CHAPTER 1

Introduction
Chapter 1: Introduction

In a eukaryotic cell the DNA is assembled as a nucleoprotein complex chromatin, which is a highly compact structure made of repeating units of nucleosomes. This structure restricts the interaction of DNA with nuclear factors during processes like transcription, replication, recombination and repair where nuclear factors need to access the DNA molecule. In an actively dividing eukaryotic cell chromatin disruption occurs during DNA replication (Flanagan and Peterson, 1999, Falbo and Shen, 2006), DNA repair (Osley et al. 2007), recombination (Cherry and Baltimore 1999), and transcription (Jones and Kadonaga, 2000; Tyler and Kadonaga, 1999; Wu and Winston, 1997). Two important classes of enzymes have been employed by the cell to ensure the smooth functioning of the cell machinery: i) histone modifying enzymes; and ii) ATP-dependent chromatin remodelers. Both these groups of proteins transiently disrupt the nucleosomal structure by mobilizing the nucleosomes and thus making it accessible to the various nuclear factors. In this chapter, I will be discussing the chromatin organization and the ATP-dependent chromatin remodelling in detail.

Chromatin organization

The eukaryotic genome is packaged into chromatin allowing huge DNA to fit into the small nucleus. The initial step for the condensation is the packing of the DNA in each octamer to form the nucleosome, which is made up of the four histone proteins i.e. H2A, H2B, H3 and H4 forming an octamer around which 146 base pair of DNA is wrapped. The histone molecules carry a net positive charge, which helps in packaging of negatively charged DNA (Wolffe, 1998). The crystal structure of the histone octamer possesses a tripartite organization i.e. a centrally located dimer of (H3-H4) flanked by two H2A-H2B dimers (Arents et al.1991). These studies have further revealed that 146 base pair of DNA is wrapped around the octamer with a
1.65 superhelical turn, leading to about 6-fold packing of the DNA in the cell (Luger et al. 1997; Kornberg, 1974; Schalch et al. 2005). The nucleosomes are separated approximately by 200 base pair DNA from each other forming a ‘bead on a string’ like structure or a linear nucleosome array. Depending upon the ionic strength of the solution the nucleosomes interact with each other forming a helix like structure in which 6-8 nucleosomes are arranged per turn of the helix leading to the formation of the 30 nm chromatin fiber (Hayes and Hansen, 2001). The nucleosomal arrays are also found to oligomerize forming large oligomeric structures in presence of high concentration of divalent cations (e.g. MgCl₂) (Schwarz et al. 1996). Oligomerization of the nucleosomal arrays may be one of the pre-requisite stages for the formation of chromonema fibres (Hansen, 2002). Also, in vitro, oligomerization provides substantial evidences for the occurrence of long-range transitions that leads to global condensation of the interphase chromosome (Belmont et al. 1994, Schwarz et al. 1996). The chromatin fibers and nucleosomal oligomers are also found to be further stabilized by the linker histone (Thoma et al. 1979; Carruthers et al. 1998). The linker histone interacts with the chromatin fiber and/or oligomers through their highly basic C-terminal domain probably by binding to the linker DNA (Allan et al. 1986). The globular domain of the linker histone plays a role in stabilizing the peripheral nucleosomal DNA and its tendency to self-associate may promote particular fiber-fiber interaction (Hayes and Hansen, 2001). Higher order structure of chromatin condensation is achieved by various inter-nucleosomal interactions mediated by the N-terminal domain of histones (Tse and Hansen, 1997). The core histones tail have been found to be important for the fiber-fiber interaction even in the absence of the linker histone (H1) and the loss of histones tail has been shown to hamper the formation of 30 nm fiber (Tse and Hansen, 1997). Studies with tailless histone, where the histone tails were proteolytically cleaved, have revealed the
importance of these domains (Tse and Hansen, 1997). Nucleosomal arrays constructed from these tailless histones or from hybrid histone proteins, where the N-terminal were deleted, were unable to form the maximum folded structures or oligomerize in the presence of high concentrations of MgCl₂ (Hansen, 2002). The N-terminal domains of histones are involved in chromatin condensation through non-electrostatic interaction. These proteins have been shown to mediate the formation of a condensed chromatin through protein-protein interaction with other chromatin-associated proteins. For example, Sir2p and Tup1p were found to interact with histone H3 and H4 amino terminal domain (Edmonson et al. 1996; Hong et al. 1993). These proteins promote higher order chromatin condensation by their ability to self-associate and thereby forming crosslink between two chromatin fibers (Hansen, 2002). Another heterochromatin protein called MENT forms inter-fiber bridges by binding to the linker DNA (Hansen, 2002). The packaging of DNA into histone octamer, nucleosomal interactions, oligomerization of the nucleosomal arrays, inter-fiber cross linking and other unknown processes ensures that the DNA is successfully packaged into the relatively small nucleus and thereby rendered inaccessible to various nuclear factors. However, for the cell to carry out various cellular activities that uses DNA as the substrate, it is important for the cell to get access to the DNA. Therefore, eukaryotic cells have evolved certain mechanisms to overcome this barrier.

