Functional Analysis of the SWI/SNF family protein LSH

Kevin Brian Myant

Thesis presented for the degree of Doctor of Philosophy

University of Edinburgh

2008

Acknowledgements

Firstly, I would like to thank Irina for taking me on as a PhD student and for mentoring me throughout my time here. In particular, for providing advice and guidance during experiments and for reading the first draft of this thesis. I would also like to thank all the members of the Stancheva lab (Jose, Eleni, Matt, Kat, Cara, Mel, Xinsheng, Manos, Thomas, Simone and Vikram) for help with experiments, helpful suggestions and ideas throughout my time here.

In addition I would like to extend a special thanks to Tom Owen-Hughes for allowing me to carry out some experiments in his laboratory in Dundee. I also thank Helder Ferriera and Chris Stockdale in Tom's lab for providing me with histone octamers, nucleosomes and help and advice on analysing LSH biochemically. I would also like to thank Adrian Bird and members of his lab for the use of equipment and much useful input to this project. Additionally, I would like to thank Juri Rappsilber and Flavia Alves for analysing LSH by Mass Spectrometry.

I would also like to thank my parents, Maureen and Martin for supporting me to this point and for providing encouragement when needed. Also, I thank my brother Peter, for providing useful relief from work, usually in the pub, and my sister Katherine for leading the way.

Finally, I would like to thank Alana for supporting me throughout my PhD experience and sharing it with me. You were always great fun and gave me support when I needed it.

This work was supported by a Cancer Research UK Studentship, whom I thank for providing financial support.

Table of Contents

List of Tables and Figures

Abstract

Methylation of the carbon-5 position of cytosine is the major covalent modification of DNA found in multicellular organisms and is involved in a diverse range of biological processes in mammals, plants and fungi. In mammals, this modification plays a number of diverse roles in development and cellular differentiation. These include transcriptional gene silencing, chromosomal integrity, repression of transposable elements, parental imprinting and X chromosome inactivation (Bird, 2002). Patterns of DNA methylation are established and maintained by proteins of the DNA methyltransferase (DNMT) family (Bestor, 2000a; Bird, 2002). Interestingly, genetic studies in *A. thaliana* and mouse have identified a protein from a functionally distinct family that is also required for high levels of DNA methylation. This protein, termed DDM1 in plants (deficient in DNA methylation 1) or Lsh in mouse (lymphoid specific helicase), belongs to the SNF2 chromatin remodelling family of ATPases (Dennis et al., 2001; Jeddeloh et al., 1999; Vongs et al., 1993). SNF2 enzymes have a variety of chromatin related functions that are based upon the ability to disrupt DNA:histone contacts in an ATP dependant manner (Narlikar et al., 2002). Despite the genetic characterisation of Lsh, little is known about its molecular function. Cytological studies and experiments with ES cells have shown that although Lsh co localises with Dnmt1 at replication foci during late S phase, it is not required for maintenance methylation of satellite DNA or on replicating episomal plasmids. Lsh has also been shown to be required for *de novo* methylation and to co-immunoprecipitate with Dnmt3a and Dnmt3b (Zhu et al., 2006). Taken together these studies indicate that the primary role of Lsh is in *de novo* methylation and it is dispensable for maintaining methylation during DNA replication.

As a member of the SNF2 family of ATPases an attractive hypothesis is that Lsh may disrupt chromatin in a manner that makes it more accessible to DNMT enzymes. Alternatively, LSH may act as a recruitment factor for DNMTs or increase their

catalytic activity in some other way. The aims of this study were to characterise the molecular function of LSH and attempt to relate this to its role in DNA methylation *in vivo*. To address these aims the work described in this thesis uses a variety of approaches to address two key questions. (1) Is LSH an active SNF2 ATPase? (2) How does LSH interact with, and modulate, the DNA methylation machinery?

To determine if LSH is an active SNF2 ATPase, recombinant LSH was purified from insect cells and biochemically characterised. These experiments revealed that LSH can hydrolyse ATP and its activity is stimulated by DNA. However, the rate of ATP hydrolysis is low and LSH does not exhibit chromatin remodelling activity. Thus, either LSH is a relatively weak SNF2 ATPase that cannot remodel chromatin or recombinant LSH may not be fully active. To identify proteins that interact with LSH, biophysical analysis of the native protein was performed. This indicated that the majority of LSH was present as a free monomeric peptide *in vivo*. However, I was able to demonstrate that LSH is an HDAC transcriptional repressor *in vivo*. Also, a weak, transient or low abundance complex including LSH, DNMT3B, DNMT1, HDAC1 and HDAC2 was identified. Thus, how LSH interacts with the DNA methylation machinery was demonstrated. Finally, I attempted to investigate how LSH modulates the activity of DNMTs *in vitro*. These experiments did not identify a role for LSH in stimulating DNMTs *in vitro*. These studies shed new light on the role of LSH in DNA methylation.

Chapter one – Introduction

1.1 Overview of Epigenetics

1.1.1 Epigenetics

The completion of the human genome sequence has given us a unique opportunity to understand the complexities of our genetic make up (Lander et al., 2001). The information derived from this DNA sequence however, only partly explains the complex biology of multicellular organisms. We consist of many diverse cell types, each with a specific function and role but each containing essentially the same DNA code. The defining factors in specifying particular cell types are the protein composition of the cell, which in turn is defined by the cells gene expression pattern. Transcription factors (TFs) help define gene expression patterns by recognizing specific DNA sequences in the promoters of target genes. Interaction of these factors with gene promoters leads to a cascade of events culminating in the expression or silencing of the gene. However, TFs alone are insufficient to establish and maintain this phenomenon as the transcription potential of the genome becomes stably restricted during development. Acting alongside TFs are a host of DNA:protein interactions revolving around covalent modifications to DNA and chromatin. These include methylation of DNA itself and numerous histone modifications. These modifications help control gene expression by regulating the accessibility of chromatin to transcription factors. Together these modifications are termed 'epigenetic', literally meaning 'in addition to genetics'. Epigenetic modifications change the final outcome of a locus without altering the underlying DNA sequence. Thus without altering the sequence of DNA the same gene can be either expressed or silenced in different tissues. Ideally, epigenetic modifications should also be heritable, but so far only DNA methylation shows stable inheritance through multiple cell divisions (Bird, 2002). Most covalent modifications that alter the expression of genes are considered epigenetic, and mechanisms of their inheritance may be discovered in the future.

1.1.2 Mechanisms of action

The human cell must store >2m of DNA in a nucleus with a diameter of only 10μm. To achieve this DNA is tightly packaged into a DNA:protein structure known as chromatin. The basic repeating unit of chromatin is the nucleosome, consisting of 146bp of DNA wrapped around an octamer of core histones; a tetramer of histone H3 and H4 and 2 dimers of histones H2A and H2B (Luger et al., 1997; Richmond and Davey, 2003a; Schalch et al., 2005). The N-terminal tails of H3 and H4 protrude from the nucleosome core where they are the targets of a multitude of post-transcriptional modifications including acetylation, methylation, phosphorylation and ubiquitylation (Kouzarides, 2007). These modifications can alter the interactions of the tails with DNA and affect chromatin folding (Tremethick, 2007) or recruit effector proteins to further modify chromatin function. Together, it is believed that combinations of modifications of the N-terminal tails of H3 and H4 constitute a 'histone code'. This 'code' is thought to regulate the accessibility of DNA to factors required for gene expression, DNA replication, recombination and repair.

Eukaryotic DNA can be divided into 2 different classes broadly based on confirmation and accessibility; euchromatin and heterochromatin. These compartments are defined cytologically (Heitz, 1928) and on the basis of DNA methylation content, histone modification patterns and replication timing (Gilbert, 2002). Importantly, the location of a gene in a particular chromatin class defines its transcriptional potential, with euchromatic genes being potentially transcribed, whereas heterochromatic genes are silenced.

Euchromatin generally consists of gene rich, GC rich DNA that is less condensed than heterochromatin and can be transcriptionally active. It stains light-coloured bands when stained with Giemsa in metaphase spreads and replicates early in S-phase. Euchromatin is enriched for 'active' histone marks including acetylation of H3 at lysines 9 and 14 (H3K9Ac/H3K14Ac) and H4 at lysine 12 (H4K12Ac). Methylation of H3 at lysine 4 (H3K4me2/me3) and lysine 79 (H3K79me2) and a lack of promoter DNA methylation are prevalent in euchromatin. The occurrence of these histone modifications and the open chromatin configuration of euchromatin

beget DNA accessible to transcription machinery. Thus genes found in euchromatin are liable to be transcribed.

Heterochromatin is generally gene poor, AT rich DNA that is highly compacted, inaccessible and prohibitive to gene expression. Heterochromatin can be divided into 2 subclasses; constitutive and facultative. Constitutive heterochromatin consists of the gene poor, repetitive portion of the genome which is found at centromeres and telomeres. Constitutive heterochromatin stains dark-coloured bands with Giemsa and replicates late in S-phase. Facultative heterochromatin refers to certain developmentally controlled loci that are not consistent within all cell types in an organism such as the inactive X chromosome and the mouse β-globin locus (Trojer and Reinberg, 2007). Both classes of heterochromatin consist of histones that are hypoacetylated, enriched in H3K9me3 and H3K27me3 and DNA methylation (Trojer and Reinberg, 2007). An interesting feature of heterochromatin is its ability to propagate and spread, thus influencing the expression of nearby genes. The spreading of facultative heterochromatin from a nucleation site on the female inactive X chromosome is an example of the importance of this spreading mechanism (Grewal and Jia, 2007).

Interestingly, although certain histone modifications have historically been thought to correlate with specific classes of chromatin recent work has shown this not to be the full picture. For example H3K9me3 has been found in the body of transcribed euchromatic genes and RNA transcripts have been found associated with heterochromatic regions (Grewal and Jia, 2007; Vakoc et al., 2005). These examples outline the addition complexities of so called epigenetic phenomena and tell us there is much still to understand about them.

1.2 DNA Methylation

Methylation at the carbon-5 position of cytosine (m^5C) is the major covalent modification of DNA in multicellular organisms and is involved in a diverse range of biological processes in mammals, plants and fungi (Figure 1.1A). In mammals m^5C is predominantly found on both strands of DNA in the context of symmetrically methylated cytosine-guanine (CpG) dinucleotides (Bird, 2002). The methyl group of m^5C in the double stranded helix has been shown to protrude into the major groove of DNA and influences DNA:protein interactions without affecting base pairing (Mayer-Jung et al., 1997). Importantly, the symmetry of DNA methylation permits a simple copying mechanism where DNA methylation patterns can be transmitted during DNA replication. The levels of CpG methylation varies widely within eukaryotes from none in yeast to ~60-80% in some mammals. Also, CpG methylation levels differ between different cell types in multicellular organisms. For example, in undifferentiated mammalian ES cells about 60% of CpGs are methylated whereas in differentiated lung cells ~80% of CpGs are methylated (Ramsahoye et al., 2000b).

DNA methylation plays a number of diverse roles in mammalian development and cellular differentiation. It is required for transcriptional gene silencing, chromosomal integrity, repression of transposable elements, parental imprinting and X chromosome inactivation (Bird, 2002). Highlighting the importance of DNA methylation to normal cellular function is the fact that removal of DNA methylation results in embryonic lethality in mammals (Li et al., 1992). DNA methylation is clearly important for normal development during cell differentiation but it also plays a role in establishing abnormal gene expression patterns observed in cancerous cells (Jones and Baylin, 2002).

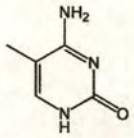
1.2.1 Functions of DNA methylation

1.2.1.1 Chemistry of m^5C

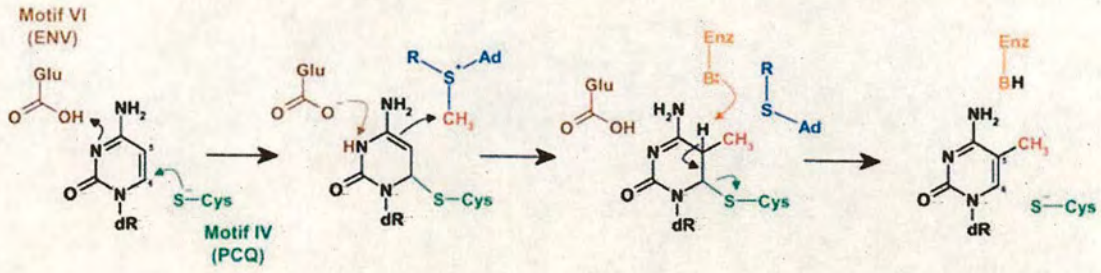
A class of enzymes known as DNA cytosine methyltransferases (DNMTs) catalyse the DNA methylation reaction. The catalytic motifs of DNMT enzymes are conserved between prokaryotes and eukaryotes and use *S*-adenosyl-L-methionine (SAM) as a methyl donor. In prokaryotes DNA methylation is involved in DNA-replication, post-replicative repair mechanisms, and restriction/modification systems. The role of DNA methylation in eukaryotes differs somewhat but the catalytic reaction and the DNMT enzymatic motifs are conserved (Kumar et al., 1994). DNMTs contain 10 characteristic motifs, of which 6 (I, IV, VI, VII, IX and X) are highly conserved (Kumar et al., 1994). Motifs I and X bind SAM, motif IV contains the PCQ motif that provides the thiolate at the active site, motif VI contains the glutamyl residue that stabilises the DNA:DNMT interaction, motif IX is involved in substrate recognition and motif VII has an as yet unknown function (Bestor, 2000a).

The DNA methylation reaction was first elucidated in the prokaryotic DNMT *M.HhaI* (Klimasauskas et al., 1994; Wu and Santi, 1985, 1987). The C-5 position of cytosine is rather unreactive and its methylation is considered improbable in physiological conditions. The DNMT enzyme must overcome the inherent inertness of the substrate and force it to react with SAM. To achieve this, the DNMT enzyme utilises a nucleophilic attack on the C-6 of the target cytosine. This attack is performed by the thiol group of the cysteine residue in the conserved PCQ motif (motif IV). This forms a covalent bond between the DNMT and C-6 and pushes electrons to the C-5 position. This activates C-5, causing it to attack the SAM and receive the methyl group (Bestor, 2000a). As this occurs the DNA:DNMT reaction intermediate is further stabilised by transient protonation of N-3 by the glutamic acid of the ENV motif (motif VI) (Hermann et al., 2004). Following completion of the methyl transfer the reaction cycle is completed by deprotonation of C-5 and release of the DNMT by β-elimination (Bestor, 2000a) (Figure 1.1B). A further hindrance to the DNA methylation reaction is the relative inaccessibility of the cytosine substrate that is buried within the double helix. Co-crystallisation studies of the *M.HhaI* DNMT in complex with DNA revealed the evolution of an elegant solution to this

A



B



C

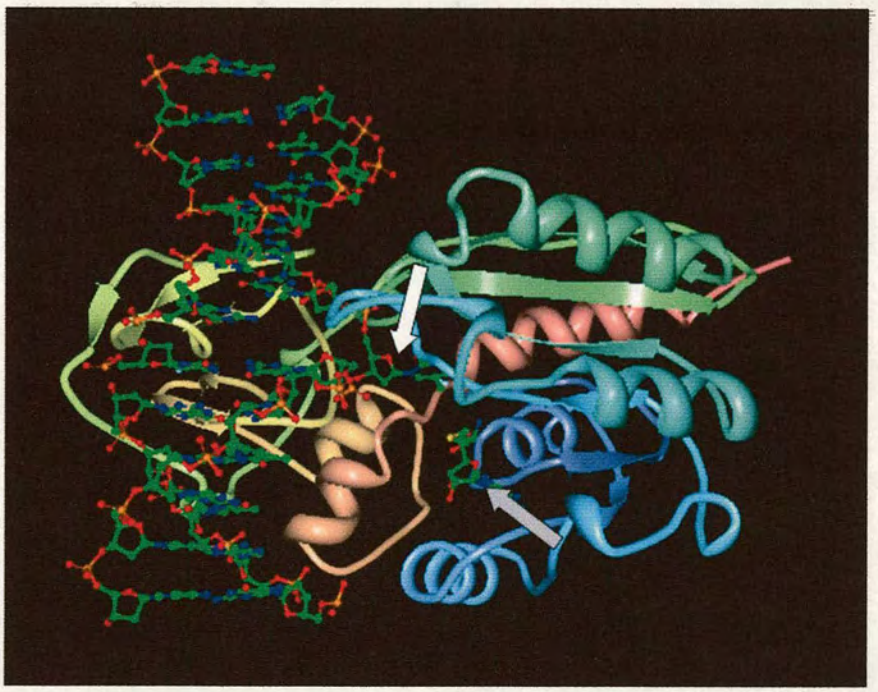


Figure 1.1. The chemistry of DNA methylation.
A. Methylation of cytosine occurs at the carbon at the 5' position in the DNA base.
B. The chemistry of the DNA methylation reaction. Figure adapted from (Klimasauskas et al., 1994)
C. Structure of the bacterial DNA methyltransferase *M.*Hha*I* in complex with DNA. The 'base flipping' of the cytosine into the active site of the methyltransferase is indicated (white arrow). The methyl-donor *S*-adenosyl-L-methionine is indicated (grey arrow). Figure adapted from (Bestor, 2000b)

problem (Klimasauskas et al., 1994; Wu and Santi, 1985, 1987). The DNMT was shown to pull out the target base in a process called 'base flipping'. This process involves the rotation of the sugar-phosphate bonds to project the target base into the catalytic pocket of the enzyme (Figure 1.1C). All other known DNMTs also utilise base flipping as do a number of other enzymes that require access to DNA aromatic rings such as DNA repair enzymes (Cheng and Roberts, 2001).

