Claudet, C., Delacour-Larose, M., Angelov, D., Hamiche, A., Bednar, J., and Faivre-Moskalenko, C. (2006). Dissection of the unusual structural and functional properties of the variant H2A. Bbd nucleosome. The EMBO journal *25*, 4234-4244.

Driscoll, R., and Cimprich, K.A. (2009). HARPing on about the DNA damage response during replication. Genes & development *23*, 2359-2365.

Dunaief, J.L., Strober, B.E., Guha, S., Khavari, P.A., Ålin, K., Luban, J., Begemann, M., Crabtree, G.R., and Goff, S.P. (1994). The retinoblastoma protein and BRG1 form a complex and cooperate to induce cell cycle arrest. Cell *79*, 119-130.

Dürr, H., Flaus, A., Owen-Hughes, T., and Hopfner, K.-P. (2006). Snf2 family ATPases and DExx box helicases: differences and unifying concepts from high-resolution crystal structures. Nucleic acids research *34*, 4160-4167.

Dürr, H., Körner, C., Müller, M., Hickmann, V., and Hopfner, K.-P. (2005). X-ray structures of the Sulfolobus solfataricus SWI2/SNF2 ATPase core and its complex with DNA. Cell *121*, 363-373.

Ebbert, R., Birkmann, A., and Schüller, H.J. (1999). The product of the SNF2/SWI2 paralogue INO80 of Saccharomyces cerevisiae required for efficient expression of various yeast structural genes is part of a high-molecular-weight protein complex. Molecular microbiology *32*, 741-751.

Elfring, L.K., Deuring, R., McCallum, C.M., Peterson, C.L., and Tamkun, J.W. (1994). Identification and characterization of Drosophila relatives of the yeast transcriptional activator SNF2/SWI2. Molecular and cellular biology *14*, 2225-2234.

Elizondo, L.I., Cho, K.S., Zhang, W., Yan, J., Huang, C., Huang, Y., Choi, K., Sloan, E.A., Deguchi, K., and Lou, S. (2009). Schimke immuno-osseous dysplasia: SMARCAL1 loss-of-function and phenotypic correlation. Journal of medical genetics *46*, 49-59.

Elizondo, L.I., Huang, C., Northrop, J.L., Deguchi, K., Clewing, J.M., Armstrong, D.L., and Boerkoel, C.F. (2006). Schimke immuno-osseous dysplasia: A cell autonomous disorder? American journal of medical genetics Part A *140*, 340-348.

Evan, G.I., and Littlewood, T.D. (1993). The role of c-myc in cell growth. Current opinion in genetics & development *3*, 44-49.

Evan, G.I., Wyllie, A.H., Gilbert, C.S., Littlewood, T.D., Land, H., Brooks, M., Waters, C.M., Penn, L.Z., and Hancock, D.C. (1992). Induction of apoptosis in fibroblasts by c-myc protein. Cell *69*, 119-128.

Fairman-Williams, M.E., Guenther, U.-P., and Jankowsky, E. (2010). SF1 and SF2 helicases: family matters. Current opinion in structural biology *20*, 313-324.

Fan, H.-Y., He, X., Kingston, R.E., and Narlikar, G.J. (2003). Distinct strategies to make nucleosomal DNA accessible. Molecular cell *11*, 1311-1322.

Fan, J.Y., Rangasamy, D., Luger, K., and Tremethick, D.J. (2004). H2A. Z alters the nucleosome surface to promote HP1α-mediated chromatin fiber folding. Molecular cell *16*, 655-661.

Farris, S.D., Rubio, E.D., Moon, J.J., Gombert, W.M., Nelson, B.H., and Krumm, A. (2005). Transcription-induced chromatin remodeling at the c-myc gene involves the local exchange of histone H2A. Z. Journal of Biological Chemistry *280*, 25298-25303.

Feng, W., Khan, M.A., Bellvis, P., Zhu, Z., Bernhardt, O., Herold-Mende, C., and Liu, H.-K. (2013). The chromatin remodeler CHD7 regulates adult neurogenesis via activation of SoxC transcription factors. Cell stem cell *13*, 62-72.

Fernández, A., Guo, H.S., Sáenz, P., Simón-Buela, L., de Cedrón, M.G., and García, J.A. (1997). The motif V of plum pox potyvirus CI RNA helicase is involved in NTP hydrolysis and is essential for virus RNA replication. Nucleic acids research *25*, 4474-4480.

Fernández, A., Laín, S., and Garcia, J.A. (1995). RNA helicase activity of the plum pox potyvirus Cl protein expressed in Escherichia coli. Mapping of an RNA binding domain. Nucleic acids research *23*, 1327-1332.

Finch, J., Lutter, L., Rhodes, D., Brown, R., Rushton, B., Levitt, M., and Klug, A. (1977). Structure of nucleosome core particles of chromatin. Nature *269*, 29-36.

Fischle, W., Wang, Y., and Allis, C.D. (2003a). Binary switches and modification cassettes in histone biology and beyond. Nature *425*, 475-479.

Fischle, W., Wang, Y., Jacobs, S.A., Kim, Y., Allis, C.D., and Khorasanizadeh, S. (2003b). Molecular basis for the discrimination of repressive methyl-lysine marks in histone H3 by Polycomb and HP1 chromodomains. Genes & development *17*, 1870-1881.

Flaus, A., Martin, D.M., Barton, G.J., and Owen-Hughes, T. (2006). Identification of multiple distinct Snf2 subfamilies with conserved structural motifs. Nucleic acids research *34*, 2887-2905.

Flaus, A., and Owen-Hughes, T. (2001). Mechanisms for ATP-dependent chromatin remodelling. Current opinion in genetics & development *11*, 148-154.

Flaus, A., and Owen-Hughes, T. (2004). Mechanisms for ATP-dependent chromatin remodelling: farewell to the tuna-can octamer? Current opinion in genetics & development *14*, 165-173.

Fryer, C.J., and Archer, T.K. (1998). Chromatin remodelling by the glucocorticoid receptor requires the BRG1 complex. Nature *393*, 88-91.

Fuks, F., Hurd, P.J., Wolf, D., Nan, X., Bird, A.P., and Kouzarides, T. (2003). The methyl-CpG-binding protein MeCP2 links DNA methylation to histone methylation. Journal of Biological Chemistry *278*, 4035-4040.

Fyodorov, D.V., and Kadonaga, J.T. (2002). Dynamics of ATP-dependent chromatin assembly by ACF. Nature *418*, 896-900.

Gallivan, J.-P., and McGarvey, M.J. (2003). The importance of the Q motif in the ATPase activity of a viral helicase. FEBS letters *554*, 485-488.

Gangaraju, V.K., and Bartholomew, B. (2007). Dependency of ISW1a chromatin remodeling on extranucleosomal DNA. Molecular and cellular biology *27*, 3217-3225.

Gautier, T., Abbott, D.W., Molla, A., Verdel, A., Ausio, J., and Dimitrov, S. (2004). Histone variant H2ABbd confers lower stability to the nucleosome. EMBO reports *5*, 715-720.

Geiss-Friedlander, R., and Melchior, F. (2007). Concepts in sumoylation: a decade on. Nature reviews Molecular cell biology *8*, 947-956.

George, J.W., Brosh, R.M., and Matson, S.W. (1994). A dominant negative allele of the Escherichia coli uvrD gene encoding DNA helicase II: a biochemical and genetic characterization. Journal of molecular biology *235*, 424-435.

Getzenberg, R.H., Pienta, K.J., Ward, W.S., and Coffey, D.S. (1991). Nuclear structure and the three-dimensional organization of DNA. Journal of cellular biochemistry *47*, 289-299.

Gibbons, R.J. (2008). Epigenetics and its Genetic Syndromes. Epigenetics in Biology and Medicine, 155-174.

Gibbons, R.J., McDowell, T.L., Raman, S., O'Rourke, D.M., Garrick, D., Ayyub, H., and Higgs, D.R. (2000). Mutations in ATRX, encoding a SWI/SNF-like protein, cause diverse changes in the pattern of DNA methylation. Nature genetics *24*, 368-371.

Gibbons, R.J., Picketts, D.J., Villard, L., and Higgs, D.R. (1995). Mutations in a putative global transcriptional regulator cause X-linked mental retardation with α-thalassemia (ATR-X syndrome). Cell *80*, 837-845.

Gilbert, N., Boyle, S., Fiegler, H., Woodfine, K., Carter, N.P., and Bickmore, W.A. (2004). Chromatin architecture of the human genome: gene-rich domains are enriched in open chromatin fibers. Cell *118*, 555-566.

Girdham, C.H., and Glover, D.M. (1991). Chromosome tangling and breakage at anaphase result from mutations in lodestar, a Drosophila gene encoding a putative nucleoside triphosphate-binding protein. Genes & Development *5*, 1786-1799.

González, V., and Hurley, L.H. (2010). The c-MYC NHE III1: function and regulation. Annual review of pharmacology and toxicology *50*, 111-129.

Gorbalenya, A.E., and Koonin, E.V. (1993). Helicases: amino acid sequence comparisons and structure-function relationships. Current opinion in structural biology *3*, 419-429.

Grand, C.L., Powell, T.J., Nagle, R.B., Bearss, D.J., Tye, D., Gleason-Guzman, M., and Hurley, L.H. (2004). Mutations in the G-quadruplex silencer element and their relationship to c-MYC overexpression, NM23 repression, and therapeutic rescue. Proceedings of the National Academy of Sciences of the United States of America *101*, 6140-6145.

Grant, P.A., Duggan, L., Côté, J., Roberts, S.M., Brownell, J.E., Candau, R., Ohba, R., Owen-Hughes, T., Allis, C.D., and Winston, F. (1997). Yeast Gcn5 functions in two multisubunit complexes to acetylate nucleosomal histones: characterization of an Ada complex and the SAGA (Spt/Ada) complex. Genes & development *11*, 1640-1650.

Graves-Woodward, K.L., Gottlieb, J., Challberg, M.D., and Weller, S.K. (1997). Biochemical analyses of mutations in the HSV-1 helicase-primase that alter ATP hydrolysis, DNA unwinding, and coupling between hydrolysis and unwinding. Journal of Biological Chemistry *272*, 4623-4630.

Graves-Woodward, K.L., and Weller, S.K. (1996). Replacement of gly815 in helicase motif V alters the single-stranded DNA-dependent ATPase activity of the herpes simplex virus type 1 helicase-primase. Journal of Biological Chemistry *271*, 13629-13635.

Greaves, I.K., Rangasamy, D., Ridgway, P., and Tremethick, D.J. (2007). H2A. Z contributes to the unique 3D structure of the centromere. Proceedings of the National Academy of Sciences *104*, 525-530.

Gregory, P.D., Schmid, A., Zavari, M., Münsterkötter, M., and Hörz, W. (1999). Chromatin remodelling at the PHO8 promoter requires SWI–SNF and SAGA at a step subsequent to activator binding. The EMBO Journal *18*, 6407-6414.

Gross, C.H., and Shuman, S. (1996). The QRxGRxGRxxxG motif of the vaccinia virus DExH box RNA helicase NPH-II is required for ATP hydrolysis and RNA unwinding but not for RNA binding. Journal of virology *70*, 1706-1713.

Grüne, T., Brzeski, J., Eberharter, A., Clapier, C.R., Corona, D.F., Becker, P.B., and Müller, C.W. (2003). Crystal structure and functional analysis of a nucleosome recognition module of the remodeling factor ISWI. Molecular cell *12*, 449-460.

Günes, C., and Rudolph, K.L. (2013). The role of telomeres in stem cells and cancer. Cell *152*, 390-393.

Gupta, M., Mazumder, M., Dhatchinamoorthy, K., Nongkhlaw, M., Haokip, D.T., Gourinath, S., Komath, S.S., and Muthuswami, R. (2015). Ligand-induced conformation changes drive ATP hydrolysis and function in SMARCAL1. FEBS journal *282*, 3841-3859.

Gupta,M. (2015). Mechanism of ATP hydrolysis by SWI2/SNF2 proteins.

Hall, M.C., Özsoy, A.Z., and Matson, S.W. (1998). Site-directed mutations in motif VI of Escherichia coli DNA helicase II result in multiple biochemical defects: evidence for the involvement of motif VI in the coupling of ATPase and DNA binding activities via conformational changes. Journal of molecular biology *277*, 257-271.

Haokip, D. T., Goel, I., Arya, V., Sharma, T., Kumari, R., Priya, R., Singh,M., and Muthuswami, R. (2016). Transcriptional Regulation of Atp-Dependent Chromatin Remodeling Factors: Smarcal1 and Brg1 Mutually Co-Regulate Each Other. Scientific Reports, *6*, 20532.

Harikrishnan, K., Chow, M.Z., Baker, E.K., Pal, S., Bassal, S., Brasacchio, D., Wang, L., Craig, J.M., Jones, P.L., and Sif, S. (2005). Brahma links the SWI/SNF chromatin-remodeling complex with MeCP2-dependent transcriptional silencing. Nature genetics *37*, 254-264.

Hauk, G., McKnight, J.N., Nodelman, I.M., and Bowman, G.D. (2010). The chromodomains of the Chd1 chromatin remodeler regulate DNA access to the ATPase motor. Molecular cell *39*, 711-723.

Havas, K., Flaus, A., PHeLan, M., Kingston, R., Wade, P.A., Lilley, D.M., and Owen-Hughes, T. (2000). Generation of superhelical torsion by ATP-dependent chromatin remodeling activities. Cell *103*, 1133-1142.

Heilek, G.M., and Peterson, M.G. (1997). A point mutation abolishes the helicase but not the nucleoside triphosphatase activity of hepatitis C virus NS3 protein. Journal of virology *71*, 6264-6266.

Hendricks, K.B., Shanahan, F., and Lees, E. (2004). Role for BRG1 in cell cycle control and tumor suppression. Molecular and cellular biology *24*, 362-376.

Hereford, L., Fahrner, K., Woolford, J., Rosbash, M., and Kaback, D.B. (1979). Isolation of yeast histone genes H2A and H2B. Cell *18*, 1261-1271.

Hirschhorn, J.N., Brown, S.A., Clark, C.D., and Winston, F. (1992). Evidence that SNF2/SWI2 and SNF5 activate transcription in yeast by altering chromatin structure. Genes & development *6*, 2288-2298.

Ho, L., and Crabtree, G.R. (2010). Chromatin remodelling during development. Nature *463*, 474-484.

Hoege, C., Pfander, B., Moldovan, G.-L., Pyrowolakis, G., and Jentsch, S. (2002). RAD6-dependent DNA repair is linked to modification of PCNA by ubiquitin and SUMO. Nature *419*, 135-141.

Hoffman, B., and Liebermann, D. (2008). Apoptotic signaling by c-MYC. Oncogene *27*, 6462-6472.

Holstege, F.C., Jennings, E.G., Wyrick, J.J., Lee, T.I., Hengartner, C.J., Green, M.R., Golub, T.R., Lander, E.S., and Young, R.A. (1998). Dissecting the regulatory circuitry of a eukaryotic genome. Cell *95*, 717-728.

Hota, S.K., and Bartholomew, B. (2011). Diversity of operation in ATP-dependent chromatin remodelers. Biochimica et Biophysica Acta (BBA)-Gene Regulatory Mechanisms *1809*, 476-487.

Hsu, J.-Y., Sun, Z.-W., Li, X., Reuben, M., Tatchell, K., Bishop, D.K., Grushcow, J.M., Brame, C.J., Caldwell, J.A., and Hunt, D.F. (2000). Mitotic phosphorylation of histone H3 is governed by Ipl1/aurora kinase and Glc7/PP1 phosphatase in budding yeast and nematodes. Cell *102*, 279-291.

Huang, C., Gu, S., Yu, P., Yu, F., Feng, C., Gao, N., and Du, J. (2010). Deficiency of smarcal1 causes cell cycle arrest and developmental abnormalities in zebrafish. Developmental biology *339*, 89-100.

Infante, J.J., Law, G.L., and Young, E.T. (2012). Analysis of nucleosome positioning using a nucleosome-scanning assay. Chromatin Remodeling: Methods and Protocols, 63-87.

Ito, T., Bulger, M., Pazin, M.J., Kobayashi, R., and Kadonaga, J.T. (1997). ACF, an ISWI-containing and ATP-utilizing chromatin assembly and remodeling factor. Cell *90*, 145-155.

Jacobs, S.A., and Khorasanizadeh, S. (2002). Structure of HP1 chromodomain bound to a lysine 9-methylated histone H3 tail. Science *295*, 2080-2083.

Jiang, Y., Liu, M., Spencer, C.A., and Price, D.H. (2004). Involvement of transcription termination factor 2 in mitotic repression of transcription elongation. Molecular cell *14*, 375-386.

Jin, J., Cai, Y., Yao, T., Gottschalk, A.J., Florens, L., Swanson, S.K., Gutiérrez, J.L., Coleman, M.K., Workman, J.L., and Mushegian, A. (2005). A mammalian chromatin remodeling complex with similarities to the yeast INO80 complex. Journal of Biological Chemistry *280*, 41207-41212.

Jongmans, M.C., Hoefsloot, L.H., van der Donk, K.P., Admiraal, R.J., Magee, A., van de Laar, I., Hendriks, Y., Verheij, J.B., Walpole, I., and Brunner, H.G. (2008). Familial CHARGE

syndrome and the CHD7 gene: a recurrent missense mutation, intrafamilial recurrence and variability. American journal of medical genetics Part A *146*, 43-50.

Kagalwala, M.N., Glaus, B.J., Dang, W., Zofall, M., and Bartholomew, B. (2004). Topography of the ISW2–nucleosome complex: insights into nucleosome spacing and chromatin remodeling. The EMBO journal *23*, 2092-2104.

Kassabov, S.R., Zhang, B., Persinger, J., and Bartholomew, B. (2003). SWI/SNF unwraps, slides, and rewraps the nucleosome. Molecular cell *11*, 391-403.

Kehle, J., Beuchle, D., Treuheit, S., Christen, B., Kennison, J.A., Bienz, M., and Müller, J. (1998). dMi-2, a hunchback-interacting protein that functions in polycomb repression. Science *282*, 1897-1900.