**Chromatin remodelling**

The eukaryotic cell, as emphasised in the previous section, has devised various mechanisms to alter the chromatin structure of which the most prevalent are covalent modification of the histone protein and histone mobilization. Covalent modifications of the histones include acetylation, methylation, phosphorylation, ubiquitination, sumoylation, and ADP-ribosylation (Wolffe and Hayes, 1999; Kouzarides, 2007; Li et al. 2007). On the other
hand, mobilization of histones can occur either by removing the whole octamer or any one of the subunit from the nucleosomal array or sliding the octamer along the DNA. These processes are carried out by a group of proteins called ATP-dependent chromatin remodelers.

In this study I will be focusing on the ATP-dependent chromatin remodeling process.

**ATP-dependent Chromatin Remodelling**

ATP-dependent chromatin remodelling, as the name suggests, utilizes the energy from ATP hydrolysis to remodel the chromatin. The ATP-dependent chromatin remodelling proteins are DNA-dependent ATPases containing highly conserved helicase motifs. These proteins belong to the DEAD/H superfamily of nucleic acid stimulated ATPase, and are responsible for the hydrolysis of ATP in the presence of DNA or nucleosomes. The DEAD/H superfamily of nucleic acid stimulated ATPase is a large protein superfamily that includes helicases and is classified into superfamily 1, 2, 3, and 4. Of these, the superfamily 2 (SF2) comprises of those proteins that possess the seven helicase motifs. It is the largest of the four superfamilies of helicases and includes a variety of DNA helicases and several families of RNA helicases as well as the SWI/SNF proteins (Gorbalenya et al. 1988). However, despite the presence of the helicase motifs, the SWI/SNF proteins do not possess helicase activity (Laurent et al. 1993).
Proteins of the SWI/SNF Family

Proteins belonging to the SWI/SNF family show great similarity in the helicase domain but the sequence outside the helicase domain vary from one protein to the other and hence Eisen et al. initially divided this family into seven subfamilies on the basis of sequence characteristic outside the conserved helicase domain and their function (Eisen et al. 1995) (Figure 1.1).

Recently, Flaus et al. have re-classified the SWI/SNF proteins into 24 distinct subfamilies (Flaus et al. 2006).

Snf2 family

The Snf2-like subfamily consists of proteins with helicase like region similar to the primary sequence to *S. cerevisiae* Snf2p. In addition to the homology within the helicase domains, these proteins also contain the bromodomain, which can interact with acetylated histones (Hassan et al. 2002). Furthermore, these proteins can bind to both naked DNA and nucleosomal DNA (Cote et al. 1994). Other proteins belonging to this family are Sth1, BRG1/hBRM, BRM that constitute the ATPase subunit in RSC complex, hSWI/SNF complex
and Drosophila SWI/SNF complex respectively (Du et al. 1998; Phelan et al. 1999; Elfring et al. 1994). Complexes having these SNF2-like proteins usually have been shown to have chromatin remodeling properties but interestingly BRG1 and hBRM have been shown to remodel chromatin as independent subunit in vitro (Phelan et al. 1999). The ability of BRG1 and hBRM to remodel chromatin themselves suggests that the ATPase subunit is the fundamental unit for remodelling. Besides chromatin remodelling, proteins from this family have been found to be involved in other cellular activities. BRG1 has been found to be important for transcription of CD44 and human metalloproteinase gene (Strobeck et al. 2001). Also, BRG1 interacts with tumor suppressor genes like Rb and BRCA1 and has been found to be mutated in most lung cancer cell lines (Bochar et al. 2000; Wong et al. 2000).

The SWI/SNF complex was first identified in S. cerevisiae, through genetic screens for genes involved in sucrose fermentation (SNF) and in mating type switching (SWI) (Stern et al. 1984, Carlson et al. 1981). Mutations in SNF genes lead to defect in SUC2 expression (Carlson et al. 1981) and similarly mutations in SWI genes lead to defect in HO expression (Stern et al. 1984). Mutations in both SNF and SWI genes affect the expression of various genes indicating that these proteins play an important role in gene expression (Carlson and Laurent, 1994). Further genetic analysis and sequence identity led to the recognition that SNF2 (SWI2), SNF5, SNF6, SWI1 and SWI3 constitute a group of functionally related proteins (Cairns et al. 1994). The remaining six subunits, SWP82, SWP73, SWP 61, SWP59, TAFII30 and SNF1, were found in a biochemically isolated complex (Burns and Peterson, 1997). These genes are non-essential for viability, but mutations affect transcription of a broad range of differently regulated genes, implying a global role in transcription control (Carlson and Laurent, 1994).
The SWI/SNF complex in *S. cerevisiae* is a large multi-subunit complex comprising of 11 subunits. The motor of this complex is the nucleosome remodeling ATPase -Snf2p-which is a 200-kDa protein (Laurent *et al*. 1993; Whitehouse *et al*. 1999). The function of many of the other subunits is less understood. SNF5 is involved both in assembly and catalytic functions of the complex and is also conserved in higher eukaryotes (Geng *et al*. 2001). Homology searches with the sequence of the yeast Snf2p ATPase and biochemical analysis led to the identification of several SWI/SNF-related nucleosome remodeling machines in higher eukaryotes (Okabe *et al*. 1992). The *D. melanogaster* and mammalian complexes all contain subunits that are homologous to the yeast SWI/SNF, SWI3, SNF5, and SWP73 subunits as well as the above-mentioned related proteins. The yeast SWI/SNF complex was the first complex to be associated with chromatin remodelling activity through the isolation of histones encoding gene that suppresses SWI/SNF mutation (Hirschhorn *et al*. 1992). Their role in chromatin remodeling was further strengthened when it was found that SWI/SNF complex increases nuclease activity at the *SUC2* promoter (Hirschhorn *et al*. 1992; Wu and Winston, 1997). The SWI/SNF complex also shows ATP-dependent nucleosomal disruption *in vitro* indicating that this complex could play a role in nucleosomes disruption (Kwon *et al*. 1994).