1.2.1.2 Physiological roles of DNA methylation

1.2.1.2.1 X-inactivation

Mammalian sex determination is based on the heteromorphic sex chromosomes, X and Y (Ross et al., 2005; Vallender et al., 2005). This heteromorphic pair is thought to have evolved from a homologous pair of autosomes ~300mya. In this time, the chromosomes have diverged with the Y becoming smaller and losing most of its ancestral genes. This divergence has resulted in a potential imbalance in expression of X-linked genes between males and females. The mechanism to correct this imbalance is known as X inactivation and involves silencing of one of the X chromosomes in XX females early in development, so that gene expression is similar to that in XY males. The inactive X can be distinguished from its active counterpart on the basis of transcriptional silencing of most of the genes, heterochromatic condensation, accumulation of facultative heterochromatin marks, late replication in S-phase and hypermethylation of CpG islands (Brockdorff and Duthie, 1998). The process of X inactivation occurs in early development and is mediated through a locus known as the X-inactivation centre (Brockdorff and Duthie, 1998). A key component of the XIC is the *Xist* gene, which plays a key role in initiation of X chromosome gene silencing. The *Xist* gene encodes a 17.3kb transcript that is not translated into protein but rather coats the inactive X chromosome (Borsani et al., 1991; Brockdorff et al., 1991; Brockdorff et al., 1992; Brown et al., 1991). The accumulation of Xist RNA on the inactive X is rapidly followed by transcriptional repression, accumulation of silent chromatin marks such as lysine hypoacetylation, H3K27 trimethylation and H3K9 trimethylation. Following transcriptional repression and the accumulation of theses marks on the inactive X CpG methylation of CGIs occurs. Recent data have suggested that most CGIs on the inactive X are methylated

compared to its active counterpart (Weber et al., 2007). A number of studies have outlined the important role DNA methylation plays in the stability and maintenance of gene silencing on the inactive X. For example, early studies using the DNMT inhibitor 5-Aza-C showed de-repression of several genes on the inactive X indicative of an important role for CpG methylation in maintaining transcriptional silencing (Lock et al., 1986; Wolf et al., 1984). More recently, conditional loss of Dnmt1 has been shown to also lead to reactivation of genes on the inactive X (Csankovszki et al., 2001). Similarly, female patients with ICF, caused by a mutation in the DNMT3B genes, have an increase in the number of genes on the inactive X that are not silenced. This may be due to a failure to maintain CGI methylation on the inactive X, despite X inactivation occurring normally (Hansen et al., 2000). The functional relevance of this finding is not clear, however as ICF males and females show similar phenotypes which would not be expected if the disease affected maintenance of X inactivation (Ehrlich et al., 2008).

1.2.1.2.2 DNA methylation is a transcriptionally repressive mark

DNA methylation in mammals generally elicits a transcription response either by one of two mechanisms, directly inhibiting TF binding or indirectly recruiting co repressor complexes (Bird, 2002). In the first instance the DNA binding of sequence specific transcription factors can be inhibited by CpG methylation. One such example is c-Myc, of which the ability to bind its cognate sequence is specifically inhibited by CpG methylation (Prendergast et al., 1991; Prendergast and Ziff, 1991).

The second mechanism of repression by CpG methylation involves recruitment of proteins to methyl-CpG. The proteins that recognise methyl-CpG fall into 1 of 2 families, proteins of the Methyl-CpG-binding domain (MBD) family and Kaiso-like proteins (Bird, 2002). The MBD family consists of 5 proteins, MeCP2 and MBD1-4. Three of these, MeCP2, MBD1 and MBD2 have been shown to have a role in methylation dependent transcriptional repression. The MBD proteins act through the Methyl-CpG dependant recruitment of chromatin modifying co-repressor complexes that alter chromatin into a transcriptionally silent state (Nan et al., 1998; Sarraf and Stancheva, 2004; Tong et al., 1998). The Kaiso-like proteins are unrelated to MBDs

and bind methylated DNA through 3 zinc finger domains (Filion et al., 2006; Prokhortchouk et al., 2001). In the case of Kaiso this leads to transcriptional silencing via an interaction with the repressive N-CoR chromatin remodelling complex (Yoon et al., 2003).

1.2.1.2.3 Patterns of DNA methylation

To fully understand the physiological role of DNA methylation we must not only consider the chemical reaction that marks CpGs with a methyl group but also the distribution of methyl-CpG throughout the genome. As mentioned above the levels of CpG methylation vary in different organisms and also between the different cell types of an organism. In animals, the lowest extreme is the nematode worm *Caenorhabditis elegans* that lacks detectable m^5C and has no conventional DNMT enzyme. The insect *Drosophila melanogaster*, also has very low m^5C levels and DNA methylation mainly occurs at CpT dinucleotides rather than CpG (Gowher et al., 2000; Lyko et al., 2000). *Drosophila* also contains a DNMT enzyme (DNMT2) which has been reported to be responsible for DNA methylation at CpT and CpA dinucleotides (Hung et al., 1999; Kunert et al., 2003; Tweedie et al., 1999). Other invertebrate genomes, for example the sea squirt *Ciona intestinalis* have moderately high m^5C levels that are organised in mosaic patterns. Interestingly, m^5C in *C. intestinalis* is targeted to transcription units and may act to suppress spurious transcriptional initiation (Suzuki et al., 2007). Vertebrate genomes have the highest levels of m^5C in the animal kingdom where around 70% of CpGs are methylated. The distribution of CpG methylation in the vertebrate genome does not follow the mosaic pattern observed in invertebrates. In fact CpG density and CpG methylation are distinctly separated into 2 fractions. The first fraction is in bulk genomic DNA, where CpGs are relatively infrequent (~1 per 100bp) and the majority of CpG methylation is found. The other fraction correlates with short stretches of much higher CpG density (~1 per 10bp) where very little CpG methylation is found (Bird et al., 1985; Bird, 1986). These regions of high CpG density are known as CpG islands and frequently co localise with the transcription start sites (TSS) of genes.

1.2.1.2.4 CpG islands

The most striking feature of vertebrate CpG distribution and CpG methylation is the non random distribution of unmethylated CpGs to regions at the 5' end of genes known as CpG islands (CGIs). CGIs are stretches of DNA typically 500 – 1500 bp in length with a G+C content of >50% and a CpG observed over expected ratio of >60% (Gardiner-Garden and Frommer, 1987; Takai and Jones, 2003). There are an estimated 25,000 CGIs in the human genome of which ~76% associate with a transcriptional unit and ~50% correspond to the TSS of a gene (Illingworth et al., 2008). Most CGIs lack DNA methylation and are associated with transcriptionally active genes, including housekeeping and tissue specific genes.

A small but significant proportion of CGIs are normally methylated. These CGIs are found at the differentially methylated regions (DMR) of imprinted genes or on the inactive X chromosome and correlate with transcriptional silencing of the associated gene. Interestingly, recent studies have expanded the numbers of methylated CGIs in normal human tissues and implied that CGI methylation may be more prevalent than previously thought (Illingworth et al., 2008; Weber et al., 2007). In these studies CGI methylation levels varied between the tissues tested perhaps indicating that CGI methylation plays a role in specifying cell fate determination during development (Illingworth et al., 2008). CGI methylation was also more prevalent at 'weak' CGIs with low CpG density, but at these sites did not preclude gene activity (Weber et al., 2007). CGIs have also been reported to acquire methylation in somatic tissues. *De novo* methylation of CGIs has been reported to occur during normal aging, *in vitro* cell culture and most dramatically in cancers (Antequera et al., 1990; Fraga et al., 2005; Issa, 2000). The acquisition of CGI methylation in these cases is coincident with global genomic hypomethylation and associated genomic instability (Chen et al., 1998; Feinberg, 1988). The precise mechanism of aberrant DNA methylation targeting in cell culture and cancer and how it differs from naturally occurring DNA methylation during aging is currently unknown.

number of enzymes have been discovered that fit the roles described above and provide a functional methylation system. DNMT1, DNMT2, DNMT3A, DNMT3B and DNMT3L have been identified as proteins containing a conserved methyltransferase domain (Figure 1.2) (Hermann et al., 2004). DNMT1 is the mammalian maintenance methyltransferase and DNMT3A and DNMT3B are the *de novo* methyltransferases. DNMT2 and DNMT3L do not appear to function as DNA methyltransferases but may play other roles (Bestor, 2000b; Hermann et al., 2004).

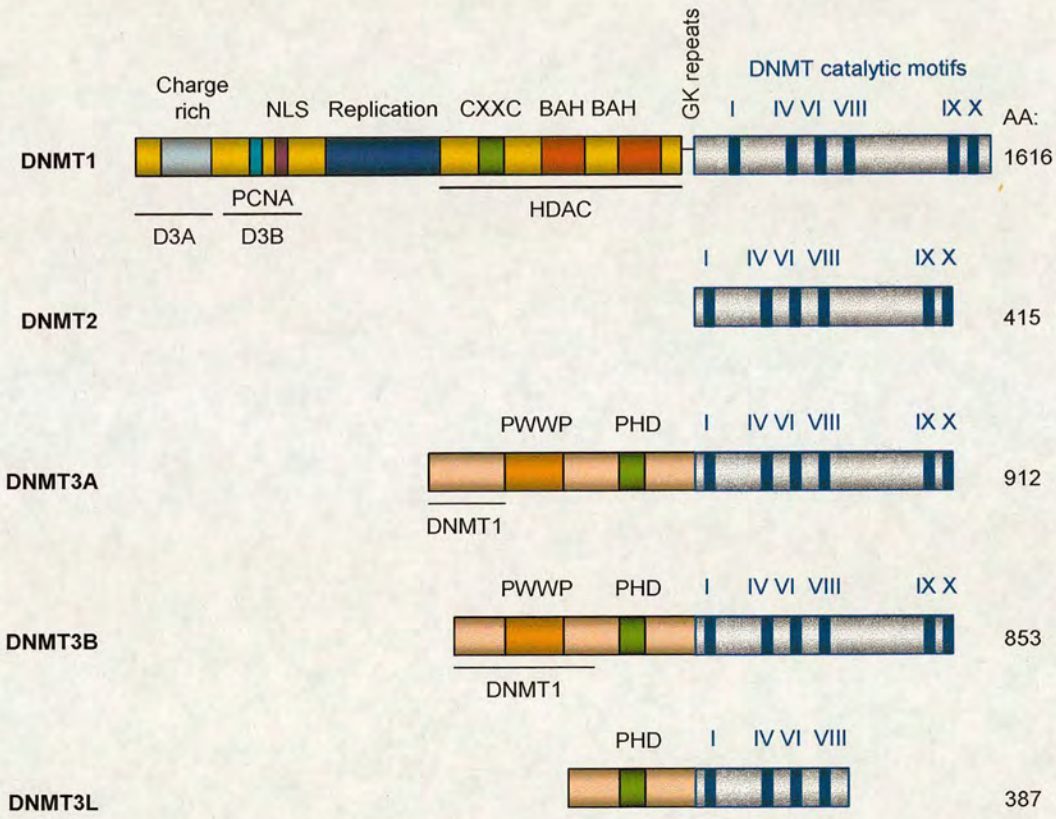


Figure 1.2. The mammalian DNA methyltransferase family.
Members of the mammalian DNMT family contain an N-terminal regulatory region and a C-terminal catalytic region. The catalytic region shows strong homology to bacterial DNMTs, in particular the catalytic motifs indicated by roman numerals. Other key domains in the different proteins and labelled above the proteins and a number of key interaction sites labelled below them. The number of amino acids contained in the human proteins is listed on the right.

All the mammalian methyltransferases share a common catalytic domain at their C-terminus and a regulatory domain at their N-terminus. This catalytic domain is similar to the prokaryotic methyltransferase domain and contains the 10 methyltransferase motifs discussed earlier. The N-terminal regions of the methyltransferases, although structurally unconserved all contain regulatory regions involved in protein:protein interactions and targeting of the enzymes.

1.2.2.1 DNMT1

1.2.2.1.1 DNMT1 is a maintenance DNA methyltransferase

DNMT1 was the first mammalian DNMT to be discovered and its major role is to fully methylate the hemimethylated CpGs produced during DNA replication (Bestor et al., 1988; Bestor, 2000b). DNMT1 has been shown to preferentially bind and methylate hemi-methylated DNA in a processive manner *in vitro* (Bacolla et al., 1999; Pradhan et al., 1999; Vilkaitis et al., 2005). DNMT1 showed a 5-30 fold preference for this substrate over a completely unmethylated one but also showed significant *de novo* activity indicating a potential role for DNMT1 in this process (Pradhan et al., 1999). Further *in vivo* evidence for a role of DNMT1 in maintenance methylation was the discovery of its association with the replication fork and interaction with PCNA (proliferating cell nuclear antigen), a major component of the replication machinery (Chuang et al., 1997; Leonhardt et al., 1992; Liu et al., 1998). This interaction is not solely responsible for it role in maintenance methylation however. Deletion studies of DNMT1 have shown that removal of the PCNA interaction region does not prevent DNMT1 actively maintaining methylation (Schermelleh et al., 2007; Spada et al., 2007). Recently, DNMT1 has been shown to interact with a protein that specifically binds hemi-methylated DNA termed Uhrf1 or Np95 (Bostick et al., 2007; Sharif et al., 2007). This interaction provides both further evidence for the role of DNMT1 in maintenance methylation and also an additional mechanism for DNMT1 targeting.

Although the major proposed role of DNMT1 is in maintenance methylation, evidence exists for a role of DNMT1 in the *de novo* process. As mentioned above

DNMT1 shows *de novo* methyltransferase activity *in vitro* (Pradhan et al., 1999). In fact, when compared, the *de novo* activity of DNMT1 appears to be higher than the activity of the bona fide *de novo* enzymes, DNMT3A and DNMT3B (Okano et al., 1998a). Further evidence for a potential role for DNMT1 in *de novo* methylation comes from a study involving a cancer cell line genetically deficient for DNMT3A and DNMT3B (Jair et al., 2006). In this study, cell extract from cells lacking the *de novo* methyltransferases could efficiently methylate CGI DNA *in vitro*. It was also shown that DNMT1 could *de novo* methylate this DNA in *Drosophila* cell lines thus implying a role for DNMT1 in *de novo* DNA methylation *in vivo* (Jair et al., 2006).

1.2.2.1.2 Role of DNMT1 in transcriptional repression

The N-terminal regulatory region of DNMT1 contains a number of interesting domains involved in regulating its function and protein:protein interactions. This region contains a nuclear localisation signal (NLS), a charge rich domain that interacts with the DMAP1 transcriptional repressor, a PCNA interacting domain, an HDAC interacting region, DNMT3A and DNMT3B interacting domains, a domain required for targeting to replication foci and a cysteine-rich zinc finger (CXXC) implicated in DNA binding (Figure 1.2) (Hermann et al., 2004). Of particular interest is the role of DNMT1 in transcriptional repression. DNMT interacts with a number of chromatin modifying enzymes that act to repress transcription, including the histone deacetylases 1 and 2 (HDAC1 and 2), the H3K9 methyltransferase SUV39H1, heterochromatin protein 1, the *de novo* methyltransferase enzymes DNMT3A and 3B and the chromatin remodelling enzyme SNF2H (Fuks et al., 2000; Fuks et al., 2003; Kim et al., 2002; Rountree et al., 2000). DNMT1 has also been shown to interact with the E2F-Rb complex and repress E2F responsive genes (Robertson et al., 2000). Interestingly, in the above studies, the catalytic domain of DNMT1 was dispensable for transcriptional gene silencing indicating that the roles of DNMT1 as a transcriptional repressor and its role in maintenance methylation are distinct.