Khammanit, R., Chantakru, S., Kitiyanant, Y., and Saikhun, J. (2008). Effect of serum starvation and chemical inhibitors on cell cycle synchronization of canine dermal fibroblasts. Theriogenology *70*, 27-34.

Khorasanizadeh, S. (2004). The nucleosome: from genomic organization to genomic regulation. Cell *116*, 259-272.

Kikin, O., D'Antonio, L., and Bagga, P.S. (2006). QGRS Mapper: a web-based server for predicting G-quadruplexes in nucleotide sequences. Nucleic acids research *34*, W676-W682.

Kilic, S.S., Donmez, O., Sloan, E.A., Elizondo, L.I., Huang, C., André, J.L., Bogdanovic, R., Cockfield, S., Cordeiro, I., and Deschenes, G. (2005). Association of migraine-like headaches with Schimke immuno-osseous dysplasia. American Journal of Medical Genetics Part A *135*, 206-210.

Kim, J.L., Morgenstern, K.A., Griffith, J.P., Dwyer, M.D., Thomson, J.A., Murcko, M.A., Lin, C., and Caron, P.R. (1998). Hepatitis C virus NS3 RNA helicase domain with a bound oligonucleotide: the crystal structure provides insights into the mode of unwinding. Structure *6*, 89-100.

Klein, H.L. (1997). RDH54, a RAD54 homologue in Saccharomyces cerevisiae, is required for mitotic diploid-specific recombination and repair and for meiosis. Genetics *147*, 1533-1543.

Knoepfler, P.S., and Eisenman, R.N. (1999). Sin meets NuRD and other tails of repression. Cell *99*, 447-450.

Kobor, M.S., Venkatasubrahmanyam, S., Meneghini, M.D., Gin, J.W., Jennings, J.L., Link, A.J., Madhani, H.D., and Rine, J. (2004). A protein complex containing the conserved Swi2/Snf2-related ATPase Swr1p deposits histone variant H2A. Z into euchromatin. PLoS Biol *2*, e131.

Kokavec, J., Podskocova, J., Zavadil, J., and Stopka, T. (2008). Chromatin remodeling and SWI/SNF2 factors in human disease. Front Biosci *13*, 6126-6134.

Kornberg, R.D. (1974). Chromatin structure: a repeating unit of histones and DNA. Science *184*, 868-871.

Korolev, S., Hsieh, J., Gauss, G.H., Lohman, T.M., and Waksman, G. (1997). Major domain swiveling revealed by the crystal structures of complexes of E. coli Rep helicase bound to single-stranded DNA and ADP. Cell *90*, 635-647.

Korolev, S., Lohman, T.M., Waksman, G., Yao, N., and Weber, P.C. (1998). Comparisons between the structures of HCV and Rep helicases reveal structural similarities between SF1 and SF2 super-families of helicases. Protein Science *7*, 605-610.

Krebs, J.E., Fry, C.J., Samuels, M.L., and Peterson, C.L. (2000). Global role for chromatin remodeling enzymes in mitotic gene expression. Cell *102*, 587-598.

Krebs, J.E., Kuo, M.-H., Allis, C.D., and Peterson, C.L. (1999). Cell cycle-regulated histone acetylation required for expression of the yeast HO gene. Genes & development *13*, 1412-1421.

Krogan, N.J., Keogh, M.-C., Datta, N., Sawa, C., Ryan, O.W., Ding, H., Haw, R.A., Pootoolal, J., Tong, A., and Canadien, V. (2003). A Snf2 family ATPase complex required for recruitment of the histone H2A variant Htz1. Molecular cell *12*, 1565-1576.

Krogh, B.O., and Symington, L.S. (2004). Recombination proteins in yeast. Annu Rev Genet *38*, 233-271.

Kulić, I., and Schiessel, H. (2003). Nucleosome repositioning via loop formation. Biophysical journal *84*, 3197-3211.

Kurdistani, S.K., and Grunstein, M. (2003). Histone acetylation and deacetylation in yeast. Nature reviews Molecular cell biology *4*, 276-284.

Kurdistani, S.K., Tavazoie, S., and Grunstein, M. (2004). Mapping global histone acetylation patterns to gene expression. Cell *117*, 721-733.

Kusch, T., Florens, L., MacDonald, W.H., Swanson, S.K., Glaser, R.L., Yates, J.R., Abmayr, S.M., Washburn, M.P., and Workman, J.L. (2004). Acetylation by Tip60 is required for selective histone variant exchange at DNA lesions. Science *306*, 2084-2087.

Lachner, M., O'carroll, D., Rea, S., Mechtler, K., and Jenuwein, T. (2001). Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins. Nature *410*, 116-120.

Lachner, M., O'Sullivan, R.J., and Jenuwein, T. (2003). An epigenetic road map for histone lysine methylation. Journal of cell science *116*, 2117-2124.

Längst, G., Bonte, E.J., Corona, D.F., and Becker, P.B. (1999). Nucleosome movement by CHRAC and ISWI without disruption or trans-displacement of the histone octamer. Cell *97*, 843-852.

Längst, G., and Manelyte, L. (2015). Chromatin remodelers: from function to dysfunction. Genes *6*, 299-324.

Laurent, B.C., Yang, X., and Carlson, M. (1992). An essential Saccharomyces cerevisiae gene homologous to SNF2 encodes a helicase-related protein in a new family. Molecular and cellular biology *12*, 1893-1902.

Lazzaro, M.A., and Picketts, D.J. (2001). Cloning and characterization of the murine Imitation Switch (ISWI) genes: differential expression patterns suggest distinct developmental roles for Snf2h and Snf2l. Journal of neurochemistry *77*, 1145-1156.

Leach, D.R. (1994). Long DNA palindromes, cruciform structures, genetic instability and secondary structure repair. Bioessays *16*, 893-900.

Lee, C.-H., Murphy, M.R., Lee, J.-S., and Chung, J.H. (1999). Targeting a SWI/SNF-related chromatin remodeling complex to the β-globin promoter in erythroid cells. Proceedings of the National Academy of Sciences *96*, 12311-12315.

Lemieux, K., Larochelle, M., and Gaudreau, L. (2008). Variant histone H2A. Z, but not the HMG proteins Nhp6a/b, is essential for the recruitment of Swi/Snf, Mediator, and SAGA to the yeast GAL1 UAS G. Biochemical and biophysical research communications *369*, 1103-1107.

Levens, D. (2008). How the c-myc promoter works and why it sometimes does not. Journal of the National Cancer Institute Monographs, 41.

Li, H., Ilin, S., Wang, W., Duncan, E.M., Wysocka, J., Allis, C.D., and Patel, D.J. (2006). Molecular basis for site-specific read-out of histone H3K4me3 by the BPTF PHD finger of NURF. Nature *442*, 91-95.

Li, J., and Broyles, S.S. (1995). The DNA binding domain of the vaccinia virus early transcription factor small subunit is an extended helicase-like motif. Nucleic acids research *23*, 1590-1596.

Liu, J., Gagnon, Y., Gauthier, J., Furenlid, L., L'Heureux, P.-J., Auger, M., Nureki, O., Yokoyama, S., and Lapointe, J. (1995). The zinc-binding site of Escherichia coli glutamyl-tRNA synthetase is located in the acceptor-binding domain. Journal of Biological Chemistry *270*, 15162-15169.

Liu, J., Kouzine, F., Nie, Z., Chung, H.J., Elisha-Feil, Z., Weber, A., Zhao, K., and Levens, D. (2006). The FUSE/FBP/FIR/TFIIH system is a molecular machine programming a pulse of c-myc expression. The EMBO journal *25*, 2119-2130.

Luger, K., Mäder, A.W., Richmond, R.K., Sargent, D.F., and Richmond, T.J. (1997). Crystal structure of the nucleosome core particle at 2.8 Å resolution. Nature *389*, 251-260.

Luo, R.X., and Dean, D.C. (1999). Chromatin remodeling and transcriptional regulation. Journal of the National Cancer Institute *91*, 1288-1294.

Marcu, K.B., Bossone, S.A., and Patel, A.J. (1992). Myc function and regulation. Annual review of biochemistry *61*, 809-858.

Marfella, C.G., and Imbalzano, A.N. (2007). The Chd family of chromatin remodelers. Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis *618*, 30-40.

Marintcheva, B., and Weller, S.K. (2003). Helicase motif Ia is involved in single-strand DNA-binding and helicase activities of the herpes simplex virus type 1 origin-binding protein, UL9. Journal of virology *77*, 2477-2488.

Marmorstein, R. (2001). Structure of histone deacetylases: insights into substrate recognition and catalysis. Structure *9*, 1127-1133.

Marmorstein, R., and Roth, S.Y. (2001). Histone acetyltransferases: function, structure, and catalysis. Current opinion in genetics & development *11*, 155-161.

Mason, A.C., Rambo, R.P., Greer, B., Pritchett, M., Tainer, J.A., Cortez, D., and Eichman, B.F. (2014). A structure-specific nucleic acid-binding domain conserved among DNA repair proteins. Proceedings of the National Academy of Sciences *111*, 7618-7623.

McBryant, S.J., Adams, V.H., and Hansen, J.C. (2006). Chromatin architectural proteins. Chromosome Research *14*, 39-51.

McDowell, T., Gibbons, R., Sutherland, H., O'Rourke, D., Bickmore, W., Pombo, A., Turley, H., Gatter, K., Picketts, D., and Buckle, V. (1999). Localization of a putative transcriptional regulator (ATRX) at pericentromeric heterochromatin and the short arms of acrocentric chromosomes. Proceedings of the National Academy of Sciences *96*, 13983-13988.

Medina, P.P., Carretero, J., Ballestar, E., Angulo, B., Lopez-Rios, F., Esteller, M., and Sanchez-Cespedes, M. (2005). Transcriptional targets of the chromatin-remodelling factor SMARCA4/BRG1 in lung cancer cells. Human molecular genetics *14*, 973-982.

Meyer-Lindenberg, A., Mervis, C.B., and Berman, K.F. (2006). Neural mechanisms in Williams syndrome: a unique window to genetic influences on cognition and behaviour. Nature Reviews Neuroscience *7*, 380-393.

Mirabella, A.C., Foster, B.M., and Bartke, T. (2016). Chromatin deregulation in disease. Chromosoma *125*, 75-93.

Mirkin, S.M. (2008). Discovery of alternative DNA structures: a heroic decade (1979–1989). Front Biosci *13*, 1064-1071.

Mitra, D., Parnell, E.J., Landon, J.W., Yu, Y., and Stillman, D.J. (2006). SWI/SNF binding to the HO promoter requires histone acetylation and stimulates TATA-binding protein recruitment. Molecular and cellular biology *26*, 4095-4110.

Mizuguchi, G., Shen, X., Landry, J., Wu, W.-H., Sen, S., and Wu, C. (2004). ATP-driven exchange of histone H2AZ variant catalyzed by SWR1 chromatin remodeling complex. Science *303*, 343-348.

Mohrmann, L., Langenberg, K., Krijgsveld, J., Kal, A.J., Heck, A.J., and Verrijzer, C.P. (2004). Differential targeting of two distinct SWI/SNF-related Drosophila chromatin-remodeling complexes. Molecular and cellular biology *24*, 3077-3088.

Moolenaar, G.F., Visse, R., Ortiz-Buysse, M., Goosen, N., and van de Putte, P. (1994). Helicase motifs V and VI of the Escherichia coli UvrB protein of the UvrABC endonuclease are essential for the formation of the preincision complex. Journal of molecular biology *240*, 294-307.

Morillon, A., Karabetsou, N., O'Sullivan, J., Kent, N., Proudfoot, N., and Mellor, J. (2003). Isw1 chromatin remodeling ATPase coordinates transcription elongation and termination by RNA polymerase II. Cell *115*, 425-435.

Morrison, A.J., Highland, J., Krogan, N.J., Arbel-Eden, A., Greenblatt, J.F., Haber, J.E., and Shen, X. (2004). INO80 and γ-H2AX interaction links ATP-dependent chromatin remodeling to DNA damage repair. Cell *119*, 767-775.

Muchardt, C., and Yaniv, M. (1993). A human homologue of Saccharomyces cerevisiae SNF2/SWI2 and Drosophila brm genes potentiates transcriptional activation by the glucocorticoid receptor. The EMBO Journal *12*, 4279.

Murawska, M., and Brehm, A. (2011). CHD chromatin remodelers and the transcription cycle. Transcription *2*, 244-253.

Murphy, D.J., Hardy, S., and Engel, D.A. (1999). Human SWI-SNF component BRG1 represses transcription of the c-fos gene. Molecular and cellular biology *19*, 2724-2733.

Muthuswami, R., Truman, P.A., Mesner, L.D., and Hockensmith, J.W. (2000). A eukaryotic SWI2/SNF2 domain, an exquisite detector of double-stranded to single-stranded DNA transition elements. Journal of Biological Chemistry *275*, 7648-7655.

Muzzin, O., Campbell, E.A., Xia, L., Severinova, E., Darst, S.A., and Severinov, K. (1998). Disruption of Escherichia coli hepA, an RNA polymerase-associated protein, causes UV sensitivity. Journal of Biological Chemistry *273*, 15157-15161.

Narlikar, G.J., Fan, H.-Y., and Kingston, R.E. (2002). Cooperation between complexes that regulate chromatin structure and transcription. Cell *108*, 475-487.

Narlikar, G.J., Sundaramoorthy, R., and Owen-Hughes, T. (2013). Mechanisms and functions of ATP-dependent chromatin-remodeling enzymes. Cell *154*, 490-503.

Natarajan, K., Jackson, B.M., Zhou, H., Winston, F., and Hinnebusch, A.G. (1999). Transcriptional activation by Gcn4p involves independent interactions with the SWI/SNF complex and the SRB/mediator. Molecular cell *4*, 657-664.

Neely, K.E., and Workman, J.L. (2002). The complexity of chromatin remodeling and its links to cancer. Biochimica et Biophysica Acta (BBA)-Reviews on Cancer *1603*, 19-29.

Neigeborn, L., and Carlson, M. (1984). Genes affecting the regulation of SUC2 gene expression by glucose repression in Saccharomyces cerevisiae. Genetics *108*, 845-858.

Neumann, H., Hancock, S.M., Buning, R., Routh, A., Chapman, L., Somers, J., Owen-Hughes, T., van Noort, J., Rhodes, D., and Chin, J.W. (2009). A method for genetically installing site-

specific acetylation in recombinant histones defines the effects of H3 K56 acetylation. Molecular cell *36*, 153-163.

Nongkhlaw, M., Dutta, P., Hockensmith, J.W., Komath, S.S., and Muthuswami, R. (2009). Elucidating the mechanism of DNA-dependent ATP hydrolysis mediated by DNA-dependent ATPase A, a member of the SWI2/SNF2 protein family. Nucleic acids research *37*, 3332-3341.

Nongkhlaw, M., Gupta, M., Komath, S.S., and Muthuswami, R. (2012). Motifs Q and I are required for ATP hydrolysis but not for ATP binding in SWI2/SNF2 proteins. Biochemistry *51*, 3711-3722.

Nowak, S.J., and Corces, V.G. (2000). Phosphorylation of histone H3 correlates with transcriptionally active loci. Genes & development *14*, 3003-3013.

Nowak, S.J., and Corces, V.G. (2004). Phosphorylation of histone H3: a balancing act between chromosome condensation and transcriptional activation. TRENDS in Genetics *20*, 214-220.

Oudet, P., Gross-Bellard, M., and Chambon, P. (1975). Electron microscopic and biochemical evidence that chromatin structure is a repeating unit. Cell *4*, 281-300.

Ouspenski, I.I., Elledge, S.J., and Brinkley, B. (1999). New yeast genes important for chromosome integrity and segregation identified by dosage effects on genome stability. Nucleic acids research *27*, 3001-3008.

Pause, A., Méthot, N., and Sonenberg, N. (1993). The HRIGRXXR region of the DEAD box RNA helicase eukaryotic translation initiation factor 4A is required for RNA binding and ATP hydrolysis. Molecular and Cellular Biology *13*, 6789-6798.

Pause, A., and Sonenberg, N. (1992). Mutational analysis of a DEAD box RNA helicase: the mammalian translation initiation factor eIF-4A. The EMBO journal *11*, 2643.

Peterson, C.L., and Herskowitz, I. (1992). Characterization of the yeast SWI1, SWI2, and SWI3 genes, which encode a global activator of transcription. Cell *68*, 573-583.

Phillips, D., and Johns, E. (1965). A fractionation of the histones of group F2a from calf thymus. Biochemical Journal *94*, 127.

Pickart, C.M. (2004). Back to the future with ubiquitin. Cell *116*, 181-190.

Plambeck, C.A., Kwan, A.H., Adams, D.J., Westman, B.J., van der Weyden, L., Medcalf, R.L., Morris, B.J., and Mackay, J.P. (2003). The structure of the zinc finger domain from human splicing factor ZNF265 fold. Journal of Biological Chemistry *278*, 22805-22811.

Poole, L.A., Zhao, R., Glick, G.G., Lovejoy, C.A., Eischen, C.M., and Cortez, D. (2015). SMARCAL1 maintains telomere integrity during DNA replication. Proceedings of the National Academy of Sciences *112*, 14864-14869.

Qin, Y., Rezler, E.M., Gokhale, V., Sun, D., and Hurley, L.H. (2007). Characterization of the G-quadruplexes in the duplex nuclease hypersensitive element of the PDGF-A promoter and modulation of PDGF-A promoter activity by TMPyP4. Nucleic acids research *35*, 7698-7713.

Racki, L.R., Naber, N., Pate, E., Leonard, J.D., Cooke, R., and Narlikar, G.J. (2014). The histone H4 tail regulates the conformation of the ATP-binding pocket in the SNF2h chromatin remodeling enzyme. Journal of molecular biology *426*, 2034-2044.

Richmond, T.J., and Davey, C.A. (2003). The structure of DNA in the nucleosome core. Nature *423*, 145-150.

Robinson, P., and Ardley, H. (2004). Ubiquitin-protein ligases. Journal of Cell Science *117*, 5191-5194.

Robinson, P.J., Fairall, L., Huynh, V.A., and Rhodes, D. (2006). EM measurements define the dimensions of the "30-nm" chromatin fiber: evidence for a compact, interdigitated structure. Proceedings of the National Academy of Sciences *103*, 6506-6511.