As mentioned earlier, other members of the SWI/SNF complex include the BRG1/hBRM containing human SWI/SNF complex (Kwon *et al*. 1994), Drosophila BRM (Dingwall *et al*. 1995) containing complex and the STH1 containing RSC complex (Cairns *et al*. 1996). These complexes have an ATPase subunit that belongs to Snf2 subfamily and they hydrolyze ATP in the presence of naked as well as nucleosomal DNA (Laurent *et al*. 1993, Whitehouse *et al*. 1999). The catalytic subunit is very crucial for the function of these complexes as mutations in the ATPase domain abrogates the ability of the complexes to carry
out their functions (Richmond and Peterson, 1996). However, BRG1 and BRM alone have the capability to mobilizes and disrupt nucleosome and nucleosomal arrays almost like the intact complex (Phelan et al. 1999). Addition of other subunits like hSnf5 (INI) and hSwi (BAF170/BAF150) restores the activity of BRG1 to that of the intact hSWI/SNF complex (Phelan et al. 1999; Phelan et al. 2000). The function of the other subunits are not well known, however Snf5 and Swi3 have been reported to interact with various transcription factors like Gal4-VP16, Swi5p, Gcn4p and Hap4p (Neely et al. 2002) indicating that these subunits may be responsible for recruiting the complex to the site of transcription. Swi3p has been shown to be important for H2A-H2B displacement in an ATP-dependent manner (Yang et al. 2007). Swi1 contains an AT-rich interaction domain (ARID) that allows nonspecific DNA binding (Wilsker et al. 2004). The SWI/SNF complexes in yeast and higher eukaryotes also contain the actin-related proteins Arp7 and Arp9. It has been suggested that these actin-related proteins might provide a link to nuclear structures, such as the nuclear matrix (Becker and Horz, 2002).

Figure 1.2: Grouping of SWI/SNF family of proteins (Flaus et al. 2006)
ISWI family

The Drosophila ISWI (imitation SWI) was the first protein to be identified in the Iswi subfamily (Elfring et al. 1994). ISWI possesses two SANT (Swi3, Ada2, NCoR and TFIIIB) domains in addition to the ATPase domain (Tsukiyama et al. 1995). ISWI proteins can hydrolyze ATP only in the presence of nucleosomes (Tsukiyama and Wu, 1995). These proteins have been found to be associated with more than one complex. For example, Drosophila ISWI protein is present in three different complexes i.e. the CHRAC (Chromatin accessibility factor), ACF (ATP-dependent chromatin assembly and remodeling factor) and NURF (Nucleosome remodeling factor) complex and similar properties were also observed in case of human, yeast and Xenopus (Varga-Weisz et al. 1997; Tsukiyama et al. 1995; Ito et al. 1997). The human homologs of Drosophila ISWI are SNF2H and SNF2L whereas yeast has two ISWI genes, ISWI1 and ISWI2 (Tsukiyama et al. 1999). SNF2H shares 70% sequence homology with Drosophila ISWI (Aihara et al. 1998). In humans, the complexes that possess SNF2H are hACF, WICH, hCRAC, RSF and SNF2H/NURD/cohesion SNF2L (Dirscherl and Krebs, 2004). The ability of ISWI protein to associate with more than one complex indicates that it can perform diverse functions using its catalytic property depending on the subunits with which it interacts. ISWI associated complexes have been found to have diverse functions that ranges from activation/repression of transcription, chromatin assembly and replication (Langst and Becker, 2001). ISWI associated complex mainly mediates chromatin remodeling through nucleosome sliding and ISWI alone has the ability to catalyze nucleosome sliding but its activity is enhanced by the presence of other subunits (Hamiche et al. 1999).
CHD subfamily

The CHD subfamily comprises those proteins that contain the chromodomain (chromatin organization modifier) in the N-terminal and the SNF2-like ATPase domain in the central region as the signature motifs (Delmas et al. 1993). However, some proteins in this family have additional domains like DNA binding, SANT-like, BRK (Brahma and Kismet), and PHD (plant homeo domain) domains (Marefella and Imbalzano, 2007). The chromodomain is useful for interaction with chromatin by directly binding with DNA, RNA and methylated histones (Marfella and Imbalzano, 2007). Proteins from this family have been shown to form chromatin remodeling complexes which have histones deacetylase as one of the subunit and hence these complexes have both ATPase and deacetylase activity (e.g. NuRD complex) (Marefella and Imbalzano, 2007).