1.2.2.1.3 Genetic studies of DNMT1

Four DNMT1 knockout models have been described; Dnmt1^{n}, Dnmt1^{s}, Dnmt1^{c} and Dnmt1^{ps} (Lei et al., 1996; Li et al., 1993; Li et al., 1992; Takebayashi et al., 2007). The first genetic study on Dnmt1 involved the creation of the Dnmt1^{n} mutant by disruption of the N-terminus of Dnmt1. This mutant had very low activity and homozygous ES cells had methyl-CpG levels decreased to around 30% of wt. Consistent with an essential role of DNA methylation in development, homozygous mice died around E11 (Li et al., 1992). A similar mutant termed Dnmt1^{s} disrupts the region involved in targeting Dnmt1 to replication foci. This mutant is more severe than Dnmt1^{n}, with homozygous mice dying earlier in development with less DNA methylation. Homozygous mice also show biallelic expression of imprinted genes and transcription from IAP repeats (Li et al., 1993). Two Dnmt1 mutants specific to the catalytic domain of Dnmt1 have also been generated to test the requirement of DNA methylation for embryonic development. Dnmt1^{c} contains a disruption of the highly conserved methyltransferase motifs IV and VI. Homozygous mice have a very severe phenotype, with a ~70% reduction of CpG methylation levels compared to wild type and embryonic lethality by E8.5 (Lei et al., 1996; Ramsahoye et al., 2000b). Homozygous embryos have subsequently been shown to exhibit imprinting and X inactivation maintenance defects due to hypomethylation of these sites (Sado et al., 2000). I earlier discussed studies that separated the role of Dnmt1 as a transcriptional repressor and its role in methylating DNA. This raises the question, is lethality in these mutants is due to the loss of DNA methylation or the loss of a global transcriptional repressor? An interesting recent study has conclusively answered this question (Takebayashi et al., 2007). The Dnmt1^{ps} mutant contains a single amino acid point mutation at C1229 in the PCQ motif. Mutation of this cysteine residue has been previously shown to abolish methyltransferase activity of Dnmt1 *in vitro* (Wyszynski et al., 1993). Importantly, mice homozygous for this mutant showed a similar phenotype to Dnmt1^{c} homozygotes with near complete loss of DNA methylation and developmental arrest around E8.5 (Takebayashi et al., 2007). These studies underline the critical importance of Dnmt1 and CpG methylation to mouse development.

1.2.2.2 Dnmt2

The *dnmt2* gene is conserved in eukaryotes that have a functional DNA methylation system, but also those that lack one such as *Schizzosaccharomyces pombe*. Dnmt2 is distinct from the maintenance and *de novo* methyltransferases in that it contains no N-terminal regulatory region, just the catalytic domain (Figure 1.2). Dnmt2 shows only very weak methyltransferase activity *in vitro* with a substrate preference for CpG (Hermann et al., 2003). The structure of Dnmt2 has been solved in complex with *S*-adenosyl-L-homocysteine and it shows structural conservation of the methyltransferase motifs indicative of a potentially active enzyme (Dong et al., 2001). ES cells null for Dnmt2 are viable and maintain normal levels of DNA methylation (Okano et al., 1998b). Two studies have outlined potential roles for Dnmt2. The first study analysed the function of Dnmt2 in *Drosophila* and showed that siRNA knockdown led to complete loss of DNA methylation as measured by an α-m^5C antibody (Kunert et al., 2003). They further found over expression of Dnmt2 led to an increase in methylation at CpA and CpT dinucleotides, suggesting a role for Dnmt2 as a CpA/T methyltransferase *in vivo*. Another, recent study looked at the role of Dnmt2 in mouse. Using a combined genetic approach they identified $tRNA^{Asp}$ as a *in vivo* target of Dnmt2 (Goll et al., 2006). As this study used a Dnmt2 null mouse that showed no obvious phenotype the functional significance of this modification is currently unknown.

1.2.2.3 *De novo* DNA methyltransferases

The Dnmt3 enzymes were identified in mouse and human of EST database searches by their homology to bacterial methyltransferases (Okano et al., 1998a). The 2 proteins Dnmt3a and Dnmt3b are highly homologous to each other and contain an N-terminal regulatory region and a C-terminal catalytic domain (Figure 1.2). Their catalytic domains are only distantly related to that of Dnmt1 and may have been acquired from a different ancestor (Okano et al., 1998a; Xie et al., 1999). The presence of residual methylation and *de novo* methyltransferase activity in $Dnmt1^s$ and $Dnmt1^c$ knockout mouse models implicated these genes as *de novo* methyltransferases. Indeed, when expressed and purified from bacteria or from baculovirus infected insect cells these proteins were found to methylate CpG without

preference for hemimethylated DNA, thus these enzymes were assigned *de novo* methyltransferases (Gowher and Jeltsch, 2001; Okano et al., 1998a; Suetake et al., 2003). Their mechanism of methylation differs however, with Dnmt3a being distributive but Dnmt3b being processive (Gowher and Jeltsch, 2002). An ability to methylate non CpGs was also detected but the significance of this is unknown (Ramsahoye et al., 2000a). Several studies have shown that these enzymes are also able to *de novo* methylate *in vivo*. Firstly, Dnmt3a and Dnmt3b are highly expressed during the early developmental stages and is it at this time that most *de novo* methylation occurs (Okano et al., 1999; Okano et al., 1998a). Secondly, when over expressed in murine ES cells, Dnmt3a and Dnmt3b but not Dnmt1 were able to methylate a stable episomal target (Hsieh, 1999). Thirdly, when expressed in *Drosophila*, Dnmt3a but not Dnmt1 could introduce DNA methylation to the genome, resulting in lethality (Lyko et al., 1999). Finally, disruption of Dnmt3a and Dnmt3b genes in mouse ES cells leads to a gradual loss of DNA methylation that can be rescued by transfection of Dnmt3a or Dnmt3b but not Dnmt1 (Chen et al., 2003; Jackson et al., 2004). This final interesting result also indicates that Dnmt1 alone is not sufficient for maintenance methylation and Dnmt3a and Dnmt3b also play a role. This is supported by a number of studies that have shown maintenance methylation by Dnmt1 to have an error rate of around 4% per cell division (Laird et al., 2004; Riggs et al., 1998; Silva et al., 1993). Together, these results indicate that Dnmt3a and Dnmt3b are *de novo* methyltransferases that also have a role in maintenance methylation.

1.2.2.3.1 Role of Dnmt3a and Dnmt3b in gene repression

The N-terminal regions of Dnmt3a and Dnmt3b both contain 2 highly conserved domains, a PWWP domain and a cysteine rich plant homeobox domain (PHD) (Figure 1.2). The PWWP domain is found in many chromatin binding proteins and is named after a Pro-Trp-Trp-Pro motif. The PWWP motif found in Dnmt3a and Dnmt3b is required for targeting the proteins to chromatin (Chen et al., 2004; Ge et al., 2004). Disruption of these domains leads to a loss of association of the enzymes with heterochromatin and is required for rescuing DNA methylation at these sites in *Dnmt3a/Dnmt3b* KO ES cells (Chen et al., 2004). The role of this domain in

associating the proteins with heterochromatin is not clear but the PWWP domain of Dnmt3b can bind DNA *in vitro* suggesting this may have a role (Chen et al., 2004).

The cysteine-rich PHD zinc finger domain is similar to the PHD-like domain in the chromatin remodelling enzyme ATRX. This domain is characterised by a conserved cysteine rich zinc-binding motif and is found mainly in proteins that regulate eukaryotic transcription. Its function is to regulate protein:protein interactions and in Dnmt3a and Dnmt3b its is required for interactions with a host of repressive chromatin modifying enzymes and associated repressors such as RP58, HP1, HDACs 1 and 2, SUV39H1 and SNF2h (Bachman et al., 2001; Fuks et al., 2001; Geiman et al., 2004a). The PWWP domain of Dnmt3a and Dnmt3b is also sufficient to repress transcription in a HDAC dependant manner (Bachman et al., 2001; Fuks et al., 2001; Geiman et al., 2004a). Similar to Dnmt1, the catalytic domains of Dnmt3a and Dnmt3b are dispensable for this. In addition to these interactions both Dnmt3a and Dnmt3b interact with Dnmt1 suggesting a dual and combined role for the proteins of the *de novo* and maintenance methyltransferase families (Kim et al., 2002).

1.2.2.3.2 Genetic studies of Dnmt3a and Dnmt3b

Mice with targeted disruption of the Dnmt3a and Dnmt3b genes revealed the enzymes have non-overlapping roles in development (Okano et al., 1999). Dnmt3a -/- mice developed to term and appeared normal at birth. At around 4 weeks of age however, the mice became runted and died. Dnmt3b mutant mice however develop normally until E9.5 but die around E13.5 with multiple developmental defects including growth impairment and neural tube defects. The Dnmt3a/Dnmt3b double mutant has an even more severe phenotype and dies before E11.5 (Okano et al., 1999).

Comparisons of the three mutants before they died (E9.5) also indicated their distinct roles in establishing DNA methylation. Interestingly, it was shown that Dnmt3b is almost solely required for methylation of the minor satellite repeats. Methylation of these repeats was as low in the Dnmt3b mutant as the double mutant and not affected in the Dnmt3a mutant. Strikingly, this loss of methylation correlates with the loss of

methylation at pericentromeric satellite regions observed in the human autosomal recessive condition ICF (immunodeficiency, centromeric instability, facial abnormalities). ICF is caused by mutations in the human DNMT3B gene (Xu et al., 1999).

Observations of DNA methylation patterns in Dnmt3 mutant ES cells confirmed the effects seen in mice (Chen et al., 2003). Progressive loss of DNA methylation at various repeats and single-copy genes was observed in these cells that could be differentially rescued using different Dnmt3s. While both enzymes could rescue methylation at pMO and IAP repeats equally well Dnmt3a was more efficient at rescuing the major satellite and Dnmt3b the minor satellite (Chen et al., 2003). In addition, Dnmt3a but not Dnmt3b could rescue methylation at two out of five imprinted genes investigated (Chen et al., 2003). Further evidence for distinct roles of Dnmt3a and Dnmt3b came from a conditional knockout of the enzymes in germ cells (Kaneda et al., 2004). Offspring from the Dnmt3a conditional mutant females died around E10.5 and lacked methylation of maternal but not paternal imprints. Dnmt3a conditional males had impaired spermatogenesis and lacked methylation at two out of three paternally imprinted loci in spermatogonia. Contrastingly, Dnmt3b conditional mutants and their offspring showed no apparent phenotype and maintained their imprints. Interestingly, the phenotype of the Dnmt3a conditional mutant was virtually indistinguishable from that of Dnmt3L suggesting a key role for Dnmt3a/Dnmt3L but not Dnmt3b in imprinting.

1.2.3.3 Dnmt3L

Dnmt3L shows homology to Dnmt3a and Dnmt3b but is truncated at both N and C termini. It lacks the catalytic motifs VIII and X and most of the regulatory domain, only the PHD-like domain remains intact (Aapola et al., 2001; Hata et al., 2002) (Figure 1.2). The absence of the catalytic motifs and mutations in many of the other key residues means this proteins is catalytically inactive (Chedin et al., 2002). The protein does, however play a role in DNA methylation as disruption of the Dnmt3L gene leads to a loss of DNA methylation at maternal imprints in mice (Hata et al., 2002). Also, like Dnmt3a, Dnmt3L is required for proper spermatogenesis as its loss

leads to meiotic failure (Bourc'his and Bestor, 2004; Webster et al., 2005). These studies strongly suggest a functional interaction between Dnmt3a and Dnmt3L.

Several experiments have shown that Dnmt3L acts as a co-factor of *de novo* methyltransferases, in particular Dnmt3a. Initial *in vivo* studies showed that over expressed Dnmt3a and Dnmt3b co-immunoprecipitate and co-localise with Dnmt3L (Hata et al., 2002). Dnmt3L has subsequently been shown to have a stimulatory effect on *de novo* methylation at episomal and endogenous targets in ES cells (Chedin et al., 2002; Chen et al., 2005; Nimura et al., 2006). Subsequent *in vitro* experiments with Dnmt3L have been quite revealing about its role in DNA methylation. Dnmt3L has been shown to bind to and stimulate the activities of Dnmt3a and Dnmt3b by up to 15 fold but has no effect on Dnmt1 (Gowher et al., 2005a; Suetake et al., 2004). The stimulatory effect of Dnmt3L involves a direct interaction between the catalytic domains of the enzymes. This interaction does not affect the binding of Dnmt3a or Dnmt3b to DNA but increases the binding affinity of SAM (Gowher et al., 2005a; Kareta et al., 2006). Recent structural studies have further clarified the role of Dnmt3L in stimulating Dnmt3a (Jia et al., 2007). This study revealed that Dnmt3L and Dnmt3a form a heterotetramer structure with 2 active sites. Dnmt3L appears to stabilise the confirmation of the active-site loop of Dnmt3a by interacting with the C-terminal portion of it. This loop stabilisation may explain the stimulatory activity of Dnmt3L (Jia et al., 2007).

1.2.3.4 Targeting *de novo* methyltransferases

The targeting of *de novo* methyltransferases to specific gene loci can be envisaged to occur in one and/or two ways. Firstly, the DNMT could bind to another molecule which directs it to its target via DNA sequence or chromatin signature. As previously mentioned the PWWP domain of Dnmt3a associates with the transcriptional repressor RP58 and can act as a co-repressor (Fuks et al., 2001). This mechanism however is not dependent on the catalytic domain of Dnmt3a and DNA methylation of these sites was not shown. Recently the Dnmt3a interacting partner Dnmt3L has been strongly implicated in its targeting. Dnmt3L has been shown to bind H3 peptides *in vitro* with the association completely lost upon di or tri methylation of

H3K4 (Ooi et al., 2007). Co-crystal structures of Dnmt3L in complex with H3 1-24 showed recognition of the peptide via the PHD domain of Dnmt3L. Steric occlusion of the interaction between aspartic acid 90 in DNMT3L and methyl-lysine 4 of histone H3 regulates this interaction (Ooi et al., 2007). This study strongly implicated Dnmt3L in targeting Dnmt3a to sites lacking H3K4me. As Dnmt3L also stimulates the activity of Dnmt3a it may perform a dual role as a targeting and activating factor (Jia et al., 2007).

Alternatively, Dnmt3a and/or Dnmt3b may possess a substrate preference. Analysis of CpG methylation sites in human genome have shown a preference for meCpG to occur when it is flanked by a pyrimidine (C or T = Y) 5' and a purine (A or G = R) 3' (YCGR) (Handa and Jeltsch, 2005). When tested *in vitro*, this consensus sequence was methylated at a ~13 fold higher rate by Dnmt3a and Dnmt3b than a non consensus sequence (RCGY). This sequence preference cannot explain the differential targeting of Dnmt3a and Dnmt3b however, as both exhibited the same sequence preference *in vitro* (Handa and Jeltsch, 2005). Further sequence specificity for Dnmt3a has been implied from its co-crystal structure with Dnm3L (Jia et al., 2007). As mentioned above this structure revealed a heterotetramer structure containing 2 active sites. Further biochemical analysis showed that Dnmt3a preferentially methylated CpG sites with 10bp periodicity *in vitro*. Interestingly, the CGIs of 12 known maternally methylated DMRs show an average CpG periodicity of 10bp (Jia et al., 2007). In contrast 3 paternally methylated DMRs and 10 randomly selected unmethylated CGIs do not show the same CpG periodicity. Together these data imply that for Dnmt3a, targeting may be achieved through recognition of unmethylated H3K4 at CGIs containing a particular CpG consensus and spacing (Handa and Jeltsch, 2005; Jia et al., 2007; Ooi et al., 2007).

1.2.3 DNA methylation in disease

DNA methylation is implicated in human carcinogenesis in at least 3 distinct ways that will now be discussed. The first is the enhanced mutability of me5C that leads to C to T transition mutations at CpG sites. The second is the role of DNA hypomethylation in genomic instability. The third is the inappropriate silencing of

tumour suppressor genes via methylation of CpGs in their promoters. I will look at these 3 issues separately, focusing particularly on the role of promoter methylation and how it may be aberrantly targeted.

1.2.3.1 Increased mutability of meCpG

The human genome is strikingly deficient in CpG dinucleotides due to the spontaneous deamination of me5C to T (Shen et al., 1994). This spontaneous mutation is often not properly corrected by DNA repair enzymes and can become a heritable TpG mutation, thus depleting the genome of CpG. The consequences of unrepaired me5C/T transitions may include mutations important for cancer progression. The p53 tumour suppressor gene is a particularly interesting and well studied example of the impact of me5C/T transitions. Around 50% of all cancers have a mutation in the p53 gene and over 290 of the 393 codons in p53 have been reported to harbour mutations. 25% of mutations in p53 occur at CpG sites which are heavily methylated in all human tissues examined so far (Pfeifer et al., 2000). In fact, methylated CpGs are the most important mutational target in p53 and 5 major mutational hotspots (codons 175, 245, 248, 273 and 282) contain meCpG. The role of me5C/T transitions in p53 mutagenesis is not restricted to spontaneous deamination. Several mutagenic agents such as UV and benzo[a]pyrene which is found in tobacco smoke have been shown to preferentially cause mutations at meCpG sites in p53 (Denissenko et al., 1997; Denissenko et al., 1996; Pfeifer et al., 2000). Thus meCpG plays a major role in DNA mutagenesis through enhanced spontaneous deamination and increased mutability by DNA reactive molecules.