Rogers, G.W., Komar, A.A., and Merrick, W.C. (2002). eIF4A: the godfather of the DEAD box helicases. Progress in nucleic acid research and molecular biology *72*, 307-331.

Rossetto, D., Avvakumov, N., and Côté, J. (2012). Histone phosphorylation: a chromatin modification involved in diverse nuclear events. Epigenetics *7*, 1098-1108.

Roth, S.Y., Denu, J.M., and Allis, C.D. (2001). Histone acetyltransferases. Annual review of biochemistry *70*, 81-120.

Rothbart, S.B., and Strahl, B.D. (2014). Interpreting the language of histone and DNA modifications. Biochimica et Biophysica Acta (BBA)-Gene Regulatory Mechanisms *1839*, 627-643.

Ryan, M.P., Jones, R., and Morse, R.H. (1998). SWI-SNF complex participation in transcriptional activation at a step subsequent to activator binding. Molecular and cellular biology *18*, 1774-1782.

Rydberg, B., Holley, W.R., Mian, I.S., and Chatterjee, A. (1998). Chromatin conformation in living cells: support for a zig-zag model of the 30 nm chromatin fiber. Journal of molecular biology *284*, 71-84.

Santangelo, L., Gigante, M., Netti, G.S., Diella, S., Puteo, F., Carbone, V., Grandaliano, G., Giordano, M., and Gesualdo, L. (2014). A novel SMARCAL1 mutation associated with a mild phenotype of Schimke immuno-osseous dysplasia (SIOD). BMC nephrology *15*, 41.

Santos-Rosa, H., and Caldas, C. (2005). Chromatin modifier enzymes, the histone code and cancer. European Journal of Cancer *41*, 2381-2402.

Sawicka, A., and Seiser, C. (2012). Histone H3 phosphorylation–a versatile chromatin modification for different occasions. Biochimie *94*, 2193-2201.

Schnitzler, G., Sif, S., and Kingston, R.E. (1998). Human SWI/SNF interconverts a nucleosome between its base state and a stable remodeled state. Cell *94*, 17-27.

Schoor, M., Schuster-Gossler, K., Roopenian, D., and Gossler, A. (1999). Skeletal dysplasias, growth retardation, reduced postnatal survival, and impaired fertility in mice lacking the SNF2/SWI2 family member ETL1. Mechanisms of development *85*, 73-83.

Sharma, T. (2016). Elucidating the role of SMARCAL1 in cell cycle regulation and transcription.

Sharma, T., Bansal, R., Haokip, D. T., Goel, I., & Muthuswami, R. (2015). SMARCAL1 negatively regulates c-Myc transcription by altering the conformation of the promoter region. Scientific reports, *5*.

Shaw, G., Gan, J., Zhou, Y.N., Zhi, H., Subburaman, P., Zhang, R., Joachimiak, A., Jin, D.J., and Ji, X. (2008). Structure of RapA, a Swi2/Snf2 protein that recycles RNA polymerase during transcription. Structure *16*, 1417-1427.

Shen, X., Mizuguchi, G., Hamiche, A., and Wu, C. (2000). A chromatin remodelling complex involved in transcription and DNA processing. Nature *406*, 541-544.

Shiio, Y., and Eisenman, R.N. (2003). Histone sumoylation is associated with transcriptional repression. Proceedings of the National Academy of Sciences *100*, 13225-13230.

Shogren-Knaak, M., Ishii, H., Sun, J.-M., Pazin, M.J., Davie, J.R., and Peterson, C.L. (2006). Histone H4-K16 acetylation controls chromatin structure and protein interactions. Science *311*, 844-847.

Shogren-Knaak, M., and Peterson, C.L. (2006). Switching on chromatin: mechanistic role of histone H4-K16 acetylation. Cell cycle *5*, 1361-1365.

Siddiqui-Jain, A., Grand, C.L., Bearss, D.J., and Hurley, L.H. (2002). Direct evidence for a G-quadruplex in a promoter region and its targeting with a small molecule to repress c-MYC transcription. Proceedings of the National Academy of Sciences *99*, 11593-11598.

Sif, S., Saurin, A.J., Imbalzano, A.N., and Kingston, R.E. (2001). Purification and characterization of mSin3A-containing Brg1 and hBrm chromatin remodeling complexes. Genes & development *15*, 603-618.

Sinha, K.M., Glickman, M.S., and Shuman, S. (2009). Mutational analysis of Mycobacterium UvrD1 identifies functional groups required for ATP hydrolysis, DNA unwinding, and chemomechanical coupling. Biochemistry *48*, 4019-4030.

Stern, M., Jensen, R., and Herskowitz, I. (1984). Five SWI genes are required for expression of the HO gene in yeast. Journal of molecular biology *178*, 853-868.

Sternberg, P.W., Stern, M.J., Clark, I., and Herskowitz, I. (1987). Activation of the yeast HO gene by release from multiple negative controls. Cell *48*, 567-577.

Stopka, T., and Skoultchi, A.I. (2003). The ISWI ATPase Snf2h is required for early mouse development. Proceedings of the National Academy of Sciences *100*, 14097-14102.

Story, R.M., and Steitz, T.A. (1992). Structure of the recA protein-ADP complex. Nature *355*, 374.

Strahl, B.D., and Allis, C.D. (2000). The language of covalent histone modifications. Nature *403*, 41-45.

Strick, R., Strissel, P.L., Gavrilov, K., and Levi-Setti, R. (2001). Cation–chromatin binding as shown by ion microscopy is essential for the structural integrity of chromosomes. J Cell Biol *155*, 899-910.

Strohner, R., Nemeth, A., Jansa, P., Hofmann-Rohrer, U., Santoro, R., Längst, G., and Grummt, I. (2001). NoRC—a novel member of mammalian ISWI-containing chromatin remodeling machines. The EMBO Journal *20*, 4892-4900.

Subramanya, H.S., Bird, L.E., Brannigan, J.A., and Wigley, D.B. (1996). Crystal structure of a DExx box DNA helicase. Nature *384*, 379.

Sudarsanam, P., and Winston, F. (2000). The Swi/Snf family: nucleosome-remodeling complexes and transcriptional control. TRENDS in Genetics *16*, 345-351.

Sukhodolets, M.V., Cabrera, J.E., Zhi, H., and Jin, D.J. (2001). RapA, a bacterial homolog of SWI2/SNF2, stimulates RNA polymerase recycling in transcription. Genes & development *15*, 3330-3341.

Sullivan, B.A., and Karpen, G.H. (2004). Centromeric chromatin exhibits a histone modification pattern that is distinct from both euchromatin and heterochromatin. Nature structural & molecular biology *11*, 1076-1083.

Sun, Z.-W., and Allis, C.D. (2002). Ubiquitination of histone H2B regulates H3 methylation and gene silencing in yeast. Nature *418*, 104-108.

Sung, P., Higgins, D., Prakash, L., and Prakash, S. (1988). Mutation of lysine-48 to arginine in the yeast RAD3 protein abolishes its ATPase and DNA helicase activities but not the ability to bind ATP. The EMBO Journal *7*, 3263.

Suto, R.K., Edayathumangalam, R.S., White, C.L., Melander, C., Gottesfeld, J.M., Dervan, P.B., and Luger, K. (2003). Crystal structures of nucleosome core particles in complex with minor groove DNA-binding ligands. Journal of molecular biology *326*, 371-380.

Svejstrup, J.Q. (2003). Rescue of arrested RNA polymerase II complexes. Journal of Cell Science *116*, 447-451.

Syed, S.H., Boulard, M., Shukla, M.S., Gautier, T., Travers, A., Bednar, J., Faivre-Moskalenko, C., Dimitrov, S., and Angelov, D. (2009). The incorporation of the novel histone variant H2AL2 confers unusual structural and functional properties of the nucleosome. Nucleic acids research *37*, 4684-4695.

Tachibana, M., Nozaki, M., Takeda, N., and Shinkai, Y. (2007). Functional dynamics of H3K9 methylation during meiotic prophase progression. The EMBO journal *26*, 3346-3359.

Tamaru, H., and Selker, E.U. (2001). A histone H3 methyltransferase controls DNA methylation in Neurospora crassa. Nature *414*, 277-283.

Tamkun, J.W., Deuring, R., Scott, M.P., Kissinger, M., Pattatucci, A.M., Kaufman, T.C., and Kennison, J.A. (1992). brahma: A regulator of Drosophila homeotic genes structurally related to the yeast transcriptional activator SNF2SWI2. Cell *68*, 561-572.

Tan, T.R., Kanaar, R., and Wyman, C. (2003). Rad54, a Jack of all trades in homologous recombination. DNA repair *2*, 787-794.

Tanner, N.K. (2003). The Newly Identified Q Motif of DEAD Box Heicases IS Involved in Adenine Recognition. Cell cycle *2*, 18-19.

Tanner, N.K., Cordin, O., Banroques, J., Doere, M., and Linder, P. (2003). The Q motif: a newly identified motif in DEAD box helicases may regulate ATP binding and hydrolysis. Molecular cell *11*, 127-138.

Tanner, N.K., and Linder, P. (2001). DExD/H box RNA helicases: from generic motors to specific dissociation functions. Molecular cell *8*, 251-262.

Taunton, J., Hassig, C.A., and Schreiber, S.L. (1996). A mammalian histone deacetylase related to the yeast transcriptional regulator Rpd3p. Science *272*, 408.

Teng, D.H.-F., Chen, Y., Lian, L., Ha, P.C., Tavtigian, S.V., and Wong, A.K. (2001). Mutation analyses of 268 candidate genes in human tumor cell lines. Genomics *74*, 352-364.

Theis, K., Chen, P.J., Skorvaga, M., Van Houten, B., and Kisker, C. (1999). Crystal structure of UvrB, a DNA helicase adapted for nucleotide excision repair. The EMBO journal *18*, 6899-6907.

Thomä, N.H., Czyzewski, B.K., Alexeev, A.A., Mazin, A.V., Kowalczykowski, S.C., and Pavletich, N.P. (2005). Structure of the SWI2/SNF2 chromatin-remodeling domain of eukaryotic Rad54. Nature structural & molecular biology *12*, 350-356.

Thomas, J.O., and Furber, V. (1976). Yeast chromatin structure. FEBS letters *66*, 274-280.

Thomas, J.O., and Kornberg, R.D. (1975). An octamer of histones in chromatin and free in solution. Proceedings of the National Academy of Sciences *72*, 2626-2630.

Tremethick, D.J. (2007). Higher-order structures of chromatin: the elusive 30 nm fiber. Cell *128*, 651-654.

Trouche, D., Le Chalony, C., Muchardt, C., Yaniv, M., and Kouzarides, T. (1997). RB and hbrm cooperate to repress the activation functions of E2F1. Proceedings of the National Academy of Sciences *94*, 11268-11273.

Tsukiyama, T. (2002). The in vivo functions of ATP-dependent chromatin-remodelling factors. Nature Reviews Molecular Cell Biology *3*, 422-429.

Tsukiyama, T., Daniel, C., Tamkun, J., and Wu, C. (1995). ISWI, a member of the SWl2/SNF2 ATPase family, encodes the 140 kDa subunit of the nucleosome remodeling factor. Cell *83*, 1021-1026.

Tsukiyama, T., Palmer, J., Landel, C.C., Shiloach, J., and Wu, C. (1999). Characterization of the imitation switch subfamily of ATP-dependent chromatin-remodeling factors in Saccharomyces cerevisiae. Genes & development *13*, 686-697.

Tsukiyama, T., and Wu, C. (1995). Purification and properties of an ATP-dependent nucleosome remodeling factor. Cell *83*, 1011-1020.

Tsurusaki, Y., Okamoto, N., Ohashi, H., Kosho, T., Imai, Y., Hibi-Ko, Y., Kaname, T., Naritomi, K., Kawame, H., and Wakui, K. (2012). Mutations affecting components of the SWI/SNF complex cause Coffin-Siris syndrome. Nature genetics *44*, 376-378.

Turner, B.M. (2002). Cellular memory and the histone code. Cell *111*, 285-291.

Udugama, M., Sabri, A., and Bartholomew, B. (2011). The INO80 ATP-dependent chromatin remodeling complex is a nucleosome spacing factor. Molecular and cellular biology *31*, 662-673.

Valdés-Mora, F., Song, J.Z., Statham, A.L., Strbenac, D., Robinson, M.D., Nair, S.S., Patterson, K.I., Tremethick, D.J., Stirzaker, C., and Clark, S.J. (2012). Acetylation of H2A. Z is a key epigenetic modification associated with gene deregulation and epigenetic remodeling in cancer. Genome research *22*, 307-321.

van Attikum, H., Fritsch, O., Hohn, B., and Gasser, S.M. (2004). Recruitment of the INO80 complex by H2A phosphorylation links ATP-dependent chromatin remodeling with DNA double-strand break repair. Cell *119*, 777-788.

van Holde, K., and Zlatanova, J. (1995). Chromatin higher order structure: chasing a mirage? Journal of Biological Chemistry *270*, 8373-8376.

Varga-Weisz, P.D., Wilm, M., Bonte, E., Dumas, K., Mann, M., and Becker, P.B. (1997). Chromatin-remodelling factor CHRAC contains the ATPases ISWI and topoisomerase II. Nature *388*, 598-602.

Velankar, S.S., Soultanas, P., Dillingham, M.S., Subramanya, H.S., and Wigley, D.B. (1999). Crystal structures of complexes of PcrA DNA helicase with a DNA substrate indicate an inchworm mechanism. Cell *97*, 75-84.

Vissers, L.E., van Ravenswaaij, C.M., Admiraal, R., Hurst, J.A., de Vries, B.B., Janssen, I.M., van der Vliet, W.A., Huys, E.H., de Jong, P.J., and Hamel, B.C. (2004). Mutations in a new member of the chromodomain gene family cause CHARGE syndrome. Nature genetics *36*, 955-957.

Voineagu, I., Narayanan, V., Lobachev, K.S., and Mirkin, S.M. (2008). Replication stalling at unstable inverted repeats: interplay between DNA hairpins and fork stabilizing proteins. Proceedings of the National Academy of Sciences *105*, 9936-9941.

Wan, Y., Saleem, R.A., Ratushny, A.V., Roda, O., Smith, J.J., Lin, C.-H., Chiang, J.-H., and Aitchison, J.D. (2009). Role of the histone variant H2A. Z/Htz1p in TBP recruitment, chromatin dynamics, and regulated expression of oleate-responsive genes. Molecular and cellular biology *29*, 2346-2358.

Wassarman, D.A., and Steitz, J.A. (1991). Alive with DEAD proteins. Nature *349*, 463-464.

Weng, Y., Czaplinski, K., and Peltz, S.W. (1996). Genetic and biochemical characterization of mutations in the ATPase and helicase regions of the Upf1 protein. Molecular and cellular biology *16*, 5477-5490.

White, P.S., Thompson, P.M., Gotoh, T., Okawa, E.R., Igarashi, J., Kok, M., Winter, C., Gregory, S.G., Hogarty, M.D., and Maris, J.M. (2005). Definition and characterization of a region of 1p36. 3 consistently deleted in neuroblastoma. Oncogene *24*, 2684-2694.

Wilkinson, K.A., and Henley, J.M. (2010). Mechanisms, regulation and consequences of protein SUMOylation. Biochemical Journal *428*, 133-145.

Wilson, C.J., Chao, D.M., Imbalzano, A.N., Schnitzler, G.R., Kingston, R.E., and Young, R.A. (1996). RNA polymerase II holoenzyme contains SWI/SNF regulators involved in chromatin remodeling. Cell *84*, 235-244.

Winston, F., and Allis, C.D. (1999). The bromodomain: a chromatin-targeting module? Nature structural biology *6*, 601-604.

Winston, F., and Carlson, M. (1992). Yeast SNF/SWI transcriptional activators and the SPT/SIN chromatin connection. Trends in Genetics *8*, 387-391.

Wong, A.K., Shanahan, F., Chen, Y., Lian, L., Ha, P., Hendricks, K., Ghaffari, S., Iliev, D., Penn, B., and Woodland, A.-M. (2000). BRG1, a component of the SWI-SNF complex, is mutated in multiple human tumor cell lines. Cancer research *60*, 6171-6177.

Woodage, T., Basrai, M.A., Baxevanis, A.D., Hieter, P., and Collins, F.S. (1997). Characterization of the CHD family of proteins. Proceedings of The National Academy of Sciences *94*, 11472-11477.

Woodcock, C.L., and Dimitrov, S. (2001). Higher-order structure of chromatin and chromosomes. Current opinion in genetics & development *11*, 130-135.

Workman, J., and Kingston, R. (1998). Alteration of nucleosome structure as a mechanism of transcriptional regulation. Annual review of biochemistry *67*, 545-579.

Wysocka, J., Swigut, T., Xiao, H., Milne, T.A., Kwon, S.Y., Landry, J., Kauer, M., Tackett, A.J., Chait, B.T., and Badenhorst, P. (2006). A PHD finger of NURF couples histone H3 lysine 4 trimethylation with chromatin remodelling. Nature *442*, 86-90.

Xu, F., Zhang, K., and Grunstein, M. (2005). Acetylation in histone H3 globular domain regulates gene expression in yeast. Cell *121*, 375-385.

Yan, Z., Wang, Z., Sharova, L., Sharov, A.A., Ling, C., Piao, Y., Aiba, K., Matoba, R., Wang, W., and Ko, M.S. (2008). BAF250B-Associated SWI/SNF Chromatin-Remodeling Complex Is Required to Maintain Undifferentiated Mouse Embryonic Stem Cells. Stem cells *26*, 1155-1165.

Yang, D., and Hurley, L.H. (2006). Structure of the biologically relevant G-quadruplex in the c-MYC promoter. Nucleosides, Nucleotides, and Nucleic Acids *25*, 951-968.

Yang, J.G., Madrid, T.S., Sevastopoulos, E., and Narlikar, G.J. (2006). The chromatin-remodeling enzyme ACF is an ATP-dependent DNA length sensor that regulates nucleosome spacing. Nature structural & molecular biology *13*, 1078-1083.