Swr1 subfamily

The Swr1-like subfamily consists of the Swr1, Ino80, and Etl1. Swr1p (SWI/SNF-related protein) from S. cerevisiae is part of the large SWR1 complex that exchanges histone H2A.Z-containing for wild-type H2A-containing dimer (Wu et al. 2005). The Ino80 complex was identified by the isolation of a novel complex from yeast that shared homology in its central domain with the Drosophila ISWI and yeast Snf2p (Shen et al. 2000). Like the SWI/SNF-complex, INO80 complex and INO80 ATPase shows DNA-dependent ATPase activity (Shen et al. 2000; Bakshi et al. 2006). Ino80 has both Drosophila (dINO80) and human (hINO80) homologs and sequence alignments reveal that these proteins contain two conserved regions beyond the motor domain, a TELY motif at the amino terminus and a GTIE motif at the carboxy terminus. The notable subunits of Ino80 complex include Rvb1 and Rvb2, which share homology with the bacterial RuvB helicase, and Act1/actin, Arp4, Arp5, and Arp8 (Shen et al.
Ino80 null mutants exhibit sensitivity to hydroxyurea, methyl methanesulfonate (MMS), ultraviolet light, and ionizing radiation, indicating roles for this complex in replication and/or the processing of DNA damage (Shen et al. 2000). Similarly, a recent study shows that deficiency in hINO80 lead to delayed S phase and also abnormal chromosomal segregation and multinucleation (Hur et al. 2010).

**SSO1653 subfamily**

SSO1653-like subfamily consists of ERCC6, SSO1635 and Mot1. Of these, Mot1 (Modifier of Transcription) does not have a direct role in chromatin remodelling but has been found to be associated with TBP to recycle it from the DNA-bound state (Auble et al. 1997). ERCC6/CSB protein and its *S. cerevisiae* homologue Rad26p is involved in transcription nucleotide excision repair coupled repair (Citterio et al. 2000). These proteins were found to be involved in helping RNA polymerase to pass or dissociate from blocking DNA lesions (Svejstrup, 2003). CSB protein has been shown to bind to histones tails to mediate ATP-dependent nucleosomal rearrangement (Citterio et al. 2000). SSO1653 is the archeal member of SNF2 protein and so far its biological function has not been understood but it does show the ability to generate DNA torsion and hydrolyze ATP in presence of DNA (Durr et al. 2005).

**Rad54 subfamily**

The Rad54-like group consists of Rad54, ATRX, Arip4, and DRD1. Rad54 hydrolyzes ATP in presence of double-stranded DNA (Tanaka et al. 2002). Rad54 can translocate on DNA, produce superhelical torsion (Petukhova et al. 1999) and increase accessibility to nucleosomal DNA (Alexiadis et al. 2004). Yeast Rad54 is involved in the repair of double strand breaks through Rad51 mediated homologous recombination (Van Komen et al. 1999). The crystal
structure of Rad54 helicase like protein from zebrafish has been determined recently (Thoma et al. 2005). ATRX protein derives its name from a genetic disorder called Alpha Thalassemia/Mental Retardation syndrome, an X-linked genetic disorder and it is involved in regulation of transcription and heterochromatin structure by moderating the accessibility of nucleosomal DNA (Flaus et al. 2006).

**Rad5/16 subfamily**

The Rad5/16 like group consists of subfamilies Rad5/16, Lodestar, Ris1 and SHPRH. The proteins in all these subfamilies, except for Lodestar subfamily, possess a RING finger (Flaus et al. 2006). The proteins in this family have various functions ranging from double-strand break repair (Rad5, Rad16 and Lodestar), to termination of transcription by removing RNA polymerase II from the transcript (TTF2), and mating-type switching (Ris1) (Jiang et al. 2004). The function of SHPRH is still unclear.

**SMARCAL1 and RapA subfamily**

The ‘distant’ members of the Snf2 families consist of SMARCAL1 and RapA family proteins. SMARCAL1 family includes SMARCAL1 and ZRANB3 proteins, which are two subtypes with similar helicase-like regions but different auxiliary domains. SMARCAL1 possesses the characteristic seven helicase motifs of Snf2 ATPase but lacks some of the trademark Snf2 domain like the conserved blocks between Motif III and IV (Flaus et al. 2006). SMARCAL1 been shown to interact with single-strand binding protein RPA during stalled replication (Ciccia et al. 2009). ADAAD, the bovine homologue of SMARCAL1 has been shown to hydrolyze ATP in the presence of DNA molecules possessing double-stranded to single-stranded transition regions (Muthuswami et al. 2000). Mutations in the helicase region of
SMARCAL1 have been implicated in a genetic disorder called Schimke immune-osseous dysplasia (SIOD) (Boerkel et al. 2002; Elizondo et al. 2009). In addition to the helicase motifs, ZRNAB3 protein possesses a zinc finger domain and a putative HNH type endonuclease domain (Flaus et al. 2006). However, nothing is known about the function of this protein.

**Mechanism of action of the ATP dependent chromatin remodelling complex**

The mechanisms by which these complexes remodel chromatin structure have not yet been clearly delineated. In the following section I will be briefly describing the various models proposed to explain chromatin remodelling by these complexes.