1.2.3.2 Genome wide hypomethylation

The genomes of human cancers tend to be hypomethylated compared to normal cells (Feinberg and Vogelstein, 1983; Gama-Sosa et al., 1983). Although this fact is often used simply as a biomarker of cancer it is possible that it plays a role in tumourogenesis as well (Manel Esteller, 2002; Moore et al., 2008). Genomic hypomethylation could be implicated in tumourogenesis in 3 ways; by inducing chromosomal instability, reactivation of transposable elements, and leading to loss of imprinting. Direct evidence of DNA hypomethylation leading to human cancers via

these three mechanisms is currently lacking. However, genetic studies in mice carrying hypomorphic alleles of Dnmt1 have shown that DNA hypomethylation leads to genomic instability and tumourogenesis. Also, studies crossing mice carrying hypomorphic *Dnmt1* alleles to the tumour susceptible $Apc^{Min/+}$ background showed inhibition of intestinal tumourogenesis but increased numbers of liver tumours. Together these studies strongly implicate DNA hypomethylation as a causative agent of cancer (Eden et al., 2003; Gaudet et al., 2003).

1.2.3.3 Promoter methylation

As previously discussed the promoters of many human genes contain CpG rich islands that tend to be unmethylated in normal tissues. Aberrant hypermethylation of the promoters of tumour suppressor genes is one of the most frequent causes of gene function loss in cancer (Jones and Baylin, 2002). It is nearly as common as genetic mutations and is responsible for silencing genes such as *p16* in melanoma, *RB1* in retinoblastoma, *APC* in colorectal cancer and *VHL* in renal-cell carcinomas (Jones and Baylin, 2002). In these cases promoter hypermethylation can account for at least one of the two 'hits' required by Knudson's two hit hypothesis for tumour suppressor inactivation (Knudson, 1971). Interestingly, promoter hypermethylation is thought to be able to account for either of the 2 inactivating hits. The mismatch-repair gene *MLH1* (mutL homologue 1, colon cancer, non-polyposis type 2) often shows promoter methylation in sporadic tumours with microsatellite instability. Promoter methylation has been implicated as a 'first hit' event as in the normal colonic epithelium of patients with colorectal cancer, the promoter of *MLH1* has also been found to be methylated. This indicates that in this case promoter methylation may be responsible for the first inactivating hit that is later followed by another mutational event (Herman et al., 1998; Nakagawa et al., 2001). It has also been reported that in many cases of familial cancer, promoter hypermethylation can account for the 'second hit' and that it only occurs at the wild type allele (Esteller et al., 2001).

There are several reasons why promoter methylation is such an important and common inactivation mechanism. Firstly, inactivation of a gene by promoter methylation appears to be equivalent to loss of function via a coding region mutation.

For example methylation of the p16 promoter results in near complete transcriptional block and an effectively null allele (Merlo et al., 1995). Secondly, promoter methylation is fully heritable through the actions of Dnmt1. This stable inheritance of promoter methylation is observed in the colon cancer cell line HCT116 (Myohanen et al., 1998). This cell line has one mutated and one wildtype allele of p16. The promoter of the wild-type allele is stably and selectively hypermethylated and silenced. Thus promoter hypermethylation is effective because it can lead to complete, heritable loss of function of tumour suppressor genes.

How aberrant promoter methylation is targeted is not currently clear. It may be a completely random event followed by growth selection or due to specific targeting via currently unknown mechanisms. A number of studies have provided tentative evidence for targeting of promoter methylation during cancer. For instance, specific groups of genes tend to be methylated in colorectal cancers leading to the 'CpG island methylator phenotype' (CIMP) (Toyota et al., 1999; Weisenberger et al., 2006). Also, during progression of a T/natural killer acute lymphoblastic leukaemia mouse model, specific genomic methylation does not appear to be random (Yu et al., 2005). Furthermore, it has recently been reported that stem cell PcG targets are 20-fold more likely to have cancer specific promoter methylation than non targets (Ohm et al., 2007; Schlesinger et al., 2007; Widschwendter et al., 2007). Together, these data indicate that promoter methylation in cancer may be a directed event, possibly via pre-existing repressive chromatin marks. Alternatively, it is possible that gene silencing via other mechanisms such as PcG merely 'predispose' promoters to methylation in cancer. In this case, DNA methylation follows inactivation by another mechanism and is therefore correlated with tumourogenesis but not causative of it.

1.3 Chromatin remodelling

The organisation of DNA and histones into chromatin provides an important role in compacting long stretches of DNA into tiny nuclei. This compact structure however affects the accessibility of DNA to the proteins that act on it. Thus in order for the proteins involved in transcription, replication, repair and recombination to access DNA, this barrier must be overcome. There are 2 recognised pathways for transiently facilitating accessibility to the genome. The first is the posttranslational modification of histone tails at numerous residues by proteins or protein complexes. The second is the modulation of histone:DNA interactions by ATP hydrolysing protein machines. Both of these processes are highly dynamic, functionally linked and influence chromatin accessibility both positively and negatively (Narlikar et al., 2002). This section will briefly discuss the influence of chromatin modifications to different chromatin related processes before discussing the roles and mechanisms of ATP dependant chromatin remodelling in more depth.

1.3.1 Chromatin remodelling by histone modification

The N-terminal tails of histones are the targets of several different classes of posttranslational modification of which methylation, acetylation and phosphorylation are the best characterised (Kouzarides, 2007). The roles of these modifications are to disrupt the local chromatin environment and/or recruit other proteins. Thus histone modifications are able to influence the biological processes inhibited by chromatin inaccessibility.

1.3.1.1 Disruption of chromatin structure by histone modifications

Chromatin structure can be disrupted by histone modifications affecting the contact between different histones or the interaction between histones and DNA. Lysine acetylation has the most potential to influence chromatin structure as it neutralises the basic charge of lysine. A number of studies have shown that internucleosomal contacts between the N-terminal tail of H4 and an acidic patch in H2A are required for higher order chromatin compaction (Dorigo et al., 2003; Dorigo et al., 2004). Due

to the neutralising effect of lysine acetylation it would be predicted that this modification would inhibit this interaction. In fact, biophysical studies on recombinant nucleosome arrays have directly addressed this expectation. Using chemical ligation to construct core histones with modified tail peptides, it has been possible to show that H4K16 acetylation directly inhibits formation of 30nm chromatin fibres and higher order chromatin structures (Shogren-Knaak et al., 2006). Thus modification of this specific histone tail residue is able to directly impact higher order chromatin structure. Thus far this is the only histone modification shown to be a direct effector of chromatin structure but others may yet to be discovered.

1.3.1.2 Recruitment of proteins by histone modifications

The second way histone modifications can influence chromatin structure is through recruitment or exclusion of non-histone proteins that recognise modification via specific domains. For example methylation is recognised by chromo-like domains and PHD domains, acetylation by bromodomains and phosphorylation by a domain in 14-3-3 proteins (Kouzarides, 2007). These domains are found in a number of proteins that are recruited to specific modifications and influence chromatin function and structure. One well characterised example is the chromodomain of HP1 that specifically binds di and tri methylated H3K9 (Bannister et al., 2001; Jacobs and Khorasanizadeh, 2002). This interaction has been shown to be required for HP1 localisation to heterochromatic regions and its transcriptional silencing ability (Bannister et al., 2001; Nakayama et al., 2001; Platero et al., 1995). Once recruited, HP1 can influence chromatin structure through interactions with a host of other chromatin modifying enzymes such as Dnmt1, Dnmt3b and the H3K9 methyltransferase Suv39h1. HP1 can also dimerise, thus potentially directly influencing chromatin structure via internucleosomal contacts (Maison and Almouzni, 2004). A number of recent studies have shown the importance of methyl-lysine recognition by PHD domain containing proteins in influencing chromatin structure. For example the NURF chromatin-remodelling complex subunit BPTF recognises H3K4me3 via a PHD domain (Li et al., 2006a; Wysocka et al., 2006). This interaction is believed to recruit the SNF2L ATPase to activate expression of the HOXC8 gene (Li et al., 2006a; Wysocka et al., 2006). H3K4me3 recognition by

PHD domain containing proteins can also lead to gene repression. For example, the PHD domain containing protein ING2 recruits the repressive deacetylase complex Sin3a-HDAC1 to the cyclinD1 promoter and silences it (Pena et al., 2006; Shi et al., 2006). Thus the actual effect on chromatin does not depend on the recognition domain, rather the proteins that interact with it.

Histone modifications can also have a negative effect on Protein:histone interactions. H3K4 methylation for instance has been shown to inhibit binding of the repressive NuRD complex to chromatin thus facilitating transcription (Nishioka et al., 2002). H3K4 methylation has also been shown to disrupt the binding of the PHD domain of Dnmt3L to histone tails, perhaps indicating a mechanism to protect actively transcribed genes from inappropriate DNA methylation (Ooi et al., 2007). Another interesting example of this phenomenon is the methylation of H3R2. This modification has been found at transcriptionally inactive regions of the genome and inhibits the binding and function of the H3K4 methyltransferase MLL (Guccione et al., 2007; Kirmizis et al., 2007).

The prevalence of modification sites on histone tails makes it very likely that several modifications may occur on each tail concurrently. These modifications could then influence the binding of chromatin modifying proteins. A number of examples of this so called 'crosstalk' have been identified. As discussed above H3K9 methylation recruits HP1 to chromatin. This interaction is lost during mitosis, but H3K9 methylation is not (Fischle et al., 2005; Hirota et al., 2005). Recent work has identified phosphorylation of H3S10 as a key modification regulating this interaction. When H3S10 is phosphorylated by Aurora B, the interaction between HP1 and H3K9me3 is lost (Fischle et al., 2005; Hirota et al., 2005). Thus combinations of histone modifications can differentially affect protein recruitment and the downstream effects of them.

1.3.2 ATP dependant chromatin remodelling

Chromatin structure can also be altered by a family of enzymes that use the energy derived from ATP hydrolysis to directly modulate histone:DNA interactions. This

family of genes was originally identified as a result of genetic screens for genes involved in regulating mating type switching (SWI) and sucrose fermentation (Sucrose Non-Fermenting). One of the mutants was found to be complemented by the protein Snf2p which was subsequently found to be the catalytic core of the multi-subunit SWI/SNF complex (Abrams et al., 1986). Since then, many proteins have been identified that are related to Snf2p through sequence similarity. Enzymes of this family tend to reside in large protein complexes and are involved in a wide range of processes including transcription, DNA repair, recombination, replication and chromatin assembly (Narlikar et al., 2002). These enzymes use the energy from ATP hydrolysis to disrupt the binding of DNA around histones. This in turn facilitates the repositioning of nucleosomes to a new translational position or the eviction of the histone octamer (Flaus and Owen-Hughes, 2004). The change in nucleosomal positioning and/or density is believed to lead to local changes in chromatin structure.

1.3.2.1 The SNF2 family of chromatin remodelling enzymes

Although functionally diverse, all of these enzymes contain a core SNF2 ATPase region homologous to the large helicase superfamily 2 (SF2) (Eisen et al., 1995; Flaus and Owen-Hughes, 2004) (Figure 1.3A). In particular, SNF2 ATPases share the seven conserved sequence motifs common to the SF2 superfamily of helicases. These are labelled sequentially I, Ia, II, III, IV, V and VI and cluster in a region ~400 amino acids in length (Bork and Koonin, 1993; Eisen et al., 1995). These helicase motifs are the core ATPase domains involved in binding to DNA and binding and hydrolysing ATP (Figure 1.3B). Although the overall structure of the helicase domains is conserved throughout the SF2 superfamily, SNF2 family members have some distinguishing features. In particular, the helicase domain is significantly longer than in other SF2 helicases due to increased spacing between helicase motifs III and IV (Flaus et al., 2006). Interestingly, this linker region contains a number of amino acid blocks conserved throughout the SNF2 family (Flaus et al., 2006) (Figure 1.3B). As this feature is so highly conserved it is likely to confer unique properties to the ATPase motor of SNF2 enzymes.

SNF2 remodelers can be split into at least 4 classes based on sequence homology of the ATPase core: SWI2/SNF2, ISWI, NuRD/Mi-2/CHD and INO80/SWR1 (Cairns, 2007; Eisen et al., 1995; Flaus et al., 2006) (Figure 1.3A). Other classes such as Rad54 also exist but these are more distantly related to the others and exhibit divergent activities (Flaus et al., 2006). Homology of the core ATPase domains is also linked to additional domains found within the different classes of enzyme. Most chromatin remodelling enzymes are found in multisubunit complexes containing additional proteins involved in targeting the complexes to chromatin and modulating their activities (Narlikar et al., 2002) (Figure 1.3C).

1.3.2.1.1 The SWI2/SNF2 class

The SWI2/SNF2 class consists of the enzymatic cores of a number of large stable multisubunit complexes. These include the yeast mating-type switching/sucrose non-fermenting (SWI2/SNF2) and STH1 proteins that are members of the SWI/SNF and RSC (restructure of chromatin) complexes respectively (Cairns et al., 1996). The human hBrm (Brahma) and Brg1 (Brahma related gene) proteins are alternative catalytic subunits of the hSWI/SNF complex and Brm is the core of *Drosophila* dSWI/SNF (Elfring et al., 1994; Kim et al., 2002). In addition to the conserved ATPase domain SWI2/SNF2 enzymes contain a bromodomain at their C-terminus (Figure 1.3C). Bromodomains recognise acetylated lysines and in the case of SWI2/SNF2 is required for stable association with a target promoter (Hassan et al., 2002). SWI2/SNF2 enzymes have a wide range of *in vivo* functions including transcriptional regulation, cell cycle progression, immune responses, chromosome segregation and DNA repair (Chi, 2004; Moshkin et al., 2007). *In vitro,* enzymes of this class can bind hyper acetylated nucleosomes in correlation with their *in vivo* function in transcriptional activation. They also exhibit DNA translocase and chromatin remodelling activities (Narlikar et al., 2002).