Yao, N., Hesson, T., Cable, M., Hong, Z., Kwong, A., Le, H., and Weber, P.C. (1997). Structure of the hepatitis C virus RNA. Nat Struct Biol *4*, 463-467.

Yuan, J., Ghosal, G., and Chen, J. (2009). The annealing helicase HARP protects stalled replication forks. Genes & development *23*, 2394-2399.

Yusufzai, T., and Kadonaga, J.T. (2008). HARP is an ATP-driven annealing helicase. Science *322*, 748-750.

Zhang, X., Yang, Z., Khan, S.I., Horton, J.R., Tamaru, H., Selker, E.U., and Cheng, X. (2003). Structural basis for the product specificity of histone lysine methyltransferases. Molecular cell *12*, 177-185.

Zhang, Y. (2003). Transcriptional regulation by histone ubiquitination and deubiquitination. Genes & development *17*, 2733-2740.

Zhang, Y., Smith, C.L., Saha, A., Grill, S.W., Mihardja, S., Smith, S.B., Cairns, B.R., Peterson, C.L., and Bustamante, C. (2006). DNA translocation and loop formation mechanism of chromatin remodeling by SWI/SNF and RSC. Molecular cell *24*, 559-568.

Zhao, J., Bacolla, A., Wang, G., and Vasquez, K.M. (2010). Non-B DNA structure-induced genetic instability and evolution. Cellular and Molecular Life Sciences *67*, 43-62.

Zofall, M., Persinger, J., and Bartholomew, B. (2004). Functional role of extranucleosomal DNA and the entry site of the nucleosome in chromatin remodeling by ISW2. Molecular and cellular biology *24*, 10047-10057.

APPENDIX-I

Site-directed mutagenesis was performed to make a make a Cys-less His-ADAAD protein for FRET experiment

All the conserved cysteine residues were mutated in ATPase core domain of ADAAD except a single cysteine residue and termed as B, C, D and J (Figure 6.1).
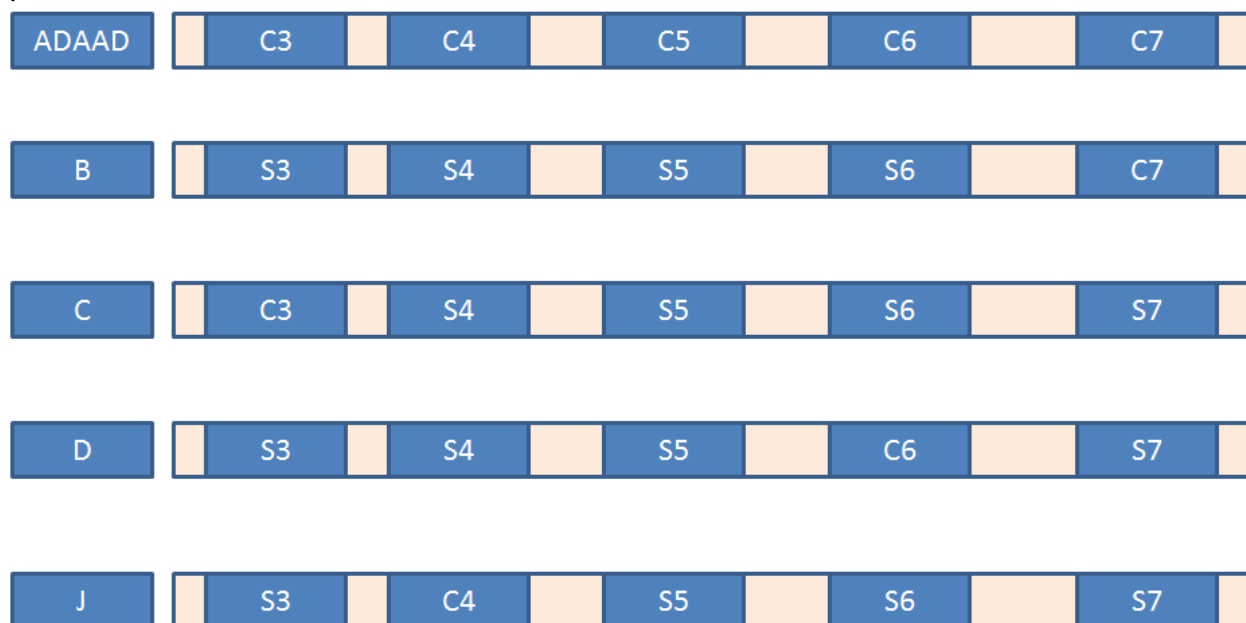
Figure 6.1: Diagrammatic representation of Cys residues present in the ATPase domain of ADAAD. B, C, D and J are the mutants where all the cysteine residues are mutated except one.

All cys mutants, thus generated, could be potentially used for FRET experiment. However, all these mutants were found to be ATPase inactive and hence, could not be used for IANS labeling and FRET experiment (Figure 6.2).
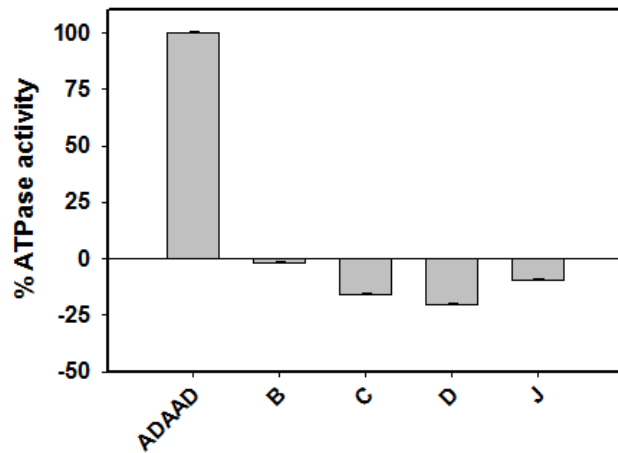
Figure 6.2 The ATPase activity of all cys mutants in the presence of stem-loop DNA as compared to the wild-type ADAAD. Each analysis represents the average of two independent experiments, whose standard deviations are shown in the figures. The protein concentration used in these analyses was 0.1 µM.

APPENDIX-II

Luria Bertani (LB) Broth (1000 ml):
Casein Enzyme Hydrolysate- 10.0 g
Yeast Extract- 5 g
NaCl – 10.0 g
pH -7
Alkaline Lysis Solution I:
50 mM glucose
25 mM Tris-Cl (pH 8.0)
10 mM EDTA (pH 8.0)
Storage-4^{0}C
Alkaline Lysis Solution II:
0.2 N NaOH
1% (w/v) SDS
Always prepare fresh
Alkaline Lysis Solution III (100 ml):
5 M Potassium Acetate- 60 ml
Glacial Acetic Acid- 11.5 ml
Distilled Water - 28.5 ml
Storage-4^{0}C
1X Tris EDTA Buffer (TE Buffer):
10 mM Tris-Cl (pH 8.0)
1 mM EDTA (pH 8.0)
10X TAE Buffer (1000ml)
48.4 g Tris Base
11.42 ml Glacial Acetic Acid
20.0 ml 0.5 EDTA
pH - 8.0
Taq Polymerase Buffer A (10X)
100.0 mM Tris HCL (pH – 9.0)
500.0 mM KCL
15.0 mM MgCl2
0.1 % Gelatin
DNA Loading Dye (6X):
0.25% (w/v) Bromophenol Blue
30% (v/v) Glycerol
SDS PAGE Running Buffer (1000ml)(pH – 8.3)
3.0 g Tris Base
14.4 g Glycine
1.0 g SDS
SDS-PAGE loading dye (4X):
40% (v/v) Glycerol
200 mM Tris-Cl (pH 6.8)
400 mM ß-me
8% (w/v) SDS
0.4% (w/v) Bromophenol Blue
Stacking Gel composition (5%):

5% Acrylamide (29:1 Acrylamide: Bis-Acryamide)
0.1% SDS
0.1% Ammonium persulphate
0.126 M Tris-Cl (pH 6.8)
TEMED (3 μl)
Resolving gel (10%):
10% Acrylamide (29:1 Acrylamide: Bis-acryamide)
0.1% SDS
0.1% Ammonium persulphate
0.375 M Tris-Cl (pH 8.8)
TEMED (4 μl)
Staining solution (100 ml):
45 ml Methanol
10 ml Glacial acetic acid
45 ml Water
Coomassie brilliant blue (R250) - 0.25 g
Destaining solution (100 ml):
45 ml Methanol
10 ml Glacial acetic acid
45 ml Water
Ammonium sulfate salt cut:

$$Y = \frac{xG\ (S_2-S_1)}{1 - (\upsilon G/1000)\ S_2}$$

where,
Y = amount of Ammonium sulfate (g)
x = volume of supernatant in L
S_2 = final % of salt
S_1 = initial % of salt
G = 514.72
υ = 0.5262
Ultracentrifugation buffer:
50 mM Tris-Cl (pH 8.0)
2 M NaCl
150 mM $MgCl_2$
5 mM ß-me
Equilibration Buffer (GST-ADAAD):
50 mM Tris-Cl (pH 8.0)
150 mM NaCl
150 mM $MgCl_2$
0.5 mM PMSF
5 mM ß-me
5% (v/v) Glycerol
Cleavage Buffer (GST-ADAAD):
25 mM Tris-Cl (pH 7.5)
150 mM NaCl

1 mM EDTA
1 mM DTT
REG Buffer (1X) for Fluorescence:
50 mM Tris-SO$_4$ (pH 7.5)
5mM ß-me
5mM MgSO$_4$
REG Buffer (1X) for ATPase Assay:
50 mM Tris-SO$_4$ (pH 7.5)
1 mM MgSO$_4$
5 mM ß-me
30 µg/ml Pyruvate Kinase
2 mM Phoshoenolpyruvate
5% Glycerol
REG Buffer (1X) for Circular dichroism
20 mM Tris-HCl (pH 7.5)
100 mM NaCl
1 mM EDTA
1mM MgCl$_2$
1 mM DTT
Phosphate Buffer Saline (PBS) pH-7.2 (1000mL)
8.5 g NaCl
0.35 g Na$_2$HPO$_4$
1.325 g NaH$_2$PO$_4$
Western-Transfer Buffer
14.4 g Glycine
3.0 g Tris Base
0.375 g SDS
200.0 ml Methanol

PUBLICATIONS

SCIENTIFIC REPORTS

OPEN

SMARCAL1 Negatively Regulates C-*Myc* Transcription By Altering The Conformation Of The Promoter Region

Tapan Sharma, Ritu Bansal, Dominic Thangminlen Haokip, Isha Goel & Rohini Muthuswami

SMARCAL1, a member of the SWI2/SNF2 protein family, stabilizes replication forks during DNA damage. In this manuscript, we provide the first evidence that SMARCAL1 is also a transcriptional co-regulator modulating the expression of c-Myc, a transcription factor that regulates 10–15% genes in the human genome. BRG1, SMARCAL1 and RNAPII were found localized onto the c-*myc* promoter. When HeLa cells were serum starved, the occupancy of SMARCAL1 on the c-*myc* promoter increased while that of BRG1 and RNAPII decreased correlating with repression of c-*myc* transcription. Using Active DNA-dependent ATPase A Domain (ADAAD), the bovine homolog of SMARCAL1, we show that the protein can hydrolyze ATP using a specific region upstream of the CT element of the c-*myc* promoter as a DNA effector. The energy, thereby, released is harnessed to alter the conformation of the promoter DNA. We propose that SMARCAL1 negatively regulates c-*myc* transcription by altering the conformation of its promoter region during differentiation.

ATP-dependent chromatin remodeling proteins regulate gene expression either by repositioning nucleosomes or by incorporating histone variants into the nucleosomes[1–3]. Baradaran-Heravi *et al.* have proposed a third possibility wherein they postulated that SMARCAL1, a distant member of the ATP-dependent chromatin remodeling protein family, could be regulating transcription of genes such as c-*myc* and c-*kit* by altering DNA structure in an ATP-dependent manner[4]. SMARCAL1 is a 105-kDa protein that hydrolyses ATP only in the presence of DNA molecules containing double-strand to single-strand transition regions[5–8]. *In vivo*, upon DNA damage the protein is recruited by RPA to stalled replication forks[9,10]. SMARCAL1 stabilizes the stalled replication forks due to its annealing helicase activity, thus playing an important role in maintaining genome stability[9–14]. Mutations in SMARCAL1 have been linked to the autosomal recessive disorder, Schimke Immunoosseous Dysplasia (SIOD)[15]. Patients afflicted with SIOD exhibit a wide range of phenotypes including skeletal dysplasia, T-cell immunodeficiency and renal dysfunction leading Boerkoel *et al.* to hypothesize that SMARCAL1 could be a transcriptional regulator of a subset of genes, both during and after development[15,16]. Experiments using zebra fish model system have shown that knockdown of *SMARCAL1* causes multi-system developmental abnormalities affecting gene expression of *gata1*, *beta-E1 globin* and other genes[17]. Recent studies have also shown that gene expression profile is altered in *SMARCAL1*−/− mice, supporting the hypothesis that the protein could function as a transcription regulator[4].

The transcription factor c-*myc*, a leucine zipper protein, regulates the expression of 10–15% of human genes, thus playing an important role in cell proliferation, differentiation, growth and survival; overexpression of the protein is associated with cancer[18–21]. The c-*myc* gene is exquisitely controlled and its expression is fine-tuned by many transcription factors[22]. The gene contains multiple promoters; in human cells four promoters have been documented: P0, P1, P2, and P3 with P2 being the maximally used promoter[21,23]. A GC-rich region, known as CT element, present −142 to −115 bp upstream of the P1 promoter, is the major regulator of c-*myc* expression by the formation of G-quadruplex and I-motif[24–26]. In addition to the CT element, a Far UpStream Element (FUSE) present 1.7 kb upstream of the P1 promoter has also been identified[27]. BRG1, an ATP-dependent chromatin remodeling protein, has been shown to remodel the nucleosomes around the FUSE region when cells are released from serum starvation[28,29].

School of Life Sciences, JNU, New Delhi 110067. Correspondence and requests for materials should be addressed to R.M. (email: rohini_m@mail.jnu.ac.in)
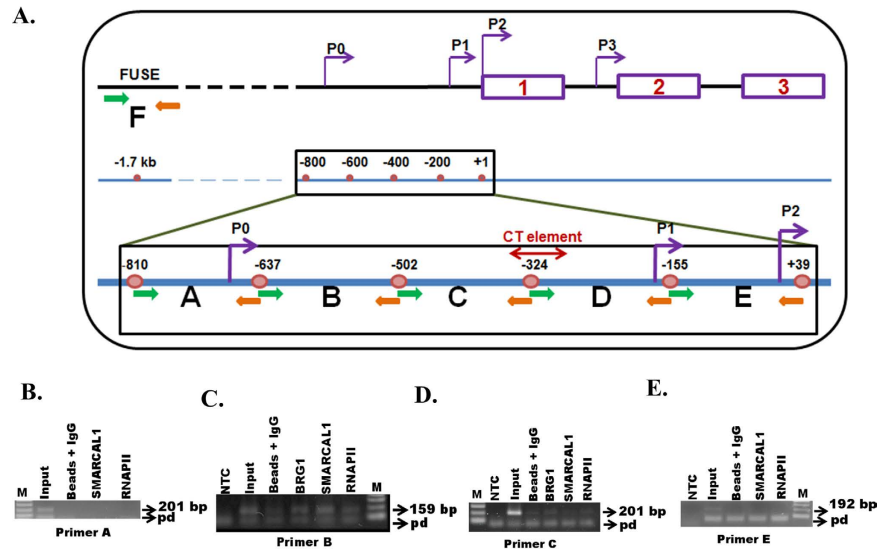
Figure 1. Analysis of occupancy of BRG1, SMARCAL1, and RNAPII on c-*myc* promoter. (**A**). Schematic representation of c-*myc* promoter showing promoters, P0, P1, P2, and P3 along with the CT element and FUSE region. The ChIP primers designed for analyzing the occupancy of BRG1, SMARCAL1 and RNAPII are also indicated in the figure. (**B**). Occupancy of SMARCAL1 and RNAPII at primer A region. (**C**). at primer B region, (**D**). at primer C region, and (**E**). at primer E region. The PCR products were resolved on 1% agarose gel. The primer-dimer formation is indicated by pd.

In this paper, we have explored the role of BRG1 and SMARCAL1 in regulating the expression of c-*myc*. We have shown that both BRG1 and SMARCAL1 bind to a 159 bp DNA segment upstream of the CT element which will be referred to as Myc_B_{159} in the remaining manuscript. Activation of c-*myc* gene was dependent on binding of BRG1 and RNA polymerase II (RNAPII) to Myc_B_{159}. In contrast, binding of SMARCAL1 to this region of the c-*myc* promoter led to repression of c-*myc* transcription. Using ADAAD, the bovine homolog of SMARCAL1, we have shown that ADAAD binds to Myc_B_{159} with an apparent K_M of 3.6 ± 0.3 nM. CD spectroscopy showed that ADAAD-Myc_B_{159} interaction results in alteration in the conformation of DNA in an ATP-dependent manner. We found that SMARCAL1 regulates differentiation of K562 cells in response to phorbol myristate acetate (PMA) by transcriptionally repressing c-*myc* expre transcriptionally repressing c-myc expression leading us to leading us to propose that the phenotypic manifestation of SIOD could be due to the changes in gene expression profiles of key transcription factors which are directly or indirectly regulated by SMARCAL1 with the negative regulation of c-*myc* presented in this paper being one such example.

Results

Downregulation of *SMARCAL1* leads to altered gene expression pattern.

Baradaran-Heravi *et al.* have hypothesized that SMARCAL1 can possibly regulate genes like c-*kit* and c-*myc* by altering the promoter structure[4]. c-*kit* expression is regulated by G-quadruplex formation, a feature that is shared by another transcription factor, c-*myc*, which regulates 10–15% of genes in mammalian cells. To explore whether SMARCAL1 can regulate gene expression of c-*myc*, we downregulated *SMARCAL1* in HeLa cells using shRNA and obtained three monoclonals- Sh1, Sh2, and Sh3 as well as one polyclonal cell line (Sh). We confirmed that SMARCAL1 was indeed downregulated in all these cell lines using quantitative real-time RT-PCR (Supplementary Fig. S1). Since BRG1 is also known to regulate the transcription of c-*myc* by binding to the FUSE region[29], and SMARCAL1 regulates *brg1* expression (Haokip *et al.* companion paper) we analyzed the expression of c-*myc* and *brg1*. We found that both *brg1* (Supplementary Fig. S1) as well as c-*myc* were downregulated (Supplementary Fig. S1) in *SMARCAL1* downregulated cells.