**ATP hydrolysis and chromatin remodeling**

The entire ATP-dependent chromatin remodeling complexes have an ATPase subunit that is homologous to the Swi2p/Snf2p ATPase. The Swi2p/Snf2p ATPase possesses the seven helicases that are the signatures of the SF2 helicase; however Swi2p/Snf2p complex does not show any helicase activity. Interestingly, loss in ATPase activity due to mutation in the conserved helicase motifs affects the remodeling activity of the complexes, indicating that the ATPase activity is very crucial for chromatin remodeling (Smith and Peterson, 2005).

**Binding of the complex to DNA and nucleosome**

As described earlier ATP hydrolysis in chromatin remodeling by SWI/SNF complex involves the stimulation by either naked DNA or nucleosome or both. Therefore, interaction of these complexes with the DNA and nucleosomes seems inevitable. The SWI/SNF complex binds to DNA and nucleosome with high affinity (Quinn et al. 1996). Snf2p can bind to naked DNA in an ATP-independent manner with $K_d$ in nanomolar range and electron spectroscopic imaging studies have shown that Snf2p binding creates loops in either nucleosomal arrays or naked DNA.
bringing distant sites into close proximity (Quinn et al. 1996; Bazett-Jones et al. 1999).

Similarly, purified ISWI and BRG1 can hydrolyze and remodel nucleosome independent of the complex, indicating that these proteins have the ability to bind to DNA (Corona et al. 1999; Phelan et al., 1999; Phelan et al. 2000). Other members of the Snf2 family like ADAAD and Rad54 too can bind to naked DNA with a similar K_\text{d} as the Snf2p (Muthuswami et al. 2000; Durr et al. 2005). These studies indicate that the ATPase subunit mainly mediates the DNA binding of the complex. The binding of the complex seems to occur through the minor groove of DNA as the complex is displaced by distamycin A or chromomycin A3, which are minor groove binding agents (Quinn et al. 1996). The affinity of Snf2p for nucleosomes is slightly higher than that for naked DNA (2-3 fold higher) (Cote et al. 1998). This is possibly due to additional interactions of the protein with the acetyl a ted histones through the bromodomain (Cote et al. 1998). In contrast, human SWI/SNF (hSWI/SNF) is able to remodel tailless mononucleosomes, as well as nucleosomal arrays, suggesting that the mechanism of nucleosome remodeling by hSWI/SNF is not dependent on the core histone tails (Guyon et al. 1999).

**ATP-dependent nucleosome disruption**

ATP-dependent chromatin remodelling involves the perturbation of the histone-DNA interaction which leads to either sliding of the histone octamer along the DNA (cis-displacement) (Hamiche et al.1999, Whitehouse et al. 1999), or transfer of the octamer from one DNA segment (trans-displacement) to the other (Phelan et al. 2000, Lorch et al. 1999), or removal of histone protein from the octamer (Yang et al. 2007), or exchange of histone proteins (Mizuguchi et al. 2004). According to crystal structure of the nucleosome (Luger et al. 1997) there are approximately 100 histone-DNA interactions that stabilize the nucleosome, so to mobilize the nucleosomes the remodeler has to overcome the energy barrier (approx 12-14 kcal mol^{-1}) that
holds the histones and DNA together. Therefore, two potential models have been put forward to explain how ATP-dependent remodelers mobilize nucleosomes. These models are: Bulge diffusion and Twist defect diffusion (Flaus and Owen-Hughes, 2004).

Bulge diffusion is the simplest mechanism for nucleosome mobilization. The DNA bulge is made by unwrapping the DNA from the histone octamer at the entry/exit of the nucleosome and then rebound at more distal sequences to the same histone contact points. The excess DNA bulges out and moves around the nucleosomal superhelix. Bulging out of DNA favours the incorporation of at least 40-60 bp of additional DNA into the nucleosomes, which is consistent with the observed sliding movement of nucleosomes both in vitro and in vivo. A nucleosome like structure having >147 bp of DNA is the most likely intermediate for this kind of model and some reports have provided support for this (Fan et al. 2003; Kassabov et al. 2003).

In twist diffusion the torsion generated on the DNA at the surface of the nucleosomes forces the adjacent region to writhe in order to balance the torsion. So in an array of nucleosomes this phenomenon is propagated along the DNA and makes the nucleosomes to move one base pair at a time along the DNA. Remodeling enzymes may act as 'molecular ratchet' or 'DNA twistase' (Cairns, 2007) that allows the twist defect to exit in one direction and results in DNA twist tension. The 'twist defect' has been observed in crystal structure that supports the presence of this model (Richmond and Davey, 2003; Edayathumangalam et al. 2005). Chromatin remodeling complexes have the ability to generate superhelical torsion in the DNA and hence might accelerate the rate of twist defect diffusion. However, presence of a nick in the DNA that could counteract the twist, does not affect the remodeling activity of ISWI and RSC. Despite the reports that support these two models, further investigations are still required to see which one predominates or whether these models work in tandem with each other.
Recently, strong evidence has come up that supports the translocation model of chromatin remodelers where it has been shown that the ATPase subunit acts as an ATP-dependent translocase (Whitehouse et al. 2003; Cairns, 2007). Single molecule studies have shown that SWI/SNF and RSC can translocate DNA at the rate of 13 bp per second and also form superhelical twist for every 100 bp translocated (Lia et al. 2006; Zhang et al. 2006; Cairns 2007). These studies also show that RSC and SWI/SNF can translocate DNA loop into the nucleosomes. Enough force is imparted to disrupt several DNA-histone contacts that is necessary for the loop formation and once the loop is formed it is propagated to the nucleosome through diffusion (Zhang et al. 2006). Based on these reports and the crystal structures an 'inch worm movement' model has been proposed which explains how the ATP-dependent remodelers translocate DNA (Cairns, 2007).