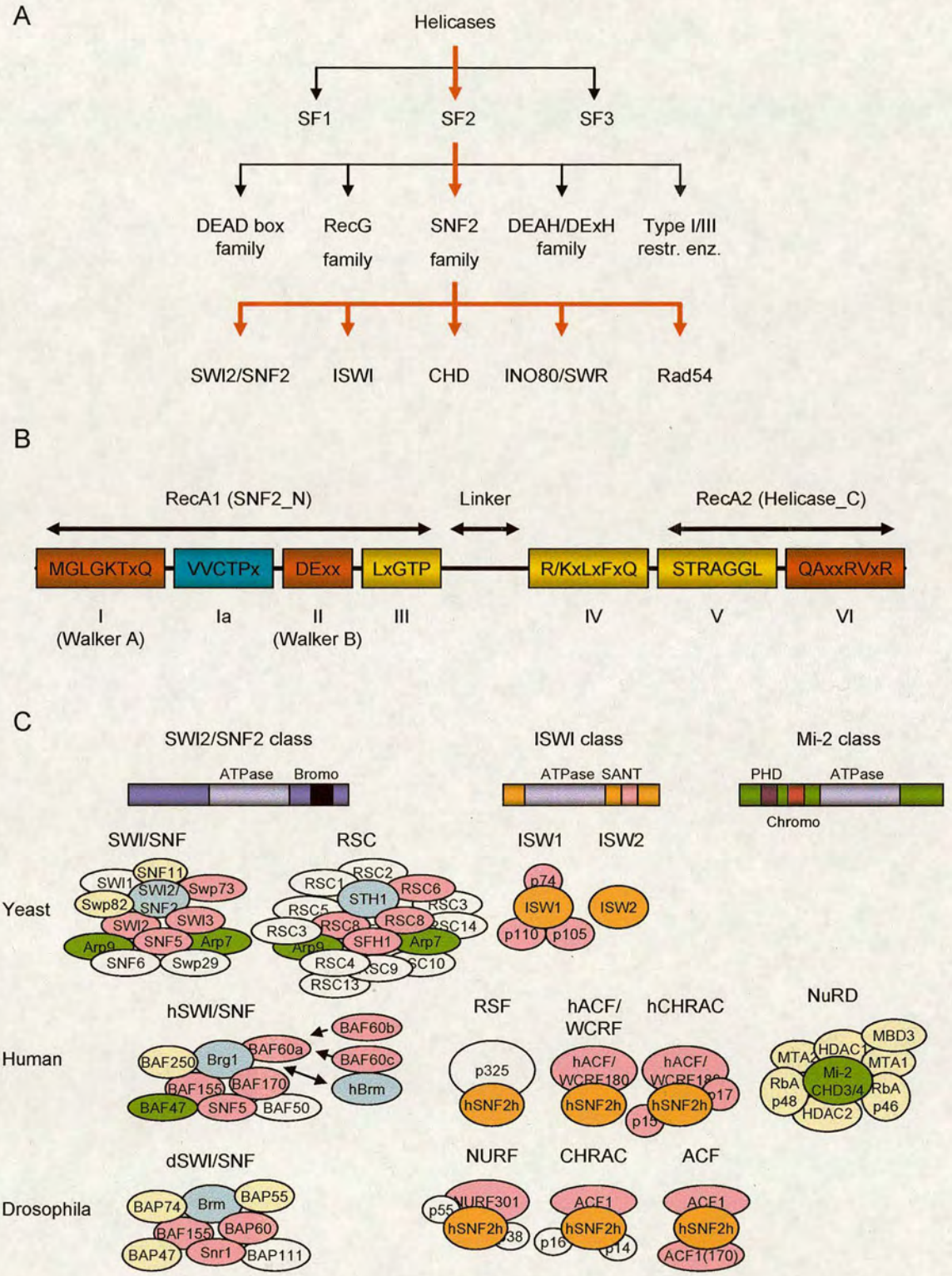

A
Helicases
SF1
SF2
SF3
DEAD box family
RecG family
SNF2 family
DEAH/DExH family
Type I/III restr. enz.
SWI2/SNF2
ISWI
CHD
INO80/SWR
Rad54
B
RecA1 (SNF2_N)
Linker
RecA2 (Helicase_C)
MGLGKTxQ
VVCTPx
DExx
LxGTP
R/KxLxFxQ
STRAGGL
QAxxRVxR
I (Walker A)
Ia
II (Walker B)
III
IV
V
VI
C
SWI2/SNF2 class
ISWI class
Mi-2 class
ATPase
Bromo
ATPase
SANT
PHD
ATPase
Chromo
SWI/SNF
RSC
ISW1
ISW2
Yeast
SWI1
SNF11
SWI2/
Swp73
Swp82
SNF2
SWI2
SWI3
Arp9
SNF5
Arp7
SNF6
Swp29
RSC1
RSC2
RSC6
STH1
RSC3
RSC5
RSC8
RSC14
SFH1
RSC4
RSC9
RSC10
RSC13
p74
ISW1
ISW2
p110
p105
hSWI/SNF
RSF
hACF/ WCRF
hCHRAC
NuRD
Human
BAF60b
BAF60a
BAF60c
BAF250
Brg1
hBrm
BAF155
BAF170
BAF47
SNF5
BAF50
p325
hSNF2h
hACF/ WCRF180
hSNF2h
hSNF2h
p17
p15
MTA2
HDAC1
MBD3
MTA1
Mi-2
CHD3/4
RbA p48
RbA p46
HDAC2
dSWI/SNF
NURF
CHRAC
ACF
Drosophila
BAP74
Brm
BAP55
BAF155
BAP60
BAP47
Snr1
BAP111
p55
NURF301
hSNF2h
p38
ACF1
hSNF2h
p16
p14
ACF1
hSNF2h
ACF1(170)

Figure 1.3. The SNF2 family of chromatin remodelling enzymes.
A. The SNF2 family of chromatin remodelling enzymes is a member of the larger, SF2 superfamily. This family can be further subdivided in to classes, of which five are shown. The red line follows the different classes through their classification.
B. Illustration of the seven conserved helicase motifs defined by Gorbalenya and Koonin indicated by roman numerals. The letters within the boxes indicate the conserved amino acids found in these motifs. The two RecA like domains are labelled alongside their ENSEMBL classification. Note the extended linker region that links the two RecA like domains. Boxes coloured red denote motifs involved in ATP binding and hydrolysis, the blue box denotes the motif involved in binding DNA.
C. Diagram of the protein complexes in which a number of SNF2 enzymes of three different classes reside. Figure adapted from (Narlikar et al., 2002).

1.3.2.1.2 The ISWI class

Members of the ISWI (Imitation Switch) class are the core ATPases of a number of different complexes. In yeast the proteins ISW1 and ISW2 are found in the ISW1 and ISW2 complexes respectively. The human ISWI homologue hSNF2h is found in a number of different complexes including hACF (ATP-utilizing chromatin assembly and remodeling factor), hCHRAC (chromatin-accessibility complex) and NoRC (Poot et al., 2000; Strohner et al., 2001). The *Drosophila* ISWI is also found in several complexes including NURF (nucleosome-remodeling factor), CHRAC and ACF (Ito et al., 1997; Tsukiyama et al., 1995; Tsukiyama and Wu, 1995; Varga-Weisz et al., 1997). Members of the ISWI class contain 2 SANT domains, one of which facilitates binding to histones (Boyer et al., 2002) (Figure 1.3C). ISWI class members are involved in transcription, DNA replication through heterochromatin, chromatin assembly and Wnt signalling (Corona and Tamkun, 2004). Purified enzymes of this class are able to assemble nucleosomes in an ATP independent manner *in vitro*. They are also able catalyse the formation of regularly spaced nucleosomal arrays and facilitate transcriptional activation (Langst and Becker, 2001; Langst et al., 1999).

1.3.2.1.3 The NuRD/Mi-2/CHD class

NuRD/Mi-2/CHD class members are components of complexes such as human and *Drosophila* NuRD (nucleosome remodelling and deacetylation). NuRD is a large complex that includes transcriptional repressors and co-repressors such as HDACs

and MBD proteins. The catalytic members of this family contain additional PHD fingers and CHD domains that are required for protein-protein interactions and complex formation (Zhang et al., 1998) (Figure 1.3C). The *in vivo* function of NuRD is transcriptional repression which in turn regulates development (Bowen et al., 2004; Le Guezennec et al., 2006). In agreement with its *in vivo* role, the NuRD complex deacetylates chromatin and promotes the formation of regularly spaced nucleosomal arrays *in vitro* (Tong et al., 1998; Zhang et al., 1998).

1.3.2.1.4 INO80/SWR1

The INO80/SWR1 family consists of 'split ATPases' which have an insertion between their ATPase domains. Out with the catalytic domains, these proteins show no homology to other SNF2 family members. INO80 is found in a large multisubunit complex in yeast, *Drosophila* and mammals. The N-terminal domains of INO80 interact with other proteins in the INO80 complex in yeast (Shen et al., 2000; Shen et al., 2003). This complex is involved in a number of remodelling processes such as transcription, recombination and DNA repair. It also has ATP-dependent nucleosome mobilisation activity *in vitro* (Bao and Shen, 2007). SWR1 is also found in large complexes in a number of species. These complexes are involved in a number of processes including deposition of histone H2AZ at sites of active transcription, chromosome stability and cell-cycle checkpoint adaptation. *In vitro*, SWR1 can catalyse histone dimer exchange of H2A-H2B for H2AZ-H2B (Mizuguchi et al., 2004).

1.3.2.2 Roles of the additional subunits present in SNF2 protein complexes

Most tested SNF2 enzymes have remodelling activities out with the complexes they are found in raising questions as to the roles of the other complex members. It can be envisaged that these proteins act either to target the complex to chromatin, or to modulate its activity.

1.3.2.2.1 Targeting chromatin remodelling complexes

Chromatin remodelling complexes can be targeted via recognition of histone modifications by recognition domains within the proteins as mentioned above. They

can also be targeted to chromatin via interactions with sequence specific transcription factors. For example the Snf5, Swi1, and Swi2/Snf2 subunits of yeast SWI/SNF have been shown to interact with the yeast activators Gcn4 and Hap4 in a redundant manner targeting the complex to chromatin (Neely et al., 2002). The murine SWI/SNF complex is also indirectly targeted via Baf60c. This protein has also been shown to be required for Brg1 interaction with cardiac transcription factors, and loss of Baf60c leads to defects in heart morphogenesis (Lickert et al., 2004). Another mammalian SWI/SNF complex PBAF is believed to be directly targeted via a specific component termed BAF200. This protein has a number of zinc fingers and is believed to specifically target PBAF to a subset of interferon inducible genes (Yan et al., 2005). Other chromatin remodelling enzymes are directly recruited by repressor proteins. For example ISW1 is recruited to the PHO8 promoter by Cbf1p where it represses transcription by causing dissociation of TBP (Moreau et al., 2003). Another ISWI enzyme ISW2 is recruited to repress early meiotic genes by the sequence specific transcription factor Ume6 (Goldmark et al., 2000). Thus the composition of different chromatin remodelling complexes can play a role in directly or indirectly targeting them to specific genomic loci.

1.3.2.2.2 Modulating remodeler activities

Another function of additional subunits present in chromatin remodelling complexes is modulation of the ATPase activity. This is best understood in the ISWI containing complexes where the additional subunits affect both the efficiency and outcome of remodelling. For example, the ISWI complexes ACF and CHRAC, which both contain Acf1 show enhanced ability to mobilise histone octamers than ISWI alone (Langst et al., 1999). Interestingly the presence of Acf1 also alters the product of the remodelling reaction. On its own ISWI moves nucleosomes to the end of the DNA fragment used for analysis, but in the presence of Acf1 it positions nucleosomes at the centre of the DNA fragment (Langst et al., 1999). Another ISWI complex, NURF has yet another reaction outcome, where nucleosomes are moved to an intermediate position. This is dependant on the NURF301 subunit in this complex (Hamiche et al., 1999). Thus in the case of ISWI different interacting partners can lead to entirely different reaction outcomes. A possible explanation for these differences is that

additional subunits may act to stabilise energetically unfavourable products, such as centrally positioned nucleosomes. Additionally, it has recently been shown that ACF activity is sensitive to the length of DNA flanking nucleosomes. The remodelling activity of the enzyme decreases as the length of flanking DNA decreases. This sensitivity drives ACF1 to reposition nucleosomes to a central position as this positioning gives rise to dynamic equilibrium (Yang et al., 2006). Monomeric ISWI does not show the same sensitivity and thus does not provide the same reaction outcome.

1.3.2.3 Mechanisms of ATP dependant chromatin remodelling

Although SNF2 enzymes vary functionally, they are all molecular motors that utilize ATP hydrolysis to rearranging the histone:DNA contacts within a nucleosome. Several biochemical assays have been used to demonstrate that modulation of histone:DNA interactions by SNF2 enzymes expose non-nucleosomal DNA. These include the generation of altered patterns of nuclease and/or restriction enzyme accessibility to nucleosomal DNA and direct observation of changes in the translational positions of histone octamers (Hamiche et al., 1999; Langst et al., 1999; Meersseman et al., 1992) . The most commonly used assay is the classical 'sliding assay' that directly observes the repositioning of the histone octamer on DNA (Eberharter et al., 2004). In this assay nucleosomes are assembled on DNA that contains a high affinity octamer positioning sequence. This positions the nucleosome at this site. Addition of the SNF2 enzyme and ATP can lead to octamer repositioning that can be observed by migration change through a native acrylamide gel (Eberharter et al., 2004). Members of all 4 classes of chromatin remodelling enzyme can change the translational position of nucleosomes on DNA in sliding assays but the outcome and efficiency of remodelling varies. For example, the ISWI containing complexes CHRAC and ACF are able to efficiently reposition nucleosomes to central positions and promote equal spacing of DNA between nucleosomes (Hamiche et al., 1999; Langst and Becker, 2001; Langst et al., 1999). On the other hand, the SWI2/SNF2 remodelers act in the opposite manner, moving nucleosomes to DNA ends and randomizing equally spaced nucleosomes (Jaskelioff et al., 2000; Owen-Hughes et al., 1996; Whitehouse et al., 1999). These different outcomes may

indicate different *in vivo* roles at altering chromatin structure for the remodelers. For example, enzymes that disrupt ordered nucleosomes may give rise to relatively open and accessible chromatin. Conversely, enzymes that generate evenly spaced nucleosomes may promote formation of tightly packed, closed chromatin. Despite the differences in outcome of the different remodeler families a common outcome was that all tested enzymes are able to transiently disrupt histone:DNA contacts, thus facilitating sliding to the nucleosome to a new position. This has led to speculation that all chromatin remodelers act on nucleosomes in a common manner. These modes of repositioning could theoretically expose small stretches of previously inaccessible DNA to factors that interact with them. It is important to acknowledge however, that nucleosome repositioning by classical sliding may not be sufficient to expose substantial tracts of DNA within tightly packed nucleosomes. To achieve this other mechanisms including covalent modification of histone tails may be required to occur simultaneously. Alternatively, SNF2 enzymes may use other means to facilitate or complement this including conformational changes to nucleosome structure, release of the histone octamer or histone dimer exchange (Lorch et al., 1999; Mizuguchi et al., 2004; Studitsky et al., 1994).

1.3.2.3.1 Bulge diffusion and twist defect diffusion

For nucleosomes to be repositioned on DNA the 14 histone:DNA contacts need to be broken. This requires ~14 kcal/mol of energy, thus providing a substantial energetic obstacle for remodelling enzymes to overcome. As ATP hydrolysis provides only 7.3 kcal/mol it is unlikely that the remodelling process involves simultaneous disruption of all the contacts. Models for nucleosome remodelling take this into account and reason that it is likely to occur through a series of smaller dissociation steps. These transient dissociations can then propagate around the nucleosome, leading to repositioning. Two models have been proposed to explain the mechanisms of SNF2 chromatin remodelling enzymes: "Bulge Diffusion" and "Twist Defect Diffusion".

The first model (Bulge Diffusion) involves the creation of a DNA bulge or loop on the surface of the nucleosome that can be propagated around the surface of the nucleosome. In this model a small number of DNA:histone contacts are transiently

disrupted by peeling the DNA off the nucleosome in an ATP dependant manner. If this transient peeling of DNA is followed by rebinding of a more distal DNA sequence a DNA loop will form. This loop is able to stochastically diffuse around the nucleosome giving rise to a distribution of repositioned nucleosomes (Figure 1.4A). The actions of the particular remodeler, including where it binds the nucleosome will then drive the reaction towards a particular outcome. Depending on the size of the loop generated on the nucleosome surface the DNA may be more accessible to nucleases, other DNA binding proteins and various DNA binding chemicals. A number of studies have exploited this property to provide evidence for the creation of a DNA loop as an intermediate step in chromatin remodelling. One study utilised a novel assay based on the preferred intercalation of ethidium bromide (EB) into free DNA (Strohner et al., 2005). In this assay, nucleosomes with no flanking DNA were incubated with the remodelling complex ACF and EB. Using a laser to induce single strand breaks at intercalation sites they observed ACF dependent DNA exposure within the nucleosome. This provides evidence for the formation of a DNA loop by ACF. Another study observed the formation of DNA loops on the surface on nucleosomes by the RSC and SWI/SNF complexes. Using optical tweezers they detected shortening of free and nucleosomal DNA over a range of force conditions in the presence of the remodelers (Zhang et al., 2006b). DNA shortening was interpreted as formation of a DNA loop on the surface of the nucleosome. Interestingly, the shortening of DNA was more stable to higher forces when the DNA was assembled into nucleosomes indicating that the remodelers were being tethered to the nucleosome (Zhang et al., 2006b).