We will focus on c-*myc* transcription in this paper and explain how SMARCAL1 possibly regulates BRG1 in the companion paper.

BRG1 and SMARCAL1 are present on the c-*myc* promoter.

The above result indicated that either BRG1 or SMARCAL1 or both were possibly regulating c-*myc* transcription. Therefore, the occupancy of BRG1, SMARCAL1, and RNAPII on the c-*myc* promoter was probed using 5 pairs of overlapping primers (25-30 bp overlaps) designed with respect to the c-*myc* P2 promoter spanning the region from -810 bp to $+39$ bp each giving ~200 bp amplicon (Fig. 1A–E). We found that all three proteins were localized on the promoter at primer B and C region though the occupancy of SMARCAL1 appeared to be greater around the primer B region than the primer C region (Fig. 1C,D). The occupancy of BRG1 and RNAPII appeared to be similar around primer B and C regions (Fig. 1C,D).

Henceforth, we will refer to primer B and C regions as Myc_B_{159} and Myc_C_{201} respectively.
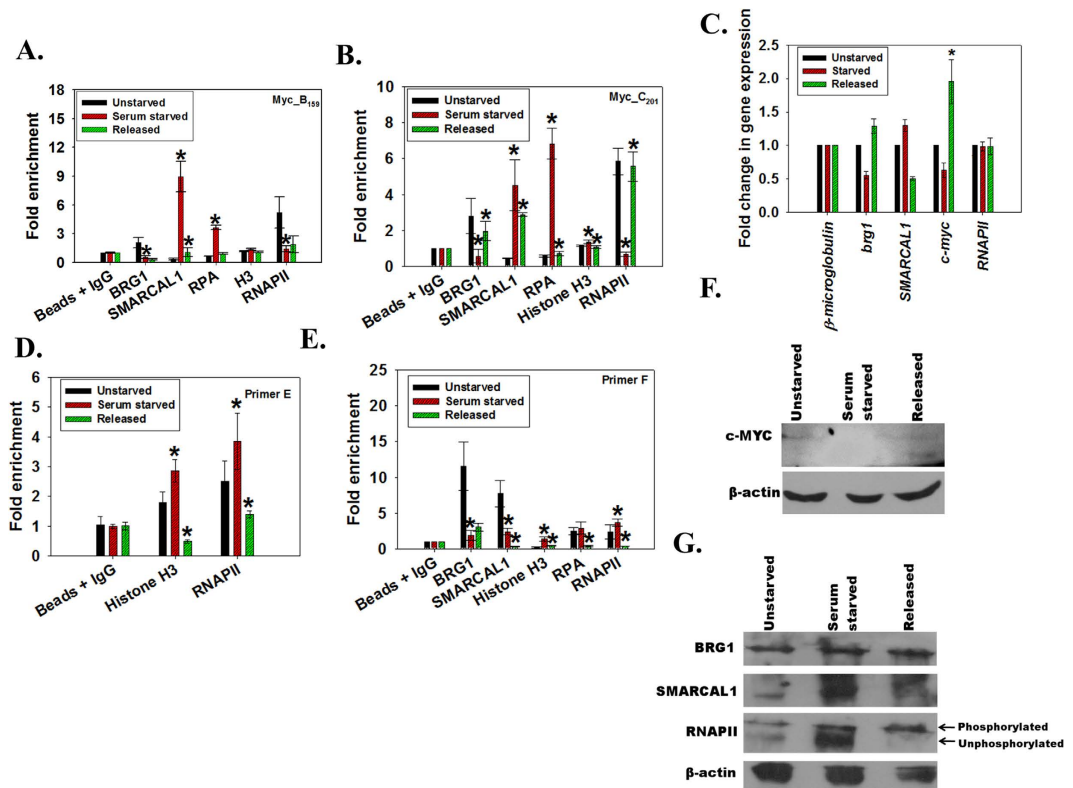
Figure 2. Analysis of occupancy of BRG1, SMARCAL1 and RNAPII on c-*myc* promoter during and after serum starvation. (**A**). Occupancy of BRG1, SMARCAL1, RPA, histone H3, and RNAPII was analyzed on Myc_B_{159} region using quantitative real-time RT-PCR following ChIP using respective antibodies. The star indicates statistical significance at $p < 0.05$. (**B**). Occupancy of BRG1, SMARCAL1, RPA, histone H3, and RNAPII on Myc_C_{201} region of c-*myc* promoter was analyzed using quantitative real-time RT-PCR following ChIP using respective antibodies. The star indicates statistical significance at $p < 0.05$. (**C**). Expression of c-*myc*, *SMARCAL1, brg1,* and *RNAPII* was monitored using quantitative real-time RT-PCR. *β- microglobulin* was used as control in this experiment. The star indicates statistical significance at $p < 0.05$. (**D**). Occupancy of histone H3 and RPA on Primer E region was analyzed using quantitative real-time RT-PCR following ChIP with respective antibodies. The star indicates statistical significance at $p < 0.05$. (**E**). Occupancy of BRG1, SMARCAL1, histone H3, RPA and RNAPII on FUSE region was analyzed using quantitative real-time RT-PCR following ChIP with respective antibodies. The star indicates statistical significance at $p < 0.05$. (**F**). c-Myc expression in unstarved, serum starved, and cells released from the block was analyzed using western blot. (**G**). SMARCAL1, BRG1, and RNAPII expression in unstarved, serum starved, and cells released from serum starvation were analyzed using western blots. In this experiment, β-actin was used as the control. Quantitation of pixel values is provided in Supplementary Fig. S2. Uncropped western blots are provided in Supplementary Fig. S11.

Increased SMARCAL1 occupancy on c-*myc* promoter upon serum starvation. As both BRG1 and SMARCAL1 were present on the c-*myc* promoter, we wanted to know whether both the proteins were required simultaneously for c-*myc* transcription. It has been previously shown that c-*myc* transcription is reduced when cells are serum starved as they enter into G0 phase[29]. Upon release from serum starvation, c-*myc* transcription restarts as the protein is required for cells to enter into the cell cycle[29]. This experimental model of change in c-*myc* transcription upon serum starvation for 48 hours and subsequent release was, therefore, used to analyze the occupancy of SMARCAL1, BRG1, RNAPII, histone H3 and RPA on the c-*myc* promoter using chromatin immunoprecipitation. The analysis showed that occupancy of SMARCAL1 and RPA increased on the c-*myc* promoter increased both at Myc_B_{159} and Myc_C_{201} regions (Fig. 2A,B). Further, histone H3 occupancy increased 2-fold upon serum starvation around Myc_C_{201} as compared to the unstarved cells (Fig. 2B). Concomitantly, BRG1 and RNAPII occupancy on Myc_B_{159} and Myc_C_{201} regions decreased coinciding with decreased c-*myc* expression upon serum starvation (Fig. 2A–C). As RNAPII is known to stall at the proximal promoter region, we also examined the occupancy of RNAPII and H3 on the promoter P2 start site (Primer E in Fig.1). As expected, under serum starvation condition, RNAPII and H3 occupancy around promoter start site increased indicating that the polymerase was paused (Fig. 2D). Finally, we examined the occupancy of these proteins at the FUSE region as BRG1 is known to mediate its effect through this DNA segment. We found that occupancy of BRG1 and SMARCAL1 decreased on the FUSE region on serum starvation as compared to the unstarved cells (Fig. 2E). Further, at this region, the occupancy of histone H3 and RNAPII increased while that of RPA remained unchanged (Fig. 2E).

On release from serum starvation, SMARCAL1 occupancy decreased substantially on Myc_B_{159} as compared to Myc_C_{201} while the BRG1 occupancy started increasing on the Myc_C_{201} as compared to Myc_B_{159} (Fig. 2A,B). Further, RPA occupancy decreased on both Myc_B_{159} and Myc_C_{201} regions while H3 occupancy decreased on Myc_C_{201} region (Fig. 2A,B). Simultaneously, RNAPII occupancy on Myc_C_{201} increased 5-fold as compared to the serum starved condition, correlating with increased c-*myc* transcription (Fig. 2B,C). In addition, at the promoter start site, RNAPII and H3 occupancy decreased, indicating that the polymerase was no longer paused (Fig. 2D). The occupancy of SMARCAL1, RPA, H3, and RNAPII also decreased at the FUSE region while that of BRG1 increased slightly (Fig. 2E).

The transcript analysis showed that upon serum starvation the expression of *SMARCAL1* increased while that of c-*myc* decreased (Fig. 2C). We corroborated the transcript analysis with western blots. The expression of c-Myc decreased on serum starvation as compared to unstarved cells (Fig. 2F and Supplementary Fig. S2). However, the expression was not restored to unstarved cells on release from the block even though the transcript levels were upregulated (Fig. 2F and Supplementary Fig. S2). The expression of SMARCAL1 increased on serum starvation and decreased on release from the block while the expression of BRG1 was unchanged during and after serum starvation (Fig. 2G and Supplementary Fig. S2). The antibody for RNAPII recognizes both unphosphorylated and phosphorylated forms of the protein and we observed two bands in unstarved condition (Fig. 2G and Supplementary Fig. S2). The intensity of the lower band (unphosphorylated form) increased on serum starvation and decreased on release from serum starvation (Fig. 2G and Supplementary Fig. S2).

From this experiment we concluded that BRG1 was a positive regulator and SMARCAL1 was a negative regulator of c-*myc* transcription.

BRG1 and SMARCAL1 direct their effects through c-*myc* promoter.

To understand how BRG1 and SMARCAL1 were mediating their transcriptional regulation, the c-*myc* promoter (P2 TSS to -765 bp upstream) was cloned into pGL3 basic reporter plasmid. The construct was transfected into HeLa cells and the serum starvation experiment was performed. We found that on serum starvation the luciferase activity was downregulated by 15% and on release from serum starvation the luciferase activity was upregulated (Supplementary Fig. S2 and S3). To delineate the roles of Myc_B_{159} and Myc_C_{201} in transcription regulation, we cloned these DNA regions separately into pGL3 promoter plasmid and performed the same experiment (Supplementary Fig. S2). We found that the luciferase activity was downregulated when Myc_B_{159}-promoter construct transfected HeLa cells were serum starved (Supplementary Fig. S2and S3). This effect was not observed when Myc_C_{201} promoter was transfected into HeLa cells and serum starved, suggesting that SMARCAL1 primarily mediates its effect through Myc_B_{159} (Supplementary Fig. S2 and S3). Upon release from serum starvation, the luciferase activity remained downregulated in case of Myc_B_{159}-promoter construct but was upregulated in case of Myc_C_{201}-promoter construct, suggesting that BRG1 mediates its effect primarily through Myc_C_{201} region (Supplementary Fig. S2 and S3).

The c-*myc* promoter accesibility is altered upon serum starvation.

To understand the changes in the chromatin architecture of the c-*myc* promoter, the promoter region from -653 to -299 bp, with respect to P2 transcription start site, was scanned for protein occupancy in unstarved, serum starved, and released cells. We hypothesized that the DNA bound by protein would be inaccessible to micrococcal nuclease (MNase) in a manner similar to that of nucleosome-bound DNA whereas the protein-free region would be accessible for digestion by MNase. The protein-bound DNA after MNase digestion could, therefore, be purified and amplified using primers specific to the c-*myc* promoter (Supplementary Fig. S4). Comparing this result with the ChIP data would enable us to identify regions bound by BRG1, RNAPII, SMARCAL1, RPA and histone H3 in unstarved, serum starved, and released from starvation conditions on the c-*myc* promoter. In these experiments DNA of approximately 150 bp was purified, and therefore, proteins binding to a region larger or smaller than 150 bp would not be detected in this experiment (Supplementary Fig. S4). Finally, the protein occupancy was compared with respect to genomic DNA assuming that the genomic DNA was completely free of proteins as explained in methods.

We found that the protein occupancy on the c-*myc* promoter changed from unstarved to serum starved conditions and again in released condition. In unstarved condition, the regions mapped by primers I and II contained proteins while those mapped by primers III, IV, V, and VI were free of proteins (Fig. 3A). When cells were serum starved, the protein occupancy in the regions mapped by primers I and II decreased while it increased in regions scanned by primers IV and V (Fig. 3A) . When cells were released from serum starvation, protein occupancy on primer I region was unchanged with respect to starved conditions but increased in the region amplified by primer II, IV, V, and VI (Fig. 3A). Combining this data with the ChIP data enabled us to model BRG1, RNAPII, and SMARCAL1 on the promoter. In normal, unstarved conditions, BRG1 and RNAPII are present spanning -653 to -506 bp upstream of P2 (Fig. 3B). In addition, RNAPII and H3 are present at the transcription start site. Upon serum starvation, the protein occupancy changes and SMARCAL1, RPA and H3 are found present between -653 to -299 bp (Fig. 3C). Further, RNAPII and H3 occupancy around the transcription start site increases. When cells are released from serum starvation, BRG1 and RNAPII occupies -653 to -299 while SMARCAL1 is present from -523 to -299 bp (Fig. 3D). RPA and H3 are removed from Myc_B_{159} and Myc_C_{201} regions while RNAPII is no longer paused around the promoter proximal site (Fig. 3D).

The c-*myc* promoter occupancy is altered in *SMARCAL1* downregulated cells.

As stated earlier, c-*myc* expression was repressed in *SMARCAL1* downregulated cells. The same result was obtained when we transfected c-*myc* promoter cloned into pGL3 basic vector and measured the luciferase activity (Supplementary Fig. S5). Further, we measured the occupancy of RNAPII and H3K9Ac, a histone modification associated with transcription activation[30], by ChIP and found that RNAPII occupancy on Myc_B_{159} and Myc_C_{201} was reduced in *SMARCAL1* downregulated cells (Supplementary Fig. S5). We also found that H3K9Ac levels were almost negligible in this region and there was no appreciable difference between the control and *SMARCAL1* downregulated cells
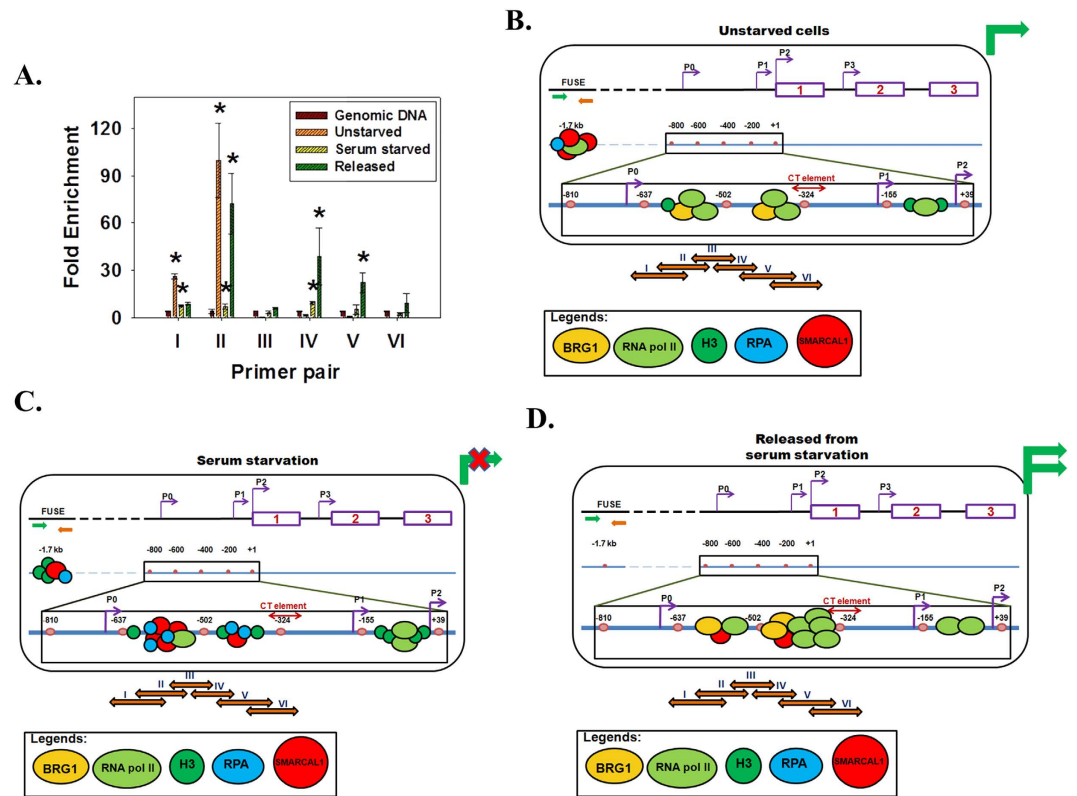
Figure 3. Protein occupancy on c-*myc* promoter is altered upon serum starvation. (**A**). Quantitative real-time RT-PCR was used to measure the fold enrichment of proteins on the c-*myc* promoter in unstarved HeLa cells, serum starved HeLa cells, and HeLa cells released from the serum starvation block. (**B**). Model explaining the occupancy of BRG1, SMARCAL1, RPA, H3, and RNAPII in unstarved HeLa cells; (**C**). in starved HeLa cells; (**D**). in cells released for 2 hours from 48 hours of serum starvation. The star indicates significant difference at $p < 0.05$.

(Supplementary Fig. S5). In addition, we hypothesized that downregulation of *brg1* should also result in reduced c-*myc* expression. Therefore, we also measured the c-*myc* transcript levels in Sh*brg1* cells and found them down-regulated as expected (Supplementary Fig. S5).

Finally, we measured the promoter accessibility in control and *SMARCAL1* downregulated cells. As in the case of the unstarved cells, in control cells also the protein occupancy was found only in the regions amplified by primer pairs I and II (compare Fig. 3A and Supplementary Fig. S6). However, in case of *SMARCAL1* downregulated cells, the protein occupancy decreased in the regions amplified by these primer pairs (Supplementary Fig. S6). Further, there was no alteration in protein occupancy in the regions amplified by primer pairs III, IV, V, and VI, indicating that in the downregulated cells, the entire region spanning −653 to −299 bp was relatively free of protein occupancy (Supplementary Fig. S6).