Besides DNA translocation and nucleosome sliding, SWI/SNF (Phelan et al. 2000) and RSC (Lorch et al. 1999) have the ability to transfer octamer to a free DNA accepter in an ATP-dependent manner.
dependent manner. Moreover, BRG1 alone shows the ability to transfer nucleosome (Phelan et al. 2000). Recent studies have also shown that ySWI/SNF can evict nucleosome with the help of an activation domain (Gutierrez et al. 2007). Another chromatin remodelling complex SWR1 has been shown to catalyze the exchange of H2A/H2B with H2AZ/H2B both in vivo and in vitro (Mizuguchi et al. 2004) whereas Ino80 complex have been reported to catalyze the reverse reaction (Papamichos-Chronakis et al. 2006).

These findings suggest that the ATP-dependent chromatin remodelers are able to alter the chromatin using various mechanisms. The ATPase subunit is a vital component for the function of these complexes and as this subunit possesses the conserved helicase motifs known to be important for the function these protein, therefore, understanding the functional significance of these motifs is crucial to delineate the mechanism of chromatin remodelling.

**Helicase Motifs**

Helicase motifs are the seven highly conserved amino acid sequences found in helicases including the Snf2 family of proteins (Figure 1.4). These motifs are labelled as Motif I, Ia, II, III, IV, V and VI. The Motif I is the well-known nucleotide binding motif having the conserved amino acid sequences called the GKT box or Walker A box (Walker et al. 1982). The Motif II is the DEAD box (DE- is highly conserved), a special version of the Walker B motif of nucleotide binding proteins (Linder et al. 1989). Motif Ia and a conserved tyrosine within Motif VI have been associated with presumed DNA binding proteins, suggesting an involvement of the sites in polynucleotide binding (Marintcheva and Weller, 2002). A conserved aspartate residue in Motif II has been implicated in phosphate binding via magnesium ions (Pause and Sonenberg, 1992). Motif III has been shown to be involved in interdomain interactions (Caruthers and McKay, 2002). Recently Motif III was also shown to be important RNA binding and ATP hydrolysis...
Motif VI is required for RNA binding and ATPase activity in elf4A and Has1p (Pause et al. 1993; Rocak et al. 2005) and Motif V have been reported to couple ATP hydrolysis to chromatin remodeling activity in ySWI/SNF (Smith and Peterson, 2005).

<table>
<thead>
<tr>
<th>I</th>
<th>Ia</th>
<th>II</th>
<th>III</th>
<th>IV</th>
<th>V</th>
<th>VI</th>
</tr>
</thead>
<tbody>
<tr>
<td>GxxxGKT</td>
<td>x+++xPo</td>
<td>+++DExH</td>
<td>+x+SAT</td>
<td>++Fxxoxo</td>
<td>+xTxxxxG+o+x</td>
<td>QxxGRxxR</td>
</tr>
</tbody>
</table>

Figure 1.4: Consensus sequence of SF2 helicase subfamily. A ‘+’ represents a hydrophobic residue, an ‘o’ represents a hydrophilic residue and an ‘x’ could be either a hydrophobic or a hydrophilic residue (Hall and Matson, 1999).