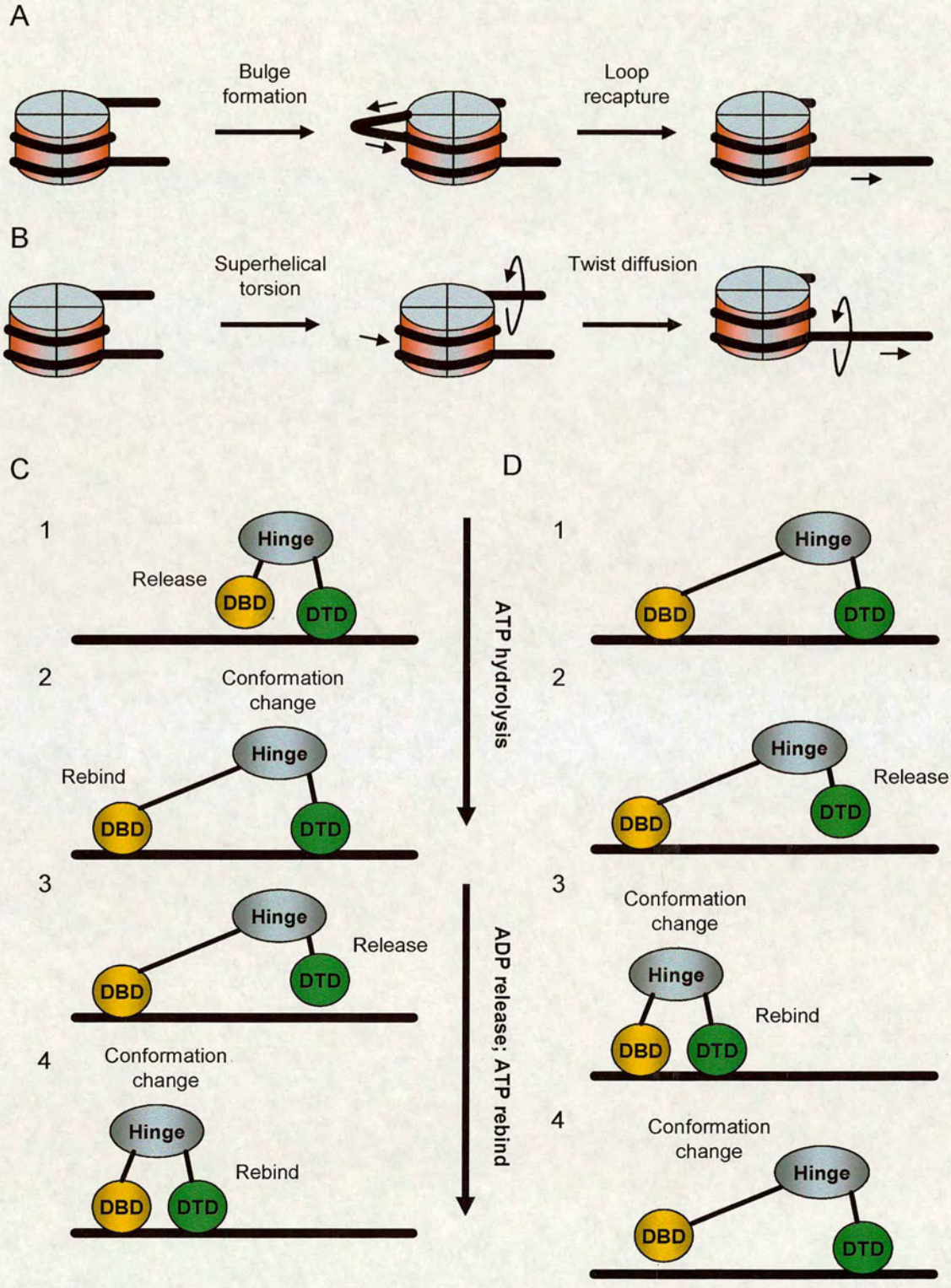

A
Bulge formation
Loop recapture
B
Superhelical torsion
Twist diffusion
C
D
1
Hinge
Release
DBD
DTD
2
Conformation change
Rebind
3
Release
4
Conformation change
Rebind
ATP hydrolysis
ADP release; ATP rebind

Figure 1.4. Proposed mechanisms of chromatin remodelling by SNF2 enzymes.
A. Bulge diffusion – This model proposes the generation and propagation of a transient DNA loop around the nucleosome can lead to its repositioning.
B. Twist defect diffusion – This model proposes that DNA distortions caused by local twist defects can disrupt DNA:histone contacts. The twist can then propagate around the nucleosome leading the repositioning.
C + D. The 'DNA inchworm' model of DNA translocation that may provide the motor force to drive the above mechanisms. Binding and hydrolysis of ATP leads to conformational changes in the SNF2 motor that cause it to translocate DNA 3'-5'. The model illustrated in **C** envisages the DNA binding domain (DBD) providing a passive anchor for the DNA tracking domain (DTD) to inch along DNA. In **D** the DBD is actively feeding DNA into the DTD which subsequently rebinds DNA in an advanced position leading to translocation.

The second model (Twist Defect Diffusion) takes advantage of the high torsional flexibility of DNA. This model assumes that DNA is distorted by local twist defects that disrupt histone:DNA contacts. The twist is then able to migrate stochastically along the DNA allowing it to propagate over the surface of the nucleosome in a manner analogous to a cork-screw (Figure 1.4B). In this model the propagation of twist would lead to incremental replacement of histone:DNA contacts with neighbouring ones. The twist defect model is supported by crystal structures of nucleosome core particles that contain 1 less base pair than would be expected for a symmetrical structure (Edayathumangalam et al., 2005; Richmond and Davey, 2003b). This is because the DNA is over twisted within the core indicating that twist defects can be accommodated within the nucleosome. Experimental evidence for the twist defect model has come from studies looking at the ability of SNF2 enzymes to generate DNA twist (Gavin et al., 2001; Havas et al., 2000). In one study the ability of a number of different enzymes to form DNA cruciforms as a measure of superhelical torsion was assayed. It was demonstrated that a range of SNF2 enzymes possess this ability. This indicates that these enzymes have the ability to twist DNA and that this may be a primary activity of these enzymes (Havas et al., 2000). Another study looked at the ability of SWI/SNF to remodel nucleosomes topologically constrained on short DNA loops. It was found that SWI/SNF is unable to remodel nucleosomes on these short DNA loops but this ability is rescued by topoisomerases indicating that SNF2 enzymes have the ability to twist DNA (Gavin et al., 2001).

1.3.2.3.2 DNA translocation provides the motor for breaking DNA:histone contacts

The two proposed mechanisms both give reasonable, testable predictions of nucleosome remodelling but do not completely explain what the motor force driving remodelling is. Evidence that SNF2 enzymes are DNA translocases indicates that this may be the mechanism by which DNA:histone contacts are broken and DNA is driven around the nucleosome (Saha et al., 2005; Whitehouse et al., 2003; Zofall et al., 2006). It could be envisaged that if the remodeler binds the histone octamer at a defined location, DNA would then be pumped around the nucleosome by the translocase activity. A number of studies have shed light on the mechanism by which DNA translocation may occur by SF2 enzymes, and how this may be translated by SNF2 members into nucleosome repositioning. Monomeric SF2 DNA translocases conform to a 'DNA inchworm' model of movement. This model involves coordinated movement of a DNA binding domain (DBD) and a DNA tracking domain (DTD) (Singleton et al., 2007) (Figure 1.4C). The DTD contains two RecA like domains flanking an ATP binding pocket and a DNA binding module. ATP binding and hydrolysis induces a conformational change between the domains resulting in a movement of the domain 1bp 3' to 5' along the DNA (Dillingham et al., 2002; Velankar et al., 1999). The role of the DBD is envisaged to be either to anchor the DNA at an advanced position following the action of the DTD or to actively feed DNA into the DTD. In the first of the possibilities the DBD binds the DNA in front of the DTD and provides an anchoring point on the DNA. During ATP hydrolysis the DBD releases and rebinds further along the DNA providing a new anchor (Figure 1.4C 1-2). The release of ADP and binding of fresh ATP completes the reaction cycle and leads to an inching forward of the motor (Fitzgerald et al., 2004) (Figure 1.4C 3-4). The alternative is that the DBD is already bound at the advanced position, from where it actively pulls DNA into the DTD upon ATP hydrolysis (Figure 1.4D 1-2). The DTD subsequently rebinds to DNA at an advanced position when ADP is released leading to movement along DNA (Figure 1.4D 3). This would then be followed by release and rebinding of the DBD further along the DNA (Figure 1.4D 4).

The effect of this inchworm type of translocation can explain nucleosome movement by SNF2 enzymes such as ISWI and SWI/SNF. Although the exact mechanism of theses enzymes may differ, the basic assumption that DNA translocation drives remodelling holds for them both. In both cases the enzymes appear to use DNA inchworm translocation to cause local disruptions in histone:DNA contacts. For SWI/SNF, these disruptions are proposed to occur by pumping small segments of DNA towards the dyad by the DTD (Saha et al., 2005). An interaction of the enzyme with the histone octamer is believed to hold it in place and release and recapture of the DBD to guide translocation across the nucleosome (Zhang et al., 2006b). In contrast ISWI is believed to initially cause disruptions that are propagated by the DTD from an internal site to the nucleosome entry/exit site. The DBD binds tightly to DNA here and this causes a DNA loop to form (Kagalwala et al., 2004; Zofall et al., 2006). This loop may then propagate back past the DTD and through the nucleosome. Thus the initial direction of DNA translocation is opposite to the direction of nucleosome movement. As is the case for SWI/SNF, ISWI can bind to the histone octamer and in fact this interaction is crucial for proper nucleosome sliding (Clapier et al., 2001; Shogren-Knaak et al., 2006). Thus, the motor force for driving chromatin remodelling is believed to be the ability of these enzymes to translocate DNA across the nucleosome (Cairns, 2007).

1.3.2.4 *In vivo* roles of chromatin remodelling enzymes

SNF2 ATPases play a role in modulating protein:DNA reactions critical for processes such as transcriptional control, DNA replication, DNA repair and recombination. As such, mutations in these enzymes often have severe consequences for the organism. For example ISWI mutations in *Drosophila* affect cell viability and gene expression during development leading to larval lethality (Deuring et al., 2000). Additionally, mice genetically deficient for the mammalian ISWI homologue SNF2h also show early embryonic lethality (Stopka and Skoultchi, 2003). Mutations in other SNF2 family members are also fatal. Homozygous null mutations in the SWI/SNF class member BRG1 leads to periimplantation stage death for example (Bultman et al., 2000). Interestingly, heterozygous BRG1 mutations predisposed the mice to tumours indicating a possible role of SNF2 enzymes in cancer (Bultman et al., 2000).

Additionally, a number of human diseases have been associated with mutations of SNF2 ATPases. For example the X-linked alpha thalassemia / mental retardation syndrome (ATR-X syndrome) is caused by mutations of the ATRX gene (Borgione et al., 2003; Gibbons et al., 2003; Gibbons et al., 1995). This disease is characterised by severe mental retardation, a characteristic facial appearance, a form of anaemia known as alpha thalassaemia and abnormal genital development (Villard and Fontes, 2002). The role mutations of ATRX play in causing this disease are currently not well understood but may be linked to its role in DNA methylation (Gibbons, 2006). Another disease caused by mutations in a SNF2 enzyme is Cockayne Syndrome B, caused by mutations in the CSB protein (Troelstra et al., 1992). The disease is characterised by increased sensitivity to UV radiation and neurodevelopmental abnormalities. CSB protein is believed to be involved in DNA repair, thus demonstrating the importance of SNF2 enzymes in a variety of processes (Osterod et al., 2002). Finally, as would be expected from a large group of proteins with fundamental roles in DNA processes, mutations in a number of SNF2 enzymes are implicated in cancer. These include the SWI/SNF proteins BRG1 and BRM that are found to be mutated or showing decreased expression in a variety of different cancer cell lines (Decristofaro et al., 2001; Reisman et al., 2003; Wong et al., 2000). Other remodelers linked to cancer include the NuRD complex ATPase subunit CHD5 and the SWI/SNF core subunit SNF5 (Bagchi et al., 2007; Biegel et al., 1999; Sevenet et al., 1999; Versteege et al., 1998).

1.4 Chromatin remodelling and DNA methylation

Although mechanistically completely different, there is strong evidence that the processes of DNA methylation and chromatin remodelling are intrinsically linked. Genetic studies in *Arabidopsis thaliana* uncovered a role for a SNF2 enzyme termed DDM1 (decrease in DNA methylation 1) in DNA methylation (Jeddeloh et al., 1999; Vongs et al., 1993). Subsequent knockout studies on the murine homologue termed LSH (lymphoid specific helicase) revealed the requirement of this protein for proper DNA methylation is conserved (Dennis et al., 2001). There are also at least 2 other SNF2 enzymes believed to play a role in DNA methylation in mammals, the ATRX protein and SNF2h. The current understanding of the role of SNF2 enzymes in DNA methylation will now be discussed, with particular focus on DDM1 and LSH.

1.4.1 ATRX and SNF2h influence DNA methylation

Mutations in the *ATRX* gene cause X-linked alpha thalassemia / mental retardation syndrome (ATR-X) (Borgione et al., 2003; Gibbons et al., 2003; Gibbons et al., 1995). Aside the characteristic disease phenotypes discussed above ATR-X patients also show specific changes in DNA methylation patterns. This is characterised by DNA hypomethylation at rDNA repeats, DNA hypermethylation at Y-specific satellite and perturbed DNA methylation at subtelomeric repeats (Gibbons et al., 2000). The precise role of ATR-X in DNA methylation is currently unknown and no interaction between ATR-X and DNMTs has been reported. However, ATR-X has been shown to interact with other members of the DNA methylation and chromatin machinery including HP1α, the polycomb protein EZH2, the transcriptional co-repressor Daxx and the MeCpG binding protein MeCP2 (Cardoso et al., 1998; Gibbons et al., 2003; Guerrini et al., 2000; McDowell et al., 1999; Nan et al., 2007; Tang et al., 2004; Xue et al., 2003). Thus ATRX may play a role in many chromatin related processes and mutations may lead indirectly to changes in DNA methylation patterns.

SNF2h, the human homologue of ISWI is a member of the Nucleolar Remodeling Complex (NoRC) (Strohner et al., 2001). This complex contains at least 2 proteins, SNF2h and TIP5 (TTF interaction protein 5) and is believed to directly mediate epigenetic gene silencing of rDNA (Santoro and Grummt, 2005; Santoro et al., 2002; Zhou et al., 2002). Epigenetic silencing of rDNA repeats has been studied artificially using over expression of TIP5. This has shown that repression of the repeats involves a temporal programme of NoRC recruitment followed by histone deacetylation and H3K9 dimethylation. These histone modifications are in turn followed by, and required for, DNA methylation at a CpG site within the rDNA promoter (Santoro and Grummt, 2005). The role of NoRC however is not merely that of recruitment as catalytically inactive $SNF2h^{K221R}$, presumably acting as a dominant negative protein still induces histone deacetylation of the rDNA promoter. Surprisingly, in this system transcriptional repression and DNA methylation do not occur (Santoro and Grummt, 2005). A subsequent study has shown that TIP5 over expression leads to repositioning of nucleosomes at the rDNA promoter (Li et al., 2006b). This repositioning is also seen in differentiating adipocytes without protein over expression. Crucially, this repositioning leads to exposure of the CpG site whose methylation is critical for gene silencing (Li et al., 2006b). Thus SNF2h may be acting to move a nucleosome at the rDNA promoter that is protecting a CpG site from methylation. The CpG site is then exposed and can be methylated and this, in conjunction with histone modifications, leads to gene silencing (Li et al., 2006b).

ATRX and SNF2h both appear to have important roles DNA methylation. In both cases, they are required for DNA methylation at specific sequences, but do not appear to have a role in global DNA methylation (Gibbons et al., 2000; Santoro et al., 2002). The SNF2 enzymes DDM1 and LSH have been shown to have a role in both global and locus specific DNA methylation indicating that they may be more general regulators of DNA methylation.

1.4.2 The plant SNF2 protein DDM1 is required for global DNA methylation in *Arabidopsis thaliana*

Screening of DNA methylation mutants of the plant *Arabidopsis thaliana* revealed mutations in a gene termed *DDM1* (decrease in DNA methylation) (Jeddeloh et al., 1999; Vongs et al., 1993). The major phenotype of *DDM1* mutants was global reduction in DNA methylation to around 30% of normal levels. The DNA methylation phenotype is irreversible as backcrossing to a wild-type background did not result in recovery of CpG methylation indicating that *de novo* CpG methylation is a slow process *in vivo* (Vongs et al., 1993) Further characterisation of the mutants showed that DNMT activity and intracellular levels of SAM were not affected, indicating a direct role for DDM1 in DNA methylation (Kakutani et al., 1995). Phenotypically, *DDM1* mutants show morphological abnormalities including defects in leaf structure, flowering time, and flower structure. These phenotypes become more severe in later generations indicating that the loss of DNA methylation is progressively deleterious (Kakutani, 1997; Kakutani et al., 1996). One of the *DDM1* mutants (*ddm1*-2) was mapped to the lower arm of chromosome 5 by chromosome walking (Jeddeloh et al., 1999). DDM1 was subsequently identified as a SNF2 ATPase as all the *DDM1* mutations mapped to the gene at this locus. The mutations all disrupted regions of the core ATPase domain of DDM1 indicating that the ATPase function of this protein is required for DNA methylation (Jeddeloh et al., 1999). Homology within the catalytic domains of DDM1 place it within the ISWI related LSH subfamily of chromatin remodelers (Flaus et al., 2006). The LSH family has homologues in a wide range of organisms including mammals and most vertebrates. Interestingly there is a homologue in *S. cerevisiae* although DNA methylation is not found in this species, perhaps indicative of a divergent role for this protein (Flaus et al., 2006).