Comparing this data with the ChIP enabled us to propose that RNAPII is present on the c-*myc* promoter in control cells while in *SMARCAL1* downregulated cells, this region is unprotected possibly because activating proteins are not present on the promoter resulting in decreased transcription (Supplementary Fig. S6).

c-*myc* expression is downregulated both in *SMARCAL1* downregulated cells as well as in serum starved cells. However, our model proposes that the mode of downregulation is different in these two cases. The expression of c-*myc* gene is downregulated in *SMARCAL1* downregulated cells because RNAPII is not present on the pro-moter. Under serum starvation condition, the c-*myc* expression is repressed because SMARCAL1 is present on the promoter region. Thus, these data led us to conclude that BRG1 and SMARCAL1 regulate c-*myc* expression in antagonistic manner.

Myc_B$_{159}$ DNA acts as an effector of ADAAD, the bovine homolog of SMARCAL1.   As stated earlier, Baradaran-Heravi *et al.* have postulated that SMARCAL1 could possibly induce conformational changes in promoter regions[4].

The QGRS[31] software (http://bioinformatics.ramapo.edu/QGRS/analyze.php) predicted a G-quadruplex in Myc_B$_{159}$ region G2-L3-G2-L1-G2-L2-G2 where G is guanine and L is a loop comprised of any nucleotide (Supplementary Table S1). Human SMARCAL1 is difficult to overexpress and purify in sufficient amount for biophysical studies. Therefore, to understand whether SMARCAL1 can induce the formation of G-quadruplex in the Myc_B$_{159}$ DNA, we used Active DNA-dependent ATPase A Domain (ADAAD) which is the bovine homolog of human SMARCAL1 (79% identity) and has been well-characterized in our laboratory[8,32].
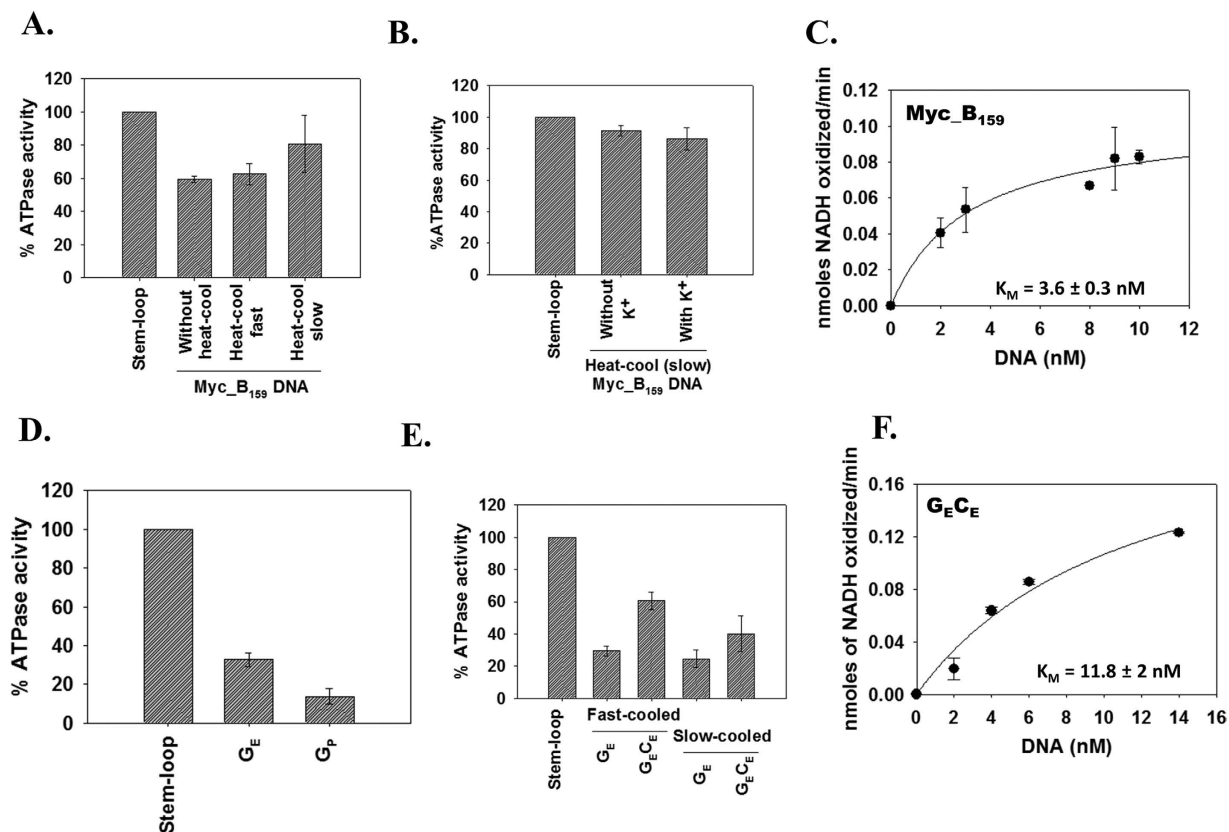
Figure 4. Myc_B$_{159}$ as well as G$_E$ are effectors for ADAAD. (**A**). ADAAD can hydrolyze ATP in the presence of Myc_B$_{159}$. (**B**). Heat-cooled Myc_B$_{159}$ DNA in the absence and presence of 100 mM K$^+$ showed similar ATPase activity. In these experiments, 0.1 μM ADAAD and 20 nM DNA was used and the reaction was incubated at 37 °C for 90 minutes. The ATPase activity due to Myc_B$_{159}$ DNA was compared with that due to stem-loop DNA. (**C**). The apparent K$_M$ for ADAAD-Myc_B$_{159}$ DNA was calculated using ATPase activity in presence of increasing concentration of DNA. The reaction was incubated at 37 °C for 45 minutes. (**D**). ADAAD prefers G$_E$ as an effector compared to G$_P$. In this experiment 2.4 μM ADAAD was incubated with 10 nM DNA for 45 minutes at 37 °C and the amount of ATP hydrolyzed was compared to that due to stem-loop DNA. (**E**). Comparison of effector properties of G$_E$ and G$_E$C$_E$ with respect to stem-loop DNA. G$_E$ and G$_E$C$_E$ was heat denatured and either rapidly cooled or slow-cooled before incubating with 0.24 μM ADAAD for 45 minutes at 37 °C and the amount of ATP hydrolyzed was compared to that due to stem-loop DNA. 10 nM DNA was used in the experiment. (**F**). The apparent K$_M$ for ADAAD-G$_E$C$_E$ DNA was calculated using ATPase activity in presence of increasing concentration of DNA. The reaction was incubated at 37 °C for 45 minutes.

ADAAD was able to hydrolyze ATP in the presence of Myc_B$_{159}$ (Fig. 4A). There was no significant difference in the amount of ATP hydrolyzed when Myc_B$_{159}$ was heat-cooled either in the absence or presence of 100 mM K$^+$ (Fig. 4A,B). The apparent K$_M$ for ADAAD- Myc_B$_{159}$ interaction was calculated to be 3.6 ± 0.3 nM (Fig. 4C).

The interaction with the predicted G-quadruplex region is weaker than with Myc_B$_{159}$. Next we sought to determine whether ADAAD specifically binds to the putative G-quadruplex forming region present within Myc_B$_{159}$ DNA. For these studies, we synthesized G$_E$, a 34 nt single-stranded DNA encompassing the predicted G-quadruplex sequence (Supplementary Table S2). G$_E$ can form double-strand in presence of its complementary sequenc C$_E$ (Supplementary Table S2). In addition, we also synthesized a single-stranded olignucleotide, G$_P$, corresponding to the CT element, as a positive control since QGRS predicted that this region can also form a G-quadruplex. ATPase assays showed that G$_E$ is a better effector than G$_P$ (Fig. 4D).

Previously, we have shown that ADAAD specifically recognizes double-stranded to single-stranded transition regions in DNA molecule[8]. As both G$_E$ and G$_P$ have been predicted to form G-quadruplex, we used Mfold[33] program to understand the differences in the secondary structure of G$_E$ and G$_P$. The Mfold program predicted that G$_E$ can form a stem-loop structure with a 11 bp stem and 3 base loop (ΔG = −8.64 kcal/mol) while G$_P$, the single-stranded oligonucleotide corresponding to the CT element, forms an unstable hairpin loop structure with the G$_8$-T$_{18}$ closing the loop (ΔG = 3.66 kcal/mol), thus providing a structural basis for the effector preference (Supplementary Fig. S7).

As under *in vivo* condition, DNA is present as a double-stranded molecule, we next compared the effector properties of single-stranded G$_E$ molecule with that of the double-stranded G$_E$C$_E$ molecule. Further, the formation of stem-loop structure is an intramolecular event that can be brought about by denaturing the DNA and rapidly cooling it. In contrast, intermolecular events are favoured by cooling the DNA molecules slowly after denaturing. To

DNA	V_{max} (nmoles of NADH oxidized/min)	K_M (nM)	k_{cat} (s^{-1})	k_{cat}/K_M (X 10^7 M^{-1} s^{-1})
Myc_B$_{159}$	0.11 ± 0.02	3.6 ± 0.3	0.075 ± 0.02	2.12 ± 0.6
G$_E$C$_E$	10.45 ± 0.9	11.75 ± 1.9	2.89 ± 0.24	24.7 ± 2.0

Table 1. Comparison of kinetic parameters for ADAAD interaction with Myc_B$_{159}$ and G$_E$C$_E$ DNA.

understand whether the stem-loop structure is really critical for the interaction, we heated the DNA molecules and then either rapidly cooled it or slowly cooled it (Supplementary Fig. S8). We found that G$_E$C$_E$ was a better effector than G$_E$ both when it was rapidly cooled and when it was slow cooled after heat denatuaration (Fig. 4E). Further, G$_E$C$_E$ was a better effector when it was fast cooled suggesting that the secondary structure of the oligonucleotide was essential for the interaction.

To understand whether ADAAD specifically recognizes G$_E$C$_E$ region within Myc_B$_{159}$ DNA, we calculated the K_M for ADAAD-G$_E$C$_E$ interaction. As shown in Fig. 4F, the K_M was found to be 11.8 ± 2 nM, about 3-fold less than that for ADAAD-Myc_B$_{159}$. However, the catalytic efficiency was higher for G$_E$C$_E$ as compared to Myc_B$_{159}$ (Table 1). From this we concluded that though G$_E$C$_E$ can act as an effector of ADAAD, this is not sufficient for the interaction within the context of Myc_B$_{159}$.

Myc_B$_{159}$ and G$_E$C$_E$ do not possess G-quadruplex structure. Circular dichroism has been used to monitor the conformational changes in DNA[34]. It has been reported that the CD spectra of anti-parallel G-quadruplex shows a negative peak at 260 nm and a positive peak at 290 nm while parallel G-quadruplex shows a positive peak at 260 nm and a negative peak at 240 nm. To understand whether double-stranded Myc_B$_{159}$ and G$_E$C$_E$ can form G-quadruplex type structure in the absence of DNA and ATP, we used CD spectroscopy to analyze their structure. G$_E$C$_E$ after heat-cooling showed a negative peak at 230 nm while Myc_B$_{159}$ showed a negative peak at 210 nm and a small rise at 275 nm, which are characteristic of double-stranded B-DNA (Supplementary Fig. S9).

As the G-quadruplex formation is an intramolecular event, we also analyzed the structure of single-stranded G$_E$ DNA after heat denaturation and rapid cooling. The CD spectra showed a negative peak at 220 nm (Supplementary Fig. S9). This was similar to the peak obtained in case of stem-loop DNA (Supplementary Fig. S9).

The secondary structure, especially G-quadruplex formation, is dependent on monovalent cations, especially K$^+$. Therefore, we analyzed the structures of G$_E$C$_E$, Myc_B$_{159}$, G$_E$, and stem-loop DNA in the presence of K$^+$ and found the structure in the absence and presence of K$^+$ were similar (Supplementary Fig. S9). Further, the structure of G$_E$C$_E$, G$_E$ and stem-loop DNA in the presence of K$^+$ was similar (Supplementary Fig. S9).

Finally, we analyzed the structure of G$_P$ and found that it can indeed form the characteristic peaks of G-quadruplex (Supplementary Fig. S9).

From the CD analysis, we concluded that none of the DNA molecules at the concentration used in the experiment show the characteristic peaks of G-quadruplex either in the absence or presence of K^{+34}.

ADAAD induces a conformational changes on binding to the c-*myc* promoter. Next, we analyzed the conformational changes induced in the double-stranded G$_E$C$_E$ and Myc_B$_{159}$ DNA upon addition of ADAAD and ATP both in the absence and presence of K$^+$ ions. When G$_E$C$_E$ (in the absence of K$^+$) was incubated with both ADAAD and ATP, the DNA showed two positive peaks- one at 258 nm with a shoulder at 269 nm and a larger peak at 210 nm (Fig. 5A). Addition of EDTA to the reaction mix resulted in disruption of the conformational change, suggesting that ADAAD mediated ATP hydrolysis is necessary for the change in conformation of G$_E$C$_E$ (Fig. 5B). To understand whether the conformational was similar to the one induced in stem-loop DNA by ADAAD, we also analyzed the structure of stem-loop DNA in the absence of K$^+$. As shown in Fig. 5C, the stem-loop DNA also showed two positive peaks–one at 250 nm and a shoulder at 272 nm and a larger peak at 215 nm (Fig. 5C,D).

Subsequently, we studied the effect of K$^+$ ions on the conformational change induced in G$_E$C$_E$ DNA by the addition of ADAAD and ATP. In the presence of K$^+$, we found that the G$_E$C$_E$ DNA showed a negative peak at 257 nm (Fig. 5E). The formation of this peak was also dependent on ATP hydrolysis as it was disrupted upon addtion of EDTA (Fig. 5F). Further, the stem-loop DNA also showed a negative peak at 257 nm in the when heat-cooled in the presence of K$^+$ (Fig. 5G,H).

Finally, we studied the conformational change induced in Myc_B$_{159}$ when incubated with ADAAD and ATP. This DNA, however, in the absence of K$^+$ showed a negative peak at 262 nm, which did not change with incubation at 37 °C (Fig. 6A,B). Formation of this peak was dependent on the presence of both ADAAD and ATP (Fig. 6A). The conformation was disrupted when ADAADiN, the known inhibitor of ADAAD[35,36] was added to the reaction, indicating continued hydrolysis is needed for maintaining the conformation (Fig. 6C). In contrast, when Myc_B$_{159}$ heat-cooled in the presence of K$^+$ was incubated with ADAAD and ATP, a positive peak at 262 nm was obtained which did not change with incubation at 37 °C and was dependent on the presence of both ADAAD and ATP (Fig. 6D,E) . Addition of ADAADiN, however, resulted in stabilization of the positive 262 nm peak (Fig. 6F).

From these experiments we conclude that ADAAD, the bovine homolog of SMARCAL1, can induce conformational change in the c-*myc* promoter DNA in an ATP-dependent manner such that structure so formed acts as an impediment to RNAPII binding, thus leading to transcription repression.

Inverse correlation between expression of SMARCAL1 and c-MYC. What is the physiological relevance of SMARCAL1 regulating c-*myc* expression? An analysis of c-*myc*[37] and *SMARCAL1*[5] levels in adult mouse tissues showed that the expression of *SMARCAL1* was high in kidney, brain, and liver while c-*myc* was not expressed in these tissues though it was expressed in the newborn mice. As symptoms of SIOD include renal dysfunction and
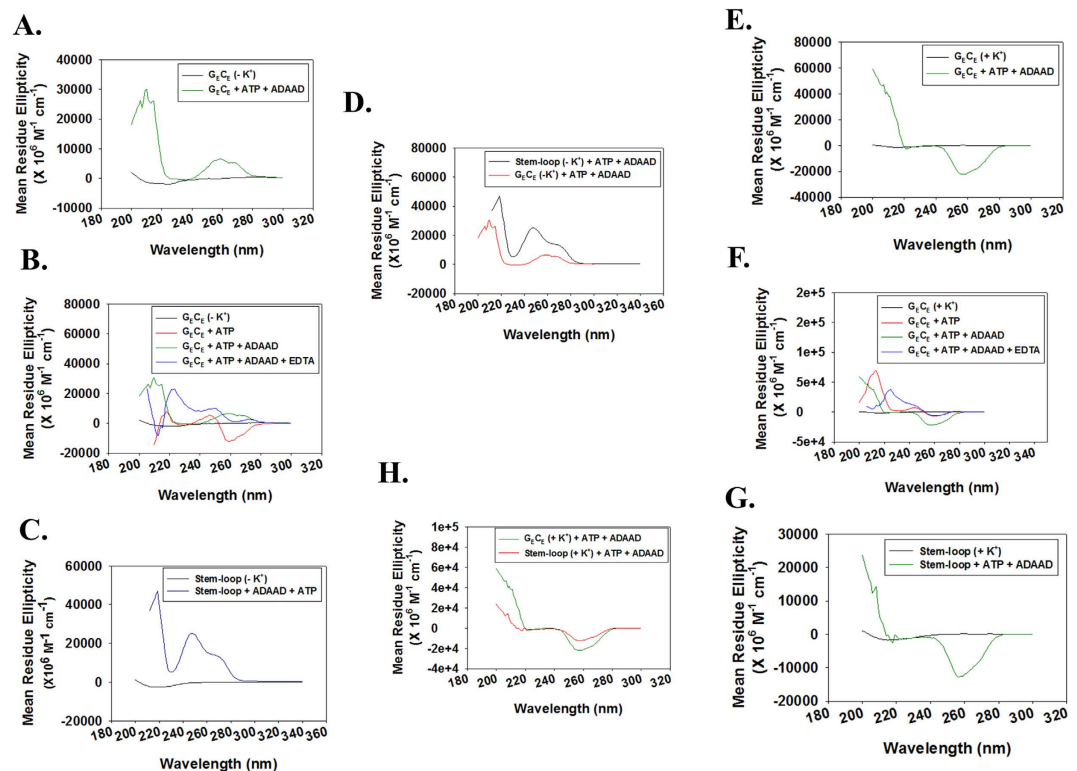
Figure 5. The conformation of G_EC_E is altered in an ATP-dependent manner in the presence of ADAAD.
(**A**). CD spectra of G_EC_E (heat-cooled in the absence of K^+) in the presence of ATP and ADAAD. (**B**). Comparison of CD spectra of G_EC_E (heat-cooled in the absence of K^+) in the presence of ATP alone, in the presence of both ATP and ADAAD, and after addition of EDTA. (**C**). CD spectra of stem-loop DNA (heat-cooled in the absence of K^+) in the presence of both ATP and ADAAD. (**D**). Comparison of CD spectra of G_EC_E and stem-loop DNA in the presence of both ATP and ADAAD. Both the DNA molecules were heat-cooled in the absence of K^+. (**E**). CD spectra of G_EC_E (heat-cooled in the presence of 100 mM K^+) in the presence of ATP and ADAAD. (**F**). Comparison of CD spectra of G_EC_E (heat-cooled in the presence of 100 mM K^+) in the presence of ATP alone, in the presence of both ATP and ADAAD, and after addition of EDTA. (**G**). CD spectra of stem-loop DNA (heat-cooled in the presence of 100 mM K^+) in the presence of both ATP and ADAAD. (**H**). Comparison of CD spectra of G_EC_E and stem-loop DNA in the presence of both ATP and ADAAD. Both DNA molecules were heat-cooled in the presence of 100 mM K^+. In these experiments, 0.5 μM DNA, 1 μM ADAAD, 2 mM ATP, 10 mM Mg^{+2}, and 50 mM EDTA was used. The DNA molecules were heat denatured at 94 °C for 3 minutes and rapidly cooled to 4 °C either in the absence or presence of 100 mM K^+.

cereberal ischemia, we hypothesized that SMARCAL1 expression during differentiation is critical for repressing c-*myc* expression.