Crystal structure of the ATPase motor domain

The crystal structure of the ATPase motor domain of zebrafish Rad54, SWI/SNF ATPase core of S. solfataricus (SSO1653) and E. coli RapA protein have been determined (Durr et al. 2005; Thoma et al. 2005; Shaw et al. 2008). According to these studies, the motor domain is a bi-lobed structure separated by a deep cleft. Each lobe is called as the RecA-like domain and their arrangement is similar to that found in other helicases (Thoma et al. 2005). The Motif I, II and III are present in domain 1 and Motifs IV, V and VI are present in the domain 2 (Durr et al. 2005; Thoma et al. 2005). The Motifs I, II and III from domain 1 form the ATP-binding site in the active site cleft. Motif IV, V and VI on domain 2 are not situated in the active site cleft in case of SSO1653, but are positioned on the outside of the protein. This unique location is due to the -180° degree flip of the domain 2, which is very unusual in comparison to helicases as helicases do not have such conformation (Durr et al. 2005). Besides, the helicase core structure, the Snf2 family have some additional structural elements inserted into the helicase core. These comprise the antiparallel alpha helical protrusions from both recA-like domains 1 and 2, a structured linker between the recA-like domains, the major insertion region at the back side of
the domain 2 alpha-helical protrusions and a triangular brace packed against the domain 2 alpha-helical protrusions. The two alpha-helical protrusions and linker are all encoded within the enlarged span between Motifs III and IV. The triangular brace is encoded immediately downstream of Motif VI (Figure 1.5). The structure and sequence of Motif I (ATP binding) and Motif II (MgCl₂ binding) between Rad54 and other helicases are highly similar. The sequence of the remaining motifs in domain 1 (Ia, TxGx and III) are different from those in SF1/SF2 helicases but the structure suggest that these motifs have similar function. Motif III of Rad54 has sequence and structural similarity with the SF2 helicase, but contains some additional residues. In helicases, Motif III interacts with the gamma phosphate of ATP, DNA, and the second domain of the protein. This motif has been suggested to be a sensor for ATP hydrolysis (Caruthers and McKay, 2002). In helicases, the Motif IV interacts with DNA through one or two basic residues (lysine and arginine), as well as through the amide backbone group at the start of the helix (Caruthers and McKay, 2002). Rad54, in contrast, has a lysine 449 residue at spatially similar position as that present in helicases. Thus, lys 449 not only provides potential DNA contact that is otherwise absent in Motif IV but also provides a mechanism for involving domain 2 in DNA binding. The domain 2 contains a DNA binding motif (Motif IV) that is structurally conserved in SF1/SF2 helicase and Rad54. Rad54 and the remainder of the SWI/SNF do not have a conserved Arg/Lys (Lys568), Ser/Thr Ser (566), and Ser (567), and Gly (570) in the Motif V that might interact with DNA in a manner analogous to PcrA and Rep proteins (Thoma et al. 2005, Caruthers and McKay, 2002). Motif VI in Rad54 and other SWI/SNF proteins have a conserved arginine. In Rad54 crystal structure, Arg 600 interacts with the sulfate ion bound in the ATP gamma phosphate binding site of Motif I in a manner analogous to that seen in PcrA (Thoma et al. 2005).
Figure 1.5: Conserved blocks contribute to distinctive structural features of Snf2 family proteins. (A) core recA-like domains 1 and 2 including colouring of helicase motifs (I in green, Ia in blue, II in bright red, III in yellow, IV in cyan, V in teal and VI in dark red). (B) Q Motif (pink). (C) antiparallel alpha helical protrusions 1 and 2 (red) projecting from recA-like domains 1 and 2, respectively. (D) Linker spanning from protrusion 1 to protrusion 2 (middle blue). (E) Major insertion region behind protrusion 2 (light green). (F) triangular brace (magenta) (Flaus et al. 2006).

The DNA predominantly binds to domain 1 and comparison between the structure of the SF1/SF2 helicases and Rad54 show that Ia and TxGx may be involved in DNA binding. Crystal structure studies of SSO1653 show that DNA binds predominantly to subdomain 1A by recognition of the two phosphate chains along the minor groove. The Ia and II region were found to be involved in DNA binding. Arg547, a highly conserve residue of the SWI/SNF in the Ic Motif, was found to interact with the minor groove of the DNA. In addition, domain 2 was
also found to bind to the DNA-phosphate backbone through Arg728 and Lys781. However, this binding is much weaker as compared to the interaction with the domain 1. The DNA binding constant of catalytic domain of SsoRad54 (SsoRad54cd) as a whole is 0.1 +/- 0.02 \mu M for dsDNA; domain 1 by itself has a binding activity of 0.22 +/- 0.1 \mu M, which is 2-fold less than that for the full length protein. Domain 2 does not show any DNA binding activity at all. ATP or non-hydrolysable ATP analogs does not affect DNA binding of SsoRad54cd (Durr et al. 2005). Comparative studies of SsoRad54 and Rad 54 with DNA helicase PcrA and RNA helicase NS3 shows that RecA-like fold and location of Motif I-III is preserved between the enzymes (Thoma et al. 2005). PcrA and NS3 have single-strand DNA (ssDNA) binding domain that firmly grabs the base of the single strand. SsoRad54cd lacks the ssDNA binding domain but can bind to 5'-3' DNA through domain 1. Similarly, Rad54 can bind to DNA through domain 1. As SsoRad54cd lacks ssDNA binding domain, therefore, it does not have ssDNA-stimulated ATP hydrolysis activity. This may be the reason that it cannot translocate ssDNA in ATP hydrolysis dependent manner. Crystallographic studies show that the DNA is fully base-paired along the entire side of SsoRas54cd and there is no upstream helix destabilizing region. This may explain why SWI/SNF proteins do not have helicase activity.

SMARCAL1

As described earlier, SMARCAL1 protein belongs to the Snf2 family of proteins, although it is distantly related owing to the absence of conserved sequence hallmarks of the Snf2 family (Flaus et al. 2006). The protein possesses all the helicase motifs and shows DNA-dependent ATPase activity, but does not unwind DNA. The SMARCAL1 SWI/SNF subfamily is most closely related to the prokaryotic HepA/RapA SWI/SNF subfamily (Coleman et al. 2000; Flaus et al. 2006). As mentioned earlier, mutations in the conserved helicase motifs in
SMARCAL1 protein leads to an autosomal recessive disease called Schimke immuno-osseous dysplasia (SIOD) (Boerkoel et al. 2002). Mutation in this gene results in pleiotropic disorder characterized by spondyloepiphyseal dysplasia, renal dysfunction, and T-cell immunodeficiency (Boerkoel et al. 2002; Elizondo et al. 2009). The involvement of SMARCAL1 in SIOD was determined using genome-wide linkage mapping and a positional candidate approach (Boerkoel et al. 2002). Most of the missense mutations that are associated with the disease have been found to occur in the highly conserved SNF2 domain. The severity of the disease depends on the type of mutations in the gene of SMARCAL1 (Boerkoel et al. 2002; Elizondo et al. 2009). For example, R561C missense mutation causes a milder type of SIOD. Mutations in the conserved arginine residues are found to be more common in SIOD (Boerkoel et al. 2002).