DDM1 plays a key role in silencing transcription of transposable elements and its loss of function leads to their mobilization (Miura et al., 2001; Singer et al., 2001). This indicates that a major role of DNA methylation in plants is the suppression of transposition events. DDM1 also plays a role in regulating H3 methylation at heterochromatin (Gendrel et al., 2002). In wildtype plants heterochromatic H3 is

dimethylated at K9 but not at K4. In *ddm1-2* mutants however, H3K9 dimethylation at a number of heterochromatic loci is lost and replaced with H3K4 dimethylation. This leads to transcriptional activation of heterochromatic sequences reinforcing the proposed role of DDM1 in transcriptional silencing (Gendrel et al., 2002; Lippman et al., 2004).

Biochemically, DDM1 has been shown to be a DNA stimulated ATPase that can reposition mononucleosomes from the end of a DNA fragment to the centre *in vitro* (Brzeski and Jerzmanowski, 2003). How this is correlated to the decrease in DNA methylation phenotype is currently unknown but a number of mechanisms have been proposed.

One mechanism relates to the ability of DDM1 to reposition nucleosomes in a way that would directly facilitate access of DNMTs to DNA. This could occur in two ways and depends on the ability of DNMTs to efficiently methylate nucleosomal DNA. The first assumes that specific cytosine residues are protected from methylation by precisely positioned nucleosomes. DDM1 could move these nucleosomes to provide access to the cytosine in a manner analogous to SNF2h at the rDNA promoter (Li et al., 2006b). This mechanism presumes that nucleosomal DNA is less accessible to DNMTs than naked DNA. This is currently not clear with some groups reporting a strong inhibition of DNMT activity on nucleosomal rather than naked DNA *in vitro* and others observing little difference (Gowher et al., 2005b; Robertson et al., 2004; Takeshima et al., 2006). As this is ambiguous, but assuming that large, highly compacted nucleosomal domains would be inaccessible to DNMTs another related mechanism can be proposed. This mechanism involves disruption of the positions of many ordered nucleosomes within a compacted domain to facilitate a more 'open' chromatin structure. This 'open' chromatin would then be generally more accessible for DNMTs to methylate the DNA. Of course, DDM1 may use a combination of these mechanisms to facilitate an 'open' chromatin structure and also to move specific positioned nucleosomes. Alternatively, DDM1 may act in the opposite manner and produce highly compact, transcriptionally silenced chromatin.

This chromatin may then become methylated due to its lack of transcription and associated chromatin modifications.

Another possible mechanism relates to the requirement of H3K9 dimethylation for DNA methylation and the necessity of DDM1 for proper H3K9 dimethylation (Gendrel et al., 2002; Jackson et al., 2002; Lindroth et al., 2001; Tamaru and Selker, 2001). In this model DDM1 is necessary for proper deposition of H3K9 dimethylation, possibly via displacement of non-modified nucleosomes during replication (Gendrel et al., 2002). A lack of H3K9 dimethylation, would in turn lead to loss of DNA methylation. Thus, DDM1 may be indirectly affecting DNA methylation via a related chromatin modification.

1.4.3 Lsh, the mouse homologue of DDM1 has a conserved role in global DNA methylation

The murine homologue of DDM1 has been cloned from fetal thymocytes and termed lymphoid specific helicase (Lsh) or proliferation associated SNF2-like gene (PASG) (Jarvis et al., 1996; Lee et al., 2000) (Figure 1.5). The human homologue has also been cloned and its genomic location mapped to chromosome 10q23-q24 (Geiman et al., 1998). The human gene has been termed helicase lymphoid specific (HELLS). For the purpose of clarity I will subsequently refer to the mammalian Lsh/PASG/HELLS as Lsh. As mentioned previously the ATPase domains of Lsh are most closely related to the ISWI subfamily of SNF2 enzymes (Flaus et al., 2006). Lsh does not however, contain the SANT domain characteristic of the ISWI family of remodelers. Lsh also contains a coiled-coil region that would be expected to facilitate dimerisation or interact with other proteins and a bipartite NLS (Figure 1.5).

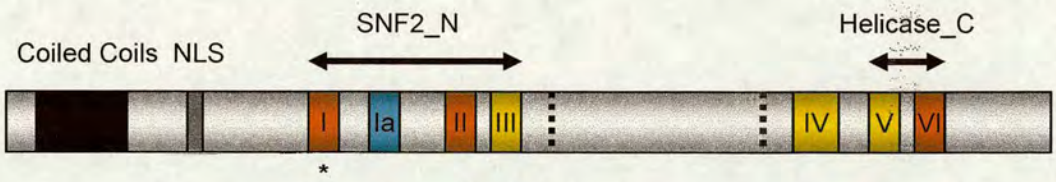


Figure 1.5. Structure of Lsh protein.
Lsh protein structure is illustrated with all relevant domains indicated. The conserved helicase motifs corresponding to the SNF2_N and Helicase_C domains are indicated by roman numerals. The extended linker between these two domains is flanked by dashed lines. Also indicated are the N-terminal coiled-coil domain and the nuclear localisation signal. In several experiments outlined during this thesis a mutation in the ATP binding site of motif I (K254Q) of Lsh is used as a control protein. The site of this mutation is indicated (*).

Initial studies showed that Lsh is predominantly expressed in proliferative tissues such as the testis, bone marrow and lymphoid tissue (Geiman et al., 1998; Jarvis et al., 1996; Lee et al., 2000). In agreement with its lymphoid expression Lsh is required for proliferation of mature T-lymphocytes and its expression appears to correlate with S-phase (Geiman and Muegge, 2000; Yan et al., 2003a). Later studies indicated that Lsh may be ubiquitously expressed, particularly during embryogenesis and development (Geiman et al., 2001).

To further investigate the role of Lsh two independent studies have generated mice carrying targeted deletions of Lsh. These deletions have resulted in functional null ($Lsh^{-/-}$) and hypomorphic alleles ($PASG^{-/-}$) (Dennis et al., 2001; Geiman et al., 2001; Sun et al., 2004).

$Lsh^{-/-}$ mice contain a deletion of exons 6 and 7 that encode the conserved helicase domains I, Ia and part of II. This knockout produced no detectable truncated forms of the protein indicating it is a null allele. Importantly the knockout strategy removes the ATP binding site of Lsh which if produced would be enzymatically inactive. Embryonic development appeared normal but the mice died soon after birth with severe renal lesions (Geiman et al., 2001). Importantly these mice showed substantial

loss of CpG methylation throughout the genome (Dennis et al., 2001). This loss of methylation was widespread affecting repetitive elements such as minor and major satellite repeats, IAP particles, Sine B1 and Line1 elements as well as single copy genes like β-globin, pgk-1, pgk-2 and H19. The amount of methylation lost was quantified by assessing the ability of $Lsh^{-/-}$ DNA to accept methylation. This technique estimated global DNA methylation loss in $Lsh^{-/-}$ mice to be ~60-70% (Geiman et al., 2001). Importantly, the levels and methyltransferase activity of Dnmt1, Dnmt3a and Dnmt3b were unaffected indicating a direct role of Lsh in mammalian DNA methylation.

In an unrelated study Sun and colleagues generated mice carrying a deletion in the PASG (Lsh) gene in exons 10-12 containing helicase motifs II, III and IV (Sun et al., 2004). This deletion resulted in a hypomorphic allele. Mice homozygous for this mutation also show genomic hypomethylation (~43% loss) at the same regions as the $Lsh^{-/-}$ mice. $PASG^{-/-}$ mice are born at just below mendelian frequency and ~60% die just after birth. Interestingly the 40% of $PASG^{-/-}$ mice that survive display low birth weight, growth retardation and a premature aging phenotype. The premature aging phenotype is associated with increased replicative senescence correlated with induction of the tumour suppressor genes $p16^{INK4A}$ and $p19^{ARF}$ (Sun at al., 2004). Embryonic fibroblasts derived from these mice and the $Lsh^{-/-}$ mice also show mitotic abnormalities leading to aneuploidy at high passage numbers (Fan et al., 2003; Sun et al., 2004). Thus two unrelated studies have shown that Lsh is involved and genome wide and locus specific DNA methylation.

Lsh has also been shown to be required for normal histone modification as its absence leads to accumulation of di and tri methyl H3K4 at pericentromeric DNA (Yan et al., 2003b). A concurrent loss of H3K9 methylation at these regions does not occur indicating that, although its role in DNA methylation and H3K4 methylation deposition is conserved between mammals and plants, its role in regulating H3K9 methylation is not (Gendrel et al., 2002; Yan et al., 2003b). $Lsh^{-/-}$ cells also show increased histone acetylation and activated transcription of repetitive elements indicating a role for Lsh in DNA methylation-dependent transcriptional repression

(Huang et al., 2004). The function of Lsh in transcriptional repression is not restricted to repetitive sequences as it is required for imprinting of the *Cdkn1c* gene and silencing of some Hox genes (Fan et al., 2005b; Xi et al., 2007). In these studies Lsh was shown by ChIP to associate with the differentially methylated region (DMR) of the *Cdkn1c* gene and the promoters of several Hox genes. It was also shown to be essential for DNA methylation at these regions and in the case of the Hox promoters for recruitment of Dnmt3b and the polycomb repressive complex -1 (PRC-1) components Bmi1, Mel18, and M33 (Fan et al., 2005b; Xi et al., 2007). Lsh has further been shown to be involved in *de novo* but not maintenance methylation of DNA in embryonic stem (ES) cells and can interact with the *de novo* methyltransferases (Zhu et al., 2006).

As Lsh has no methyltransferase activity and expression of the DNMTs in knockout mice were normal it is possible that Lsh regulates global methylation levels in a manner similar to those proposed for DDM1. Alternatively, Lsh may target DNA methyltransferase enzymes to genomic loci in a sequence specific manner or simply by increasing their affinity to DNA. Another possibility is that Lsh, as a putative SNF2 DNA translocase, may translocate DNA pulling the DNMTs along with it, thus facilitating processive DNA methylation.

Some evidence also indicates that Lsh may have a role in carcinogenesis. Firstly, human LSH maps to chromosomal region 10q23-24, which frequently shows abnormalities in human leukaemia, mycosis fungoides and Sezary syndrome (Geiman et al., 1998; Wain et al., 2005). Secondly, an alternative LSH transcript with a 75bp deletion has been observed in a variety of human acute leukaemia's (Lee et al., 2000). Also a tumour-specific splicing event introducing a premature stop codon into the LSH gene has recently been identified in non-small cell lung cancer (Yano et al., 2004). Thus, it could be envisaged that aberrant gene methylation patterns seen in many cancer types may be partly attributed to defects in targeting of DNA methyltransferases by LSH.

1.5 Project Aims

The aims of this study are to investigate the molecular function of LSH and attempt to relate this to its role in DNA methylation. To address these aims the work described in this thesis uses a variety of approaches to address two key questions. (1) Is LSH an active SNF2 ATPase? (2) How does LSH interact with, and modulate, the DNA methylation machinery?

Chapters three and five address question (1) through characterisation of the enzymatic function of LSH using purified recombinant LSH. The results obtained in these chapters illustrate that although LSH does exhibit a low level of ATPase activity it does not appear to disrupt chromatin structure *in vitro*. Chapters four and five addresses question (2) using a variety of biochemical, and cell biology approaches. Chapter four outlines that, although not stably associated with DNMT enzymes, LSH functionally and physically interacts with them. However, experiments described in chapter five do not identify a key role for LSH in modulating DNMT activity *in vitro*. Together, these studies reveal new insight into how LSH interacts with the DNA methylation machinery.

Chapter two - Materials and Methods

2.1 Materials

Frequently used materials for standard molecular biology techniques are outlined in this section. Additional materials utilised for specific tasks such as protein purification and biochemical assays are referred to in the Methods section. All materials were stored at room temperature unless otherwise stated

2.1.1 Common Buffers

Phosphate-buffered saline (PBS): 140 mM NaCl, 3 mM KCl, 2 mM KH2PO4, 10 mM Na2HPO4.

Tris-glycine-EDTA (TGE): 25mM Tris-HCl pH 8.3, 190mM Glycine, 1mM EDTA

Tris-glycine SDS (TGS): 25 mM Tris, 250 mM Glycine, 0.1 % SDS.

Tris-buffered saline (TBS): 50 mM Tris-HCl pH 8.0, 150 mM NaCl

Tris-acetate EDTA (TAE): 40 mM Tris, 20 mM glacial acetic acid, 1 mM EDTA and pH adjusted to 8.0.

Tris-borate EDTA (TBE): 89 mM Tris, 89 mM Boric Acid, 2 mM EDTA and pH adjusted to 8.0.

Tris-EDTA (TE): 10mM Tris-HCl pH 7.5, 1mM EDTA pH 8.0

2.1.2 Reagents for manipulation of DNA

Bisulphite modification solution: 3.8g sodium hydrogensulphite (this is a mixture of sodium bisulphite, $NaHSO_3$, and sodium metabisulphite, $Na_2S_2O_5$) was dissolved in 5ml dH_2O and 1.5ml 2M NaOH protected from light. 0.11g hydroquinone was dissolved in 1ml dH_2O at 50°C. The sodium bisulphite and the hydroquinone solutions were then mixed to give a solution with a final pH of 5.0. Bisulphite modification solution was prepared immediately prior to use.

DNA isolation buffer: 10mM Tris HCl pH8.0, 10mM EDTA pH8.0.

Orange G loading buffer (6×): 0.198% (w/v) orange G, 12% (w/v) Ficoll, 120mM EDTA pH8.0.

2.1.3 Reagents for manipulation of proteins

Coomassie Brilliant Blue R-250 staining solution: 30% (v/v) methanol, 10% (v/v) glacial acetic acid, 0.1% (w/v) Coomassie Brilliant Blue R-250. Filtered through a Whatman number 1 filter.

ECL solution: 100mM Tris-HCl pH8.8, 1.25mM luminal, 0.2mM p-coumaric acid, 0.04% H_2O_2

NE1: 20mM Hepes pH7.0, 10mM KCl, 1mM $MgCl_2$, 0.1% (v/v) Triton X-100, 20% (v/v) glycerol, 0.5mM DTT and complete protease inhibitors (Sigma). DTT, and complete protease inhibitors were added immediately prior to use. Stored at 4°C.

NE2: 20mM Hepes pH7.0, 420mM NaCl, 10mM KCl, 1mM $MgCl_2$, 0.1% (v/v) Triton X-100, 20% (v/v) glycerol, 0.5mM DTT and complete protease inhibitors (Sigma). DTT, and complete protease inhibitors were added immediately prior to use. Stored at 4°C.

Ponceau S staining solution: 1% (v/v) glacial acetic acid, 0.5% (w/v) Ponceau S

SDS PAGE loading buffer (5×): 225mM Tris-HCl pH6.8, 50% glycerol, 5% SDS, 0.05% bromophenol blue, 250mM DTT. Stored at room temperature, DTT added just prior to use

SDS PAGE separating gel: 7-15% (w/v) 29:1 acrylamide:bis-acrylamide, 0.1% (w/v) SDS, 390mM Tris HCl pH8.8, 0.08% (v/v) TEMED, 0.1% (w/v) APS. Prepared immediately prior to use.

SDS PAGE stacking gel: 5% (w/v) 29:1 acrylamide:bis-acrylamide, 0.1% (w/v) SDS, 129mM Tris HCl pH6.8, 0.1% (v/v) TEMED, 0.1% (w/v) APS. Prepared immediately prior to use.

Transfer buffer: 25 mM Tris, 250 mM Glycine.

2.1.4 Bacterial Strains

DH5α: *supE*44 Δ*lac*U169 (ϕ80 *lacZ*ΔM15) *hsdR*17 *recA*1 *endA*1 *gyrA96 thi*-1 *relA*1. Laboratory stock.

Nova-blue: *endA1 hsdR17* (r_{K12-} m_{K12+}) *supE44 thi-1 recA1 gyrA96 relA1 lac* F′ [*proA+B+ lacI^{q}ZΔM15*::Tn*10*]. Laboratory stock.

ER2566: F– λ– fhuA2 [lon] ompT lacZ::T7 gene1 gal sulA11 Δ(mcrC-mrr)114::IS10 R(mcr-73::miniTn10–TetS)2 R(zgb-210::Tn10)(TetS) endA1 [dcm].

2.1.5 Bacterial cell culture media and related reagents

Ampicillin stock solution: 50mg/ml ampicillin in dH_2O. 0.2μm filter sterilised and stored at –20°C. Added to LB medium to 100μg/ml.

Blue/white selection LB agar plates: 40μl 100mM IPTG, and 40μl 40mg/ml X-gal, was spread over LB agar plates containing the appropriate antibiotic(s). The plates were dried at r/t before being used the same day.

Competent cell buffer: 100mM MgCl2, 70mM $CaCl_2$. 0.2μm filter sterilised and stored at 4°C.

Gentamicin: Purchased from Invitrogen (SKU# 15750-037) as a stock concentration of 50mg/ml and stored at 4 °C. Added to LB medium to7μg/ml.