Differentiation of K562 into myeloid cell lineage in response to PMA treatment has been well-documented and correlates with transcription repression of c-*myc*[38–40]. We used this system to understand whether downregulation of c-*myc* expression during differentiation was dependent on SMARCAL1 expression.

K562 cells were treated with PMA for upto 5 days and the expression of *brg1*, *SMARCAL1*, and *c-myc* were estimated using quantitative real-time RT-PCR. As shown in Supplementary Fig. S10, c-*myc* expression was down-regulated and *SMARCAL1* expression was upregulated after treatment with PMA validating the inverse correlation we observed between *SMARCAL1* and c-*myc* expression when cells were serum starved and subsequently, released from the block (Fig. 2C). Interestingly, *brg1* expression was unaltered when K562 cells differentiated.

Discussion

The transcription factor, c-*myc*, regulates 10–15% genes in the mammalian cells[18]. The expression of c-*myc* itself is regulated by exquisite machinery consisting of DNA secondary structures, transcription factors, ATP-dependent chromatin remodeling factors, and multiple promoters[23].

The CT element present 150 bp upstream of promoter P1 has been shown to form G-quadruplex and i-motif structures[24–26,41]. Nucleolin has been shown to bind to this element both *in vitro* and *in vivo*[42]. Further, the protein has been shown to induce G-quadruplex structure in this element *in vitro* suggesting that it might be doing the same *in vivo*[42]. In addition to the CT element, the promoter also contains a FUSE region about 1.7 kb upstream of the promoter P2[23]. This region possesses nucleosomes and BRG1 has been shown to reposition the nucleosomes during activation of c-*myc* expression[29].
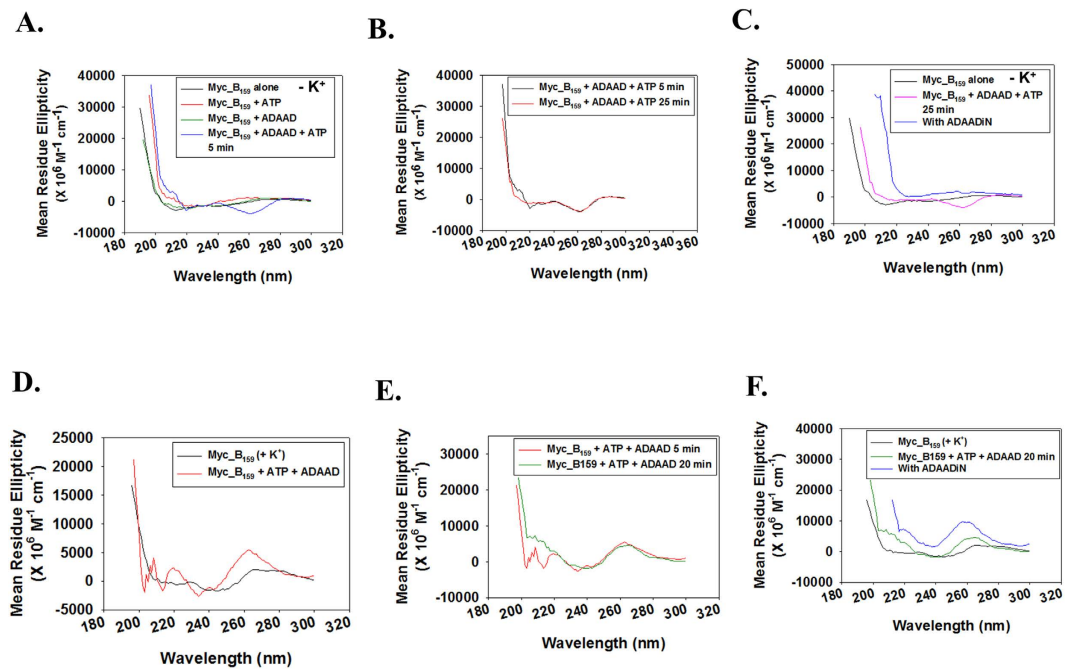
Figure 6. Conformational changes induced in Myc_B$_{159}$ DNA. (A). Conformation of Myc_B$_{159}$ DNA (heat-cooled in the absence of K$^+$) in the presence of ATP and ADAAD. (**B**). Comparison of CD spectra of Myc_B$_{159}$ DNA (heat-cooled in the absence of K$^+$) in the presence of ATP and ADAAD at 5 minutes and 25 minutes. (**C**). CD spectra of Myc_B$_{159}$ DNA (heat-cooled in the absence of K$^+$) in the presence of ATP, ADAAD, and ADAADiN. (**D**). Conformation of Myc_B$_{159}$ DNA (heat-cooled in the presence of 100 mM K$^+$) in the presence of ATP and ADAAD. (**E**). Comparison of CD spectra of Myc_B$_{159}$ DNA (heat-cooled in the presence of 100 mM K$^+$) in the presence of ATP and ADAAD at 5 minutes and 20 minutes. (**E**). CD spectra of Myc_B$_{159}$ DNA (heat-cooled in the presence of 100 mM K$^+$) in the presence of ATP, ADAAD, and ADAADiN. In these experiments 0.15 μM Myc_B$_{159}$ DNA, 0.1 μM ADAAD, 0.5 mM ATP, 10 mM Mg^{+2} and 5 μM ADAADiN were used.

In this paper, we have identified yet another element regulating the expression of *c-myc*. SMARCAL1, a distant member of the ATP-dependent chromatin remodeling protein family, has been shown to function as an annealing helicase required for stabilizing replication forks when DNA is damaged in the S phase[9–14,43]. Mutations in SMARCAL1 have been correlated with SIOD, a pleiotropic disorder characterized by spondyloepiphyseal dysplasia, renal dysfunction and T-cell immunodeficiency[15,44]. As SIOD encompasses many different organ systems, it has been hypothesized that SMARCAL1 might function as a transcriptional regulator of a particular subset of genes, possibly through its chromatin remodeling activity[15]. SIOD patients have been found to show disturbed gene expression profiles necessary for skeletal development, renal tissue maintenance and T-cell development[16]. However, till now SMARCAL1 has not been shown to localize at promoters and there exists only one report that documents the interaction of SMARCAL1 with histones[45].

Using a combination of *in vivo* and *in vitro* experiments we have shown that SMARCAL1 regulates *c-myc* transcription. The protein binds to a region termed as Myc_B$_{159}$ present upstream of the CT element. This region has not previously been shown to form either G-quadruplex or to regulate the transcription activity. Our studies show that under normal conditions BRG1 and RNAPII bind to this region allowing for transcription to occur. During serum starvation, when cells enter into the G0 phase and *c-myc* transcription is shut off, SMARCAL1 occupancy on Myc_B$_{159}$ increases while that of BRG1 and RNAPII decreases correlating with transcription repression. When cells are released from the block, the occupancy of SMARCAL1 decreases while that of BRG1 and RNAPII increases resulting in transcription activation. Thus, BRG1 functions as a positive regulator while SMARCAL1 acts as a negative regulator of *c-myc* transcription.

SMARCAL1 has been shown to interact with RPA and we find that these two proteins are present on the *c-myc* promoter on serum starvation. However, from these experiments it is not clear whether RPA recruits SMARCAL1 to the promoter or not. Further, our experiments also show that the Myc_B$_{159}$ and Myc_C$_{201}$ regions are free from nucleosomes under normal conditions. When cells were serum starved, H3 occupancy increased on Myc_C$_{201}$ as well as at the transcription start site (primer E region) indicating SMARCAL1 might induce formation of an inaccessible chromatin structure leading to transcription repression.

Using the bovine homolog of SMARCAL1 and *in vitro* assays we have found that Myc_B$_{159}$ is indeed an effector of ATPase activity and the protein binds to this region with an apparent K$_M$ of 3.6 ± 0.3 nM, which was similar to the stem-loop DNA that we reported earlier as the best effector for this protein[8]. Bioinformatic[31] studies predicted a potential G-quadruplex forming region within Myc_B$_{159}$. A 34-nucleotide double-stranded DNA, called G$_E$C$_E$, spanning the potential G-quadruplex was also found to be an effector of ADAAD, though the K$_M$ values indicated that the interaction is 3-fold weaker, suggesting the protein recognizes additional regions within Myc_B$_{159}$ DNA.

CD spectroscopy showed that interaction of ADAAD with G_EC_E in the presence of ATP leads to a conformational change in the DNA, which was similar to the one induced in stem-loop DNA leading us to hypothesize SMARCAL1 might be recognizing c-*myc* promoter region due to its structural features viz. ability to form stem-loop structure. The conformation of Myc_B_{159} adopted in the presence of ADAAD and ATP is different from that of G_EC_E. This difference between Myc_B_{159} and G_EC_E could be due to the length of the molecule. Despite the difference in the spectra we can conclude that binding of ADAAD results in a conformation change in DNA which possibly impedes the binding of RNAPII and BRG1 to the promoter resulting in transcription repression.

Using the K562 cell lines as a model system, we have shown that during differentiation the expression of *SMARCAL1 i*s upregulated while that of c-*myc* is downregulated, suggesting a negative correlation between the expression of these two proteins.

However, we have also shown in the paper that downregulation of *SMARCAL1* leads to downregulation of c-*myc*. This is due to the fact that downregulation of *SMARCAL1* leads to downregulation of *brg1* also as a positive feedback loop governs the levels of these two proteins (Haokip *et al.* Transcriptional regulation of ATP-dependent chromatin remodeling factors: positive feedback loop regulates the levels of SMARCAL1 and BRG1). As BRG1 is a positive regulator of c-*myc* expression, we observe a downregulation in the levels of c-*myc* in *SMARCAL1* downregulated cells. The positive feedback regulation between SMARCAL1 and BRG1, however, seems to be operative only under specific conditions. For example, when K562 cells undergo differentiation, the expression of BRG1 does not change while SMARCAL1 is significantly upregulated. Similarly, during serum starvation, SMARCAL1 expression is upregulated but not that of BRG1. The upregulation of SMARCAL1 expression correlates with downregulation of c-*myc* expression during both K562 differentiation and serum starvation of HeLa cells.

The relevance of regulation of c-*myc* expression by SMARCAL1 became clear when we compared northern blot analysis of these c-*myc* and *SMARCAL1* in mouse adult tissues reported by Zimmerman *et al.*, and Coleman *et al.*, respectively[5,37]. c-*myc* expression was absent in tissues where *SMARCAL1* was expressed at high level. Interestingly, in kidney and brain tissues, c-*myc* was not expressed in adult mice[37], while *SMARCAL1* was expressed[5]. Further, lung and spleen tissues show low *SMARCAL1* expression that correlates with high c-*myc* expression. Renal dysfunctions as well as impaired brain development are symptoms observed in SIOD patients[15,46]. We, therefore, hypothesize that some of the symptoms observed in SIOD patients might be due to the inability of the mutant SMARCAL1 to regulate c-*myc* expression.

This is the first report of SMARCAL1 regulating c-*myc* expression. Many questions are unanswered. For example, what is the conformational change effected in the promoter region and how does it block RNAPII interaction. Further, how is SMARCAL1 recruited to the promoter? In the companion paper, we show that SMARCAL1 can interact with both RNAPII and histone H3. It is, thus, possible that the protein is recruited via interaction with a specific modification present on histones or it is recruited via RPA as it happens during DNA repair. We also do not know the protein partners that interact with SMARCAL1 for regulating transcription. We also do not know how many genes SMARCAL1 regulates. Further experiments would enable us to answer these questions.

Methods

Reagents.   Dulbecco's Modified Eagle's Medium (DMEM), penicillin-streptomycin cocktail, amphotericin B, sodium bicarbonate, TRIzol reagent, Hoechst 33342, and Escort transfection reagent were purchased from Sigma-Aldrich (USA). Fetal bovine serum was purchased from Gibco (USA). Restriction enzymes, Turbofect, M-MuLV Reverse Transcriptase kit, Hi-fidelity PCR enzyme mix, FastAP thermo sensitive Alkaline Phosphatase and the INSTAclone TA-cloning kit were purchased either from MBI Fermentas (USA) or from NEB (USA). QIAquick gel extraction kit was purchased from Qiagen (USA). Protein A-CL agarose bead resin was purchased from Bangalore Genei (India). 2X SYBR Green PCR master mix, micro-amp Fast 96-well reaction plates (0.1 ml) and micro-amp optical adhesive films were purchased either from Applied Biosystems (USA) or Kapa Biosystems (USA). For western blotting, Immobilon-P PVDF membrane was purchased from Merck-Millipore (USA). X-ray films, developer, and fixer were from Kodak (USA).

Antibodies.   The various primary antibodies, unless otherwise mentioned, were purchased from Sigma-Aldrich (USA), Cell Signalling Technology (USA), Abcam (UK) or Bangalore Genei (India). The HRP-conjugated anti-mouse IgG and anti-rabbit IgG antibodies were obtained from Bangalore Genei (India). The catalog # of antibodies used in this study are as follows: RNAPII (Rpb1 CTD, Cell Signaling Technology, Cat #2629; 1: 6500 dilution), c-MYC (Santa Cruz, sc-40; 1: 2500), H3 (Cell Signaling Technology, Cat #3638), RPA (Cell Signaling Technology, Cat #2208), BRG1 (Sigma Aldrich, Cat #B8184: 1: 5000), and β-actin (Sigma Aldrich, Cat #A1978: 1: 10,000 dilution). SMARCAL1 antibody (1: 1800 dilution) was raised against the N-terminus HARP domain as discussed previously[35].

Primers.   The list of primers used in cloning, quantitative real-time RT-PCR, ChIP and protein occupancy assay is given in Supplementary Table S3-6 respectively. All oligonucleotides were synthesized by Sigma-Aldrich (USA).

Cell culture and transfection.   HeLa and K562 cells obtained from NCCS were cultured in Dulbecco's modified Eagle's medium and RPMI1640 respectively supplemented with 10% fetal bovine serum and 1% antibiotic cocktail. HeLa cells seeded to a confluency of 50–70% were transfected with various plasmid constructs.

Preparation of *SMARCAL1* downregulated cell line.   HeLa cells were seeded at 50–70% confluency and transfected with ShRNA clones obtained from Sigma Aldrich (USA). Stably transfected cells were selected using DMEM supplemented with 2 μg/ml puromycin. These cells, termed as Sh, were further subjected to clonal selection and three clones-Sh1, Sh2, and Sh3 were obtained.

Construction of pGL3-*myc* promoter construct.   The c-*myc* promoter was amplified from HeLa genomic DNA using specific primers with Hi-fidelity PCR enzyme. The 765 bp amplified product was cloned into pTZ57R/T vector. The clone was confirmed by restriction digestion followed by sequencing. c-*myc* promoter region was then released from the T/A clone and cloned into pGL3 basic vector using SacI and NheI restriction sites. The construct was confirmed using restriction digestion and used for further analysis.

Cloning of Myc_B$_{159}$ and Myc_C$_{201}$ in pGL3 promoter vector.   The Myc_B$_{159}$ and Myc_C$_{201}$ regions were amplified with Pfu DNA polymerase using pGL3-*myc* promoter as template with specific primers. The amplification products were digested with SacI/KpnI and were cloned into similarly double-digested pGL3 promoter vector. The constructs were confirmed using restriction digestion and used for further analysis.

RNA isolation and cDNA preparation.   Total RNA was extracted using the TRIzol reagent (Sigma-Aldrich). 90% confluent cells in a 35 mm plate were lysed with 1 ml of the TRIzol reagent to give a homogenized lysate. The lysate was transferred to a tube. 200 μl of chloroform was added to each tube per ml of TRIzol reagent, shaken vigorously and allowed to stand for 10–15 minutes at room temperature. The samples were centrifuged at 11,000 rpm for 15 minutes at 4 °C. The top aqueous layer obtained was transferred to a fresh tube and 0.5 ml of isopropanol was added per ml of TRIzol reagent, mixed and allowed to stand at room temperature for 10–15 minutes. The samples were then centrifuged at 11,000 rpm for 10 minutes at 4 °C. The RNA pellet obtained was washed with 70% ethanol and resuspended in DEPC-treated water. RNA concentrations were determined using NanoDrop 2000 (Thermo Fisher Scientific, USA) and equal amount of RNA from various samples was used for preparing cDNA using random hexamer primers according to the manufacturer's protocol. The prepared cDNA was checked for quality by performing a PCR using suitable primers.