Previous studies have shown that the bovine homolog of SMARCAL1 known as DNA-dependent ATPase A can hydrolyze ATP in presence of DNA double-strand to single-strand transition region, for example stem-loop DNA (Muthuswami et al. 2000). SMARCAL1 has also been shown to interact with the chromatin in vivo (Elizondo et al. 2009). Recently, human SMARCAL1 has been found to interact with single-stranded DNA binding RPA protein at stalled replication fork and double-stranded break sites (Postow et al. 2009; Ciccia et al. 2009; Bansbach et al. 2009). Furthermore, SMARCAL1 has been found to interact with RPA through its conserved N-terminal domain and these domains are conserved in RPA-interacting DNA repair proteins (Yusufzai et al. 2009; Ciccia et al. 2009; Postow et al. 2009). RPA binding has been found essential for function of SMARCAL1 in vivo as it has been shown to recruit SMARCAL1 to the site of stalled replication fork and DSB site (Bansbach et al. 2009; Yusufzai et al. 2009; Ciccia et al. 2009). However, the recruitment of SMARCAL1 to the site of DSB happens in absence of its ATPase activity as it was shown that mutant with no activity could also
localize to the site of damage (Ciccia et al. 2009). SMARCAL1 have also been reported to be an ‘annealing helicase’, in which it was shown that it can re-anneal the single stranded plasmid held apart by RPA (Yusufzai and Kadonaga, 2008). Further, SMARCAL1 has been shown to be phosphorylated by Ataxia-telangiectasia mutated (ATM), ATM and Rad3-related (ATR), and DNA-dependent protein kinase (DNA-PK) in response to replication stress and DSB (Postow et al. 2009; Bansbach et al. 2009). Recently, knockdown studies of SMARCAL1 in D. rerio has shown that deficiency of the SMARCAL1 leads to multi-system developmental defects, including growth retardation, craniofacial abnormality, reduced thymic development, and defects in both primitive and definitive hematopoiesis and angiogenesis (Huang et al. 2010). Furthermore, SMARCAL1 has also found to be involved in cell cycle progression, even though the stage at which the protein is involved remains unclear as studies have shown involvement both in S phase (Ciccia et al. 2009), and in G0/G1 phase (Huang et al. 2010). Downregulation of SMARCAL1 was also reported to result in apoptosis (Huang et al. 2010). SMARCAL1 is also a target of E2F6, a transcription regulator that is involved in cell cycle progression indicating it might play a role in cell cycle. In our lab, SMARCAL1 has been found to interact with β-actin and β-tubulin. Moreover, SMARCAL1 co-localized with β-tubulin during spindle fiber formation and knockdown studies of SMARCAL1 caused aberration in spindle formation, suggesting a role in mitosis. (Reshma Kumari, Ph.D.thesis submitted, 2010).

**DNA-dependent ATPase A**

DNA-dependent ATPase A, the bovine homologue of human SMARCAL1, was originally isolated from the calf thymus tissue (Hockensmith et al. 1986). The 105-kDa ATPase A protein undergoes proteolysis yielding two polypeptides with molecular masses of 68- and 83-
kDa (Mesner et al. 1993). All three polypeptides exhibit ATPase activity only in the presence of DNA (Hockensmith et al. 1986; Mesner et al. 1993). Polypeptide sequencing of the 105-kDa protein and its immunoaffinity purified 82-kDa revealed that the 82-kDa lacks the N-terminal domain (Muthuswami et al. 2000). However, the 82-kDa protein retains all the seven helicase motifs and thus, shows DNA-dependent ATPase activity. The cDNA sequence encoding the 82-kDa (727 amino acid) polypeptide has been cloned and termed Active DNA-dependent ATPase A Domain (ADAAD). The ATPase activity of ADAAD is maximal in the presence of DNA molecules possessing a double-strand to single-strand transition region but no activity is detected in the presence of single-stranded or double-stranded DNA suggesting that the structure of the DNA molecule is important for the activity. Further, ADAAD is able to hydrolyze ATP in the presence of various DNA effector like 3′-recessed DNA, AT-rich duplex, 4 base pair match duplex, and 10 base pair match duplex A (Muthuswami et al. 2000).

**Hypothesis**

Based on the above results, we hypothesized that the effector driven catalytic activity of ADAAD is due to the conformational changes the protein undergoes in the presence of DNA and ATP.

In this study, I have tried to explore the mechanism of DNA-dependent ATP hydrolysis by the SWI/SNF proteins using ADAAD as the experimental model and stem-loop DNA as the effector. I have also attempted to delineate the role of motor domain in DNA and ATP binding. Finally, I have proposed a possible mechanism of ATP hydrolysis by the SWI/SNF proteins.