Quantitative real-time RT-PCR.   Quantitative real-time RT-PCR was performed with 7500 Fast Real-Time PCR system (ABI Biosystems, USA) using gene-specific primers designed for exon-exon junctions. For each reaction 15 μl of samples were prepared in triplicates and the data obtained was analyzed using Fast7500 software provided by manufacturer. The p-value was calculated using Sigma-Plot (Sigma-Plot, USA).

Cell extract preparation for western blot.   Cell extracts were made using either RIPA lysis buffer (50 mM Tris-Cl pH 7.5, 300 mM NaCl, 2 mM EDTA, 1% v/v NP-40, 0.5% w/v sodium deoxycholate, 1% w/v sodium dodecyl sulphate) or urea lysis buffer (90% 8.8 M urea, 2% 5 M NaH$_2$PO$_4$ and 8% 1 M Tris-Cl pH 8.0). Briefly, cells were grown in 100 mm culture dishes to a confluency of 75–80%, harvested and thoroughly washed thrice with PBS. The cells were pelleted at 2500 rpm for 10 minutes at 4 °C and then resuspended in the appropriate lysis buffer. The cells were incubated on ice for 15 minutes with regular mild tapping followed by sonication (5 cycles of 30 sec on/off). The sonicated cell suspension was spun at 13000 rpm for 10 minutes at 4 °C. The supernatant was collected and used for further experiments. The protein concentration was determined using Bradford reagent.

Chromatin Immunoprecipitation (ChIP).   ChIP was performed according to the X-ChIP protocol provided online by Abcam (http://www.abcam.com/ps/pdf/protocols/x_CHip_protocol.pdf) with minor modifications. Briefly, cells were cross-linked by adding formaldehyde (final concentration 1%) and later quenched by adding glycine (final concentration 125 mM) to the media. The cells were then washed thoroughly using ice-cold PBS, scraped into 1 ml PBS and collected in eppendorf tubes. Cells were pelleted at 2500 rpm at 4 °C for 10 minutes. The pelleted cells were treated with freshly prepared lysis buffer (10 mM Tris-Cl, pH 8.0; 140 mM NaCl; 1 mM EDTA, pH 8.0; 1% Triton-X100; 0.1% sodium deoxycholate; 0.1% sodium dodecyl sulphate, 1 mM PMSF and protease inhibitor cocktail 10 μl/plate) for 10 minutes at 4 °C followed by sonication using a water-sonicater (40 cycles of 30 sec pulse/ 20 sec rest). The sonicated samples were centrifuged and the supernatant was used for further analysis. 50 μl of the sonicated sample was purified and the DNA concentration was determined. A small part of the purified DNA was run on 1.5% agarose gel to check for sonication efficiency. The remaining DNA was stored to be used as the "Input" sample. An equal amount of chromatin was taken for performing IP using various antibodies. One sample was kept as Beads-IgG negative control. Pre-adsorbed protein A bead resin (pre-adsorbed with 75 ng/μl sonicated salmon sperm DNA and 0.1 μg/μl of BSA) and 5 μg of the desired antibody was added to each sample. The cocktails were incubated overnight at 4 °C on an eppendorf-rotator. This was followed by washing the pelleted bead resin 3 times in wash buffer (0.1% (w/v) sodium dodecyl sulphate; 1% Triton X-100; 2 mM EDTA, pH 8.0; 150 mM NaCl; 20 mM Tris-Cl, pH 8.0) and once in final wash buffer (0.1% sodium dodecyl sulphate; 1% Triton X-100; 2 mM EDTA, pH 8.0; 500 mM NaCl; 20 mM Tris-Cl, pH 8.0). The bound DNA was eluted using fresh elution buffer (1% sodium dodecyl sulphate; 100 mM NaHCO$_3$). The eluted DNA was purified using phenol-chloroform and precipitated as mentioned in protocol. The resuspended DNA was used for ChIP PCR using standardized primers.

Serum starvation assay.   HeLa cells were cultured in DMEM with 10% FBS. For starvation assay, cells were cultures with 0.4% serum for 48 hours. Cells were released from serum starvation by adding 10% FBS and harvested at indicated time points for analysis.

Differentiation of K562 using PMA.   K562 cells were cultured in RPMI1640 media containing 10% FBS and 1% antibiotic cocktail. For PMA-induced differentiation, cells were treated with 10 nM PMA for atleast 24 hours. As a control, cells were mock-treated with DMSO.

Promoter accessibility assay.   The assay was executed using the method described by Infante *et al.* [47]. Briefly, HeLa cells grown to 2×10^7 to 3×10^7 cells were cross-linked with formaldehyde (1% (v/v) final concentration) for 15 min at room temperature and later quenched with glycine (125 mM final concentration). Cells were then washed with 10 ml cold PBS twice at 4 °C, scraped out, resuspended in NPS buffer (0.5 mM spermidine, 0.075%

NP-40, 50 mM NaCl, 10 mM Tris-Cl, pH 7.5, 5 mM MgCl$_2$, 1 mM CaCl$_2$, and 1 mM β-mercaptoethanol) and digested with 30 units of micrococcal nuclease for 15 min at 37 °C. Digestions were stopped by shifting the tubes to 4 °C and adding EDTA and EGTA to final concentrations of 15 mM and 2.3 mM respectively. Subsequently, the digested samples were treated with 60 μl 10% (v/v) SDS, 10 mg/ml proteinase K and 10 μl of 10 mg/ml RNase for 15 min at 37 °C. The DNA was extracted twice with phenol saturated with 0.1 M Tris-Cl, pH 7.5 and once with equal volume of chloroform. The DNA was precipitated with 0.1 volume of 3 M NaOAc, pH 5.3 and 2.5 volume of 100% ice-cold ethanol. The precipitated DNA was resuspended in TE pH 8.0 buffer and analyzed on a 1.5% agarose gel. The mono nucleosomes were purified from the gel and subjected to RT- PCR using primers designed for the c-*myc* promoter region (Supplementary Table S6).

The amount of protein-protected DNA for each primer pair was measured as a ratio between MNase digested and undigested genomic DNA and then normalization was done with the highest amount of protein protected DNA. We have assumed that the genomic DNA purified from the cells was devoid of proteins and therefore, this DNA was not digested with MNase. Each experiment was done in triplicates and average data along with standard deviation has been reported. Further, the p-value (Sigma-Plot, USA) was calculated to determine the statistical significance.

Dual-luciferase reporter assay.   Equal numbers of HeLa cells were seeded in 12-well plates and co-transfected, with pGL3-c-*myc* promoter or pGL3-empty vector and pRL-TK, using turbofect. The luciferase assay was performed 36 hours after transfection using the Dual-luciferase reporter assay kit (Promega) and the luciferase activity was measured and normalized with respect to the controls.

ATPase assays.   ATPase assays were performed using purified ADAAD as published previously[32].

CD spectra.   CD spectra were recorded using Chirascan (Applied Photophysics). Briefly, CD spectra of DNA were recorded in 1 mM sodium phosphate buffer (pH 7.0) in the presence of ATP, Mg^{+2}, and ADAAD. The concentrations of these reagents are indicated in the figure legends. For each experiment, 5 scans were taken at 37 °C. Spectra of appropriate buffer conditions were also taken at each time point. The spectra reading for each condition was subtracted from the appropriate buffer reading and plotted as a function of the wavelength. The CD values were converted to Mean Residue Ellipticity [θ] using the following equation:

$$[\theta] = S \times mRw/(10 \times c \times l)$$

Where S is the CD signal, c is the concentration of DNA in M, l is pathlength in cm, and mRw is the mean residue weight given by mRw = molecular weight of the oligonucleotide/number of bases in the oligonucleotide.

For ATPase assays as well as for CD spectra, Myc_B$_{159}$ was amplified using appropriate primers. The amplicon was agarose gel-purified and used in these experiments. DNA fragments used for the assays were either used directly without any heat/cool treatment or were heated at 95 °C for 3 minutes followed by slow/fast cooling. For assays requiring K$^+$, 1 M KCl was added to the purified DNA to a final concentration of 100 mM.

References

1. Flaus, A. & Owen-Hughes, T. Mechanisms for ATP-dependent chromatin remodelling: the means to the end. *FEBS J.* **278,** 3579–3595 (2011).
2. Narlikar, G. J., Sundaramoorthy, R. & Owen-Hughes, T. Mechanisms and functions of ATP-dependent chromatin-remodeling enzymes. *Cell* **154,** 490–503 (2013).
3. Watanabe, S., Radman-Livaja, M., Rando, O. J. & Peterson, C. L. A Histone Acetylation Switch Regulates H2A.Z Deposition by the SWR-C Remodeling Enzyme. *Science* **340,** 195–199 (2013).
4. Baradaran-Heravi, A. *et al.* Penetrance of biallelic SMARCAL1 mutations is associated with environmental and genetic disturbances of gene expression. *Hum. Mol. Genet.* **21,** 2572–87 (2012).
5. Coleman, M. A., Eisen, J. A. & Mohrenweiser, H. W. Cloning and characterization of HARP/SMARCAL1: a prokaryotic HepA-related SNF2 helicase protein from human and mouse. *Genomics* **65,** 274–82 (2000).
6. Flaus, A., Martin, D. M. A., Barton, G. J. & Owen-Hughes, T. Identification of multiple distinct Snf2 subfamilies with conserved structural motifs. *Nucleic Acids Res.* **34,** 2887–2905 (2006).
7. Muthuswami, R., Truman, P. A., Mesner, L. D. & Hockensmith, J. W. A eukaryotic SWI2/SNF2 domain, an exquisite detector of double-stranded to single-stranded DNA transition elements. *J. Biol. Chem.* **275,** 7648–7655 (2000).
8. Nongkhlaw, M., Dutta, P., Hockensmith, J. W., Komath, S. S. & Muthuswami, R. Elucidating the mechanism of DNA-dependent ATP hydrolysis mediated by DNA-dependent ATPase A, a member of the SWI2/SNF2 protein family. *Nucleic Acids Res.* **37,** 3332–3341 (2009).
9. Ciccia, A. *et al.* The SIOD disorder protein SMARCAL1 is an RPA-interacting protein involved in replication fork restart. *Genes Dev.* **23,** 2415–2425 (2009).
10. Yusufzai, T., Kong, X., Yokomori, K. & Kadonaga, J. T. The annealing helicase HARP is recruited to DNA repair sites via an interaction with RPA. *Genes Dev.* **23,** 2400–2404 (2009).
11. Bétous, R. *et al.* Substrate-selective repair and restart of replication forks by DNA translocases. *Cell Rep.* **3,** 1958–1969 (2013).
12. Betous, R. *et al.* SMARCAL1 catalyzes fork regression and Holliday junction migration to maintain genome stability during DNA replication. *Genes Dev.* **26,** 151–162 (2012).
13. Bansbach, C. E., Bétous, R., Lovejoy, C. A., Glick, G. G. & Cortez, D. The annealing helicase SMARCAL1 maintains genome integrity at stalled replication forks. *Genes Dev.* **23,** 2405–2414 (2009).
14. Yusufzai, T. & Kadonaga, J. T. HARP is an ATP-driven annealing helicase. *Science* **322,** 748–750 (2008).
15. Boerkoel, C. F. *et al.* Mutant chromatin remodeling protein SMARCAL1 causes Schimke immuno-osseous dysplasia. *Nat. Genet.* **30,** 215–20 (2002).
16. Lou, S., Lamfers, P., McGuire, N. & Boerkoel, C. F. Longevity in Schimke immuno-osseous dysplasia. *J. Med. Genet.* **39,** 922–5 (2002).
17. Huang, C. *et al.* Deficiency of SMARCAL1 causes cell cycle arrest and developmental abnormalities in zebrafish. *Dev. Biol.* **339,** 89–100 (2010).
18. Marcu, K. B., Bossone, S. A. & Patel, A. J. Myc function and regulation. *Annu. Rev. Biochem.* **61,** 809–860 (1992).
19. Levens, D. L. Reconstructing MYC. *Genes Dev.* **17,** 1071–7 (2003).
20. Barna, M. *et al.* Suppression of Myc oncogenic activity by ribosomal protein haploinsufficiency. *Nature* **456,** 971–975 (2008).

21. González, V. & Hurley, L. H. The c-MYC NHE III(1): function and regulation. *Annu. Rev. Pharmacol. Toxicol.* **50,** 111–29 (2010).
22. Meyer, N. & Penn, L. Z. Reflecting on 25 years with MYC. *Nat. Rev. Cancer* **8,** 976–990 (2008).
23. Levens, D. How the c-myc promoter works and why it sometimes does not. *J. Natl. Cancer Inst. Monogr.* **2008,** 41–3 (2008).
24. Dai, J., Hatzakis, E., Hurley, L. H. & Yang, D. I-motif structures formed in the human c-MYC promoter are highly dynamic-insights into sequence redundancy and I-motif stability. *PloS One* **5,** e11647 (2010).
25. Mathad, R. I., Hatzakis, E., Dai, J. & Yang, D. c-MYC promoter G-quadruplex formed at the 5'-end of NHE III1 element: insights into biological relevance and parallel-stranded G-quadruplex stability. *Nucleic Acids Res.* **39,** 9023–33 (2011).
26. Yang, D. & Hurley, L. H. Structure of the biologically relevant G-quadruplex in the c-MYC promoter. *Nucleosides Nucleotides Nucleic Acids* **25,** 951–68 (2006).
27. Avigan, M. I., Strober, B. & Levens, D. A far upstream element stimulates c-myc expression in undifferentiated leukemia cells. *J. Biol. Chem.* **265,** 18538–18545 (1990).
28. Chi, T. H. *et al.* Sequential roles of Brg, the ATPase subunit of BAF chromatin remodeling complexes, in thymocyte development. *Immunity* **19,** 169–182 (2003).
29. Liu, J. *et al.* The FUSE/FBP/FIR/TFIIH system is a molecular machine programming a pulse of c-myc expression. *EMBO J.* **25,** 2119–30 (2006).
30. Wang, Z. *et al.* Combinatorial patterns of histone acetylations and methylations in the human genome. *Nat. Genet.* **40,** 897–903 (2008).
31. Kikin, O., D'Antonio, L. & Bagga, P. S. QGRS Mapper: a web-based server for predicting G-quadruplexes in nucleotide sequences. *Nucleic Acids Res.* **34,** W676–W682 (2006).
32. Nongkhlaw, M., Gupta, M., Komath, S. S. & Muthuswami, R. Motifs Q and I are required for ATP hydrolysis but not for ATP binding in SWI2/SNF2 proteins. *Biochemistry (Mosc.)* **51,** 3711–3722 (2012).
33. Zuker, M. Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res.* **31,** 3406–3415 (2003).
34. Kypr, J., Kejnovska, I., Renciuk, D. & Vorlickova, M. Circular dichroism and conformational polymorphism of DNA. *Nucleic Acids Res.* **37,** 1713–1725 (2009).
35. Dutta, P. *et al.* Global epigenetic changes induced by SWI2/SNF2 inhibitors characterize neomycin-resistant mammalian cells. *PloS One* **7,** e49822 (2012).
36. Muthuswami, R. *et al.* Phosphoaminoglycosides inhibit SWI2/SNF2 family DNA-dependent molecular motor domains. *Biochemistry (Mosc.)* **39,** 4358–4365 (2000).
37. Zimmerman, K. A. *et al.* Differential expression of myc family genes during murine development. *Nature* **319,** 780–783 (1986).
38. Cañelles, M. *et al.* Max and inhibitory c-Myc mutants induce erythroid differentiation and resistance to apoptosis in human myeloid leukemia cells. *Oncogene* **14,** 1315–1327 (1997).
39. Huo, X.-F. *et al.* Differential expression changes in K562 cells during the hemin-induced erythroid differentiation and the phorbol myristate acetate (PMA)-induced megakaryocytic differentiation. *Mol. Cell. Biochem.* **292,** 155–167 (2006).
40. Shelly, C., Petruzzelli, L. & Herrera, R. PMA-induced phenotypic changes in K562 cells: MAPK-dependent and -independent events. *Leukemia* **12,** 1951–1961 (1998).
41. Phan, A. T., Modi, Y. S. & Patel, D. J. Propeller-Type Parallel-Stranded G-Quadruplexes in the Human *c-myc* Promoter. *J. Am. Chem. Soc.* **126,** 8710–8716 (2004).
42. Gonzalez, V., Guo, K., Hurley, L. & Sun, D. Identification and Characterization of Nucleolin as a c-myc G-quadruplex-binding Protein. *J. Biol. Chem.* **284,** 23622–23635 (2009).
43. Postow, L., Woo, E. M., Chait, B. T. & Funabiki, H. Identification of SMARCAL1 as a Component of the DNA Damage Response. *J. Biol. Chem.* **284,** 35951–35961 (2009).
44. Elizondo, L. I. *et al.* Schimke immuno-osseous dysplasia: SMARCAL1 loss-of-function and phenotypic correlation. *J. Med. Genet.* **46,** 49–59 (2009).
45. Coleman, M. A., Miller, K. A., Beernink, P. T., Yoshikawa, D. M. & Albala, J. S. Identification of chromatin-related protein interactions using protein microarrays. *Proteomics* **3,** 2101–2107 (2003).
46. Deguchi, K. *et al.* Neurologic phenotype of Schimke immuno-osseous dysplasia and neurodevelopmental expression of SMARCAL1. *J. Neuropathol. Exp. Neurol.* **67,** 565–577 (2008).
47. Infante, J. J., Law, G. L. & Young, E. T. Analysis of nucleosome positioning using a nucleosome-scanning assay. *Methods Mol. Biol. Clifton NJ* **833,** 63–87 (2012).

Acknowledgements

Author Contributions

R.M and T.S. conceived the experiments, which were executed by T.S., D.T.H. and R.B. R.M. wrote the paper.

Additional Information

Supplementary information accompanies this paper at http://www.nature.com/srep

Competing financial interests: The authors declare no competing financial interests.

How to cite this article: Sharma, T. *et al.* SMARCAL1 negatively regulates c-myc transcription by altering the conformation of the promoter region. *Sci. Rep.* **5,** 17910; doi: 10.1038/srep17910 (2015).