IPTG stock solution: 1M IPTG in dH_2O. 0.2μm filter sterilised and stored at −20°C.

Kanamycin stock solution: 50mg/ml kanamycin in dH_2O. 0.2μm filter sterilised and stored at −20°C. Added to LB medium to 50μg/ml.

LB medium: 10g/l Bacto tryptone (Difco), 5g/l Bacto yeast extract (Difco), 10g/l NaCl. pH adjusted to 7.0 with NaOH. 15g/l Bacto agar (Difco) added if making LB agar then autoclaved. LB agar plates (20ml volume) were stored inverted at 4°C and LB broth was stored at r/t.

Tetracycline: 10mg/ml in ethanol. Stored at -20°C. Added to LB medium to 10μg/ml.

X-gal stock solution: 40mg/ml X-gal in dimethylformamide (DMF). Stored at −20°C protected from light.

2.1.6 Yeast strains

AH109: *MATa, trp1-901, leu2-3, 112, ura3-52, his3-200, gal4Δ, gal80Δ, LYS2 : : GAL1*$_{UAS}$-*GAL1*$_{TATA}$-*HIS3, GAL2*$_{UAS}$- *GAL2*$_{TATA}$-*ADE2 URA3 : : MEL1*$_{UAS}$- *MEL1* $_{TATA}$-*LacZ MEL1*. Clontech.

Y187: MATα, ura3-52, his3-200, ade2-101, trp1-901, leu2-3, 112, gal4Δ, gal80Δ, met–, URA3 : : $GAL1_{UAS}$-$GAL1_{TATA}$-LacZ MEL1. Clontech.

2.1.7 Yeast cell culture media and related reagents

Dropout supplements (10×): Dropout supplements are components of minimal yeast media, and contain a specific mixture of amino acids and nucleosides (one or more constituents may be missing from a particular dropout supplement, thereby enabling selection of yeasts with a particular phenotype). Dropout supplements (Sigma) were dissolved in dH_2O to 10× concentration, then 0.2μm filter sterilised and stored at 4°C.

PEG/LiAc: 40% (w/v) PEG 3,350, 100mM LiOAc, 10mM Tris HCl pH7.5, 1mM EDTA. 0.2μm filter sterilised.

SD medium: A minimal yeast medium used to select for specific phenotypes. SD medium is prepared by combining minimal SD base with a specific dropout supplement. 27g/l minimal SD base (BD Biosciences) was dissolved and made up to 900ml with dH2O. 20g/l BactoAgar added if making SD agar then autoclaved and cooled to 55°C. 10X dropout supplement (warmed to 55°C) added to 1X. SD agar plates were stored inverted at 4°C and SD broth was stored at r/t.

TE/LiAc: 100mM LiOAc, 10mM Tris HCl pH7.5, 1mM EDTA. 0.2μm filter sterilised.

YPDA medium: Complete yeast medium. 10g/l yeast extract, 20g/l peptone, 0.003% (w/v) adenine hemisulphate made up to 900ml with dH2O. 20g/l BactoAgar added if making YPDA agar then autoclaved and cooled to 55°C. 20% (w/v) glucose (warmed to 55°C) added to 2%. YPDA agar plates were stored inverted at 4°C and YPDA broth was stored at r/t.

2.1.8 Mammalian cell lines

HCT116: Human colon cancer cell line obtained from Vogelstein laboratory. Grown in RPMI medium containing 10% (v/v) FBS, non essential amino acids (Invitrogen), Sodium Pyruvate (Invitrogen), supplemented with a mix of 100u/mL penicillin, 100μg/ml of a streptomycin and 2mM L-glutamine (Sigma). DNMT1 (DNMT1-/-),

DNMT3B (DNMT3B-/-), DNMT1 and DNMT3B (DKO) and DNMT1+/-, DNMT3A and DNMT3B (TKO) derivative cell lines obtained from Vogelstein and Scheubel laboratories (Jair et al., 2006; Rhee et al., 2002; Rhee et al., 2000).

H226: Human lung squamous carcinoma cell line (laboratory stock). Grown in RPMI medium containing 10% (v/v) FBS, supplemented with a mix of 100u/mL penicillin, 100µg/ml of a streptomycin and 2mM L-glutamine (Sigma).

H520: Human non-small cell lung carcinoma cell lines (laboratory stock). Grown in RPMI medium containing 10% (v/v) FBS, supplemented with a mix of 100u/mL penicillin, 100µg/ml of a streptomycin and 2mM L-glutamine (Sigma).

HeLa: Human epithelial carcinoma cell line (laboratory stock). Grown in DMEM medium containing 10% (v/v) FBS, supplemented with a mix of 100u/mL penicillin, 100µg/ml of a streptomycin and 2mM L-glutamine (Sigma).

Lsh-/- mouse embryonic fibroblasts: Mouse embryonic fibroblasts genetically deficient for Lsh obtained from the Muegge lab (Dennis et al., 2001). Grown in DMEM medium containing 10% (v/v) FBS, supplemented with a mix of 100u/mL penicillin, 100µg/ml of a streptomycin and 2mM L-glutamine (Sigma).

MRC5: Derived from normal lung tissue of a 14-week-old male fetus. Grown in DMEM medium containing 10% (v/v) FBS, supplemented with a mix of 100u/mL penicillin, 100µg/ml of a streptomycin and 2mM L-glutamine (Sigma).

Mouse embryonic fibroblasts: Mouse embryonic fibroblasts strain matched to those genetically defiecient for Lsh obtained from the Muegge lab (Dennis et al., 2001). Grown in DMEM medium containing 10% (v/v) FBS, supplemented with a mix of 100u/mL penicillin, 100µg/ml of a streptomycin and 2mM L-glutamine (Sigma).

Mouse embryonic stem cells: Embryonic stem (ES) cells are derived from the inner cell mass of a blastocyst embryo (laboratory stock). Grown in DMEM medium containing 10% (v/v) FBS, supplemented with leukemia inhibitory factor and a mix of 100u/mL penicillin, 100µg/ml of a streptomycin and 2mM L-glutamine (Sigma).

VA13: The VA13 cell line was established from the fetal human diploid fibroblast cell line WI-38 after transformation with simian virus 40 (SV40) (laboratory stock). Grown in DMEM medium containing 10% (v/v) FBS, supplemented with a mix of 100u/mL penicillin, 100µg/ml of a streptomycin and 2mM L-glutamine (Sigma).

2.1.9 Mammalian cell culture media

PBS: 140mM NaCl, 3mM KCl, 2mM KH2PO4, 10mM Na2HPO4 (Gibco)

Trypsin-EDTA: 10× trypsin-EDTA (Gibco) diluted in PBS to 0.25% (w/v) trypsin, 1mM EDTA. Stored at −20°C.

2.1.10 Insect cell lines and culture media

Sf9: All baculovirus stock generation and expression was done using SF9 cells (Invitrogen) which are derived from the papal ovarian tissue of the fall army worm Spodoptera frugiperda. Cells were grown in serum free SF900 II medium containing L-glutamine (Gibco) supplemented with penicillin and streptomycin (100μg/ml), and incubated at 27°C.

2.1.11 Oligonucleotides

Custom oligonucleotides were purchased from MWG. Lyophilised oligonucleotides were resuspended in dH_2O to 100μM and stored at –20°C.

2.1.12 PCR primers

Table 2.1: PCR primers used in this thesis

Name	Sequence (5′-3′)	Description / use
M13 F	TGTAAAACGACGGCCAGT	M13 forward primer for PCR amplification of bacmid DNA
M13 R	GTTTTCCCAGTCACGAC	M13 reverse primer for PCR amplification of bacmid DNA
SP6	CTAGCATTTAGGTGACACTA	Sequencing
T7	AATACGACTCACTATAGGGA	Sequencing
LSH NcoI/F	CCATGGATGCCAGCGGAACG GCCCGC	Cloning LSH into Nco1/Not1 sites of pFAST-BAC HTA
LSH NotI/R	GCGGCCGCAAACAAACATTC AGGACTGGAATC	Cloning LSH into Nco1/Not1 sites of pFAST-BAC HTA
LSH NcoI/F	CCATGGATGCCAGCGGGACG GCCCGC	Cloning LSH into NcoI/SmaI sites of pGBKT7

Name	Sequence (5′-3′)	Description / use
LSH StuI/R	AGGCCTAAACAAACATTCAG GACTGGAATC	Cloning LSH into NcoI/SmaI sites of pGBKT7
LSH 1-226 StuI/R	AGGCCTGTACCATCGCATCA CTCCTC	Cloning LSH(1-226) into NcoI/SmaI sites of pGBKT7
LSH 227-838 NcoI/F	CCATGGGGTACCAAGTAGAA GGCATGG	Cloning LSH(1-226) into NcoI/SmaI sites of pGBKT7
Lsh 1-333 StuI/R	AGGCCTCAAATGACGTGATT ACCACA	Cloning LSH(1-333) into NcoI/SmaI sites of pGBKT7
G4-LSH NheI/F	GCTAGCATGAAGCTACTGTC TTCTATCGA	Cloning GAL4BD-LSH into the Nhe1/ApaI sites of pCDNA3.1
G4-LSH ApaI/R	GGGCCCAAACAAACATTCAG GACTGGAAT	Cloning GAL4BD-LSH into the Nhe1/ApaI sites of pCDNA3.1
DNMT3B 1200 BamHI/F	GGATCCATGAAGGGAGACAC CAGGCA	Cloning 1-1200bp of hDNMT3B into BamHI/SalI sites of pGEX-4T1
DNMT3B 1200 SalI/R	GTCGACCATTTGTTCTCGGCT CTGAT	Cloning 1-1200bp of hDNMT3B into BamHI/SalI sites of pGEX-4T1
DNMT3B 2502 BamHI/F	GGATCCGCTTCAGATGTTGC CAACAA	Cloning 1201-2502bp of hDNMT3B into BamHI/SalI sites of pGEX-4T1
DNMT3B 2502 SalI/R	GTCGACCTATTCACATGCAA AGTAGTCC	Cloning 1201-2502bp of hDNMT3B into BamHI/SalI sites of pGEX-4T1
DNMT1 751 EcoRI/F	GAATTCTGCGGGCAGTACCT GGACGA	Cloning 751-3375bp of hDNMT1 into EcoRI/NotI sites of pGEX-4T1
DNMT1 751 NotI/R	GCGGCCGCATGCCGCTGAAG CCCTGGCA	Cloning 751-3375bp of hDNMT1 into EcoRI/NotI sites of pGEX-4T1

Name	Sequence (5′-3′)	Description / use
DNMT3B 640C/S F	ATTGGCGGAAGCCCATCCAA CGATCTCTCAAATG	Primers for C/S mutagenisis of hDNMT3B C640 (catalytic site)
DNMT3B 640C/S R	CATTTGAGAGATCGTTGGAT GGGCTTCCGCCAAT	Primers for C/S mutagenisis of hDNMT3B C640 (catalytic site)
LSH S503A F	GAAACAATTGAGTTA GCT CCTACTGGTCGACC	Point mutagenesis of serine 503 to alanine in LSH
LSH S503A R	GGTCGACCAGTAGG AGC TAACTCAATTGTTTC	Point mutagenesis of serine 503 to alanine in LSH
LSH S503D F	GAAACAATTGAGTTA GAT CCTACTGGTCGACC	Point mutagenesis of serine 503 to aspartic acid in LSH
LSH S503D R	GGTCGACCAGTAGG ATC TAACTCAATTGTTTC	Point mutagenesis of serine 503 to aspartic acid in LSH
LSH Not1/F	ATTAGCGGCCGCGATGCCAG CGGAACGGCCCGC	For cloning LSH into p3XFLAG-CMV-10
LSH Xma1/R	ATTACCCGGGAAACAAACAT TCAGGACTGGAATC	For cloning LSH into p3XFLAG-CMV-10
pFastBAC His Off F	CCGAAACCATGTCGTACTAC GATTACGATATCCCAACGAC C	Removing 6xHis tag from pfastbac vectors
FastBac His Off R	GGTCGTTGGGATATCGTAAT CGTAGTACGACATGGTTTCG G	Removing 6xHis tag from pfastbac vectors
54A18 F	TATGTAAATGCTTATGTAAA	Amplification of MMTV NucA
54A18 R	TACATCTAGAAAAAGGAGCA	Amplification of MMTV NucA
54A54 R	AGCACATGTGAAAGTTAAAA	Amplification of MMTV NucA
0A0 F	ACTTGCAACAGTCCTAACAT	Amplification of MMTV NucA
0A0 R	CAAAAAACTGTGCCGCAGTC	Amplification of MMTV NucA
67A0 F	TGTTCTATTTTCCTATGTTC	Amplification of MMTV NucA
TK-prom F	AATTGCTCAACAGTATGAAC ATTTC	Amplification of TK promoter in ChIP experiments

Name	Sequence (5′-3′)	Description / use
TK-prom R	CAATTGTTTTGTCACGATCAA AGGA	Amplification of TK promoter in ChIP experiments
GAPDH F	GAGGCTGTGAGCTGGCTGTC	Amplification of control GAPDH promoter in ChIP experiments
GAPDH R	CAGAGCAGAGTAGCAAGAG CAAGG	Amplification of control GAPDH promoter in ChIP experiments
pTK/bis/F	TGAGTTAGTTTATTTATTAGG TATT	Bisulphite sequencing of TK promoter
pTK/bis/R	CTATAATTTATATTCAACCCA TATC	Bisulphite sequencing of TK promoter
4+5Gs/F	CGGGTGGGGGTGCGGAGCGC TCTCGCTTTGGGGCTG	DNA probe containing runs of Gs used for EMSA experiments
4+5Gs/R	CAGCCCCAAAGCGAGAGCGC TCCGCACCCCCACCCG	DNA probe containing runs of Gs used for EMSA experiments
0Gs/F	GCAGATGAATGACTGATATT ACTCAGAAGTTTCATG	DNA probe without runs of Gs used for EMSA experiments
0Gs/R	CATGAAACTTCTGAGTAATA TCAGTCATTCATCTGC	DNA probe without runs of Gs used for EMSA experiments

2.1.13 Plasmids

Table 2.2: Plasmids used in this thesis

Name	Source	Selectable marker	Description
pGEX 4T	GE Healthcare	Amp	Bacterial expression vector, GST fusion
pFAST-BAC	Invitrogen	Amp	Shuttle vector for production of recombinant BACMID
pGEM T-Easy	Promega	Amp	PCR cloning vector
pGBKT7	Clonetech	Kan	Yeast 2-hybrid bait vector

Name	Source	Selectable marker	Description
pGADT7	Clonetech	Amp	Yeast 2-hybrid activator vector
p3xFlag	Sigma	Amp	Mammalian expression vector, N-terminal 3xFlag tag
GAL4-TK-LUC	A. Bird	Amp	Luciferase gene driven by TK promoter upstream of 5xGAL4 binding sites
pact-βgeo	A. Bird	Amp	β-geo gene under control of β-actin promoter
pCDNA3.1 Hygro+	Invitrogen	Amp	Mammalian expression vector
pFAST-BAC DDM1	A. Jerzmanowski	Amp	Shuttle vector for production of recombinant BACMID
pReceiver I01 DNMT3B	RZPD	Amp	Shuttle vector for production of recombinant BACMID
pFAST-BAC DNMT1		Amp	Shuttle vector for production of recombinant BACMID
pReceiver M03 DNMT3B-GFP	RZPD	Amp	DNMT3B-GFP fusion for mammalian expression
pEGFP-C1 Dnmt1-GFP	H. Leonhardt	Amp	Dnmt1-GFP fusion for mammalian expression
pCDNA3.1 Hygro+ LSH	R.J. Arceci	Amp	Full length, human LSH cloned into a mammalian expression vector
pCDNA3.1 Hygro+ LSH^{K254Q}	R.J. Arceci	Amp	Full length, human LSH^{K254Q} cloned into a mammalian expression vector

2.1.14 Antibodies

Table 2.3: Primary antibodies used in this thesis

Name	Source, catalogue number	Type	Dilution for Western blot
α-DNMT1	NEB (M0231L)	Rabbit polyclonal	1:1,000
α-FLAG M2	Sigma (F1804)	Mouse monoclonal	1:500
α-HA	CRUK (12CA5)	Mouse monoclonal	1:500
α-HIS	Santa Cruz Biotechnology (sc-803)	Rabbit polyclonal	1:1,000
α-LSH	Santa Cruz Biotechnology (sc-46665)	Mouse monoclonal	1:500
α-Myc	CRUK (9E10)	Mouse monoclonal	1:500
α-DNMT1	Santa Cruz Biotechnology (sc-20701)	Rabbit polyclonal	1:500
α-DNMT3B	Affinity Bioreagents (PA1-884)	Rabbit polyclonal	1:1,000
α-HDAC1	Santa Cruz Biotechnology (sc-7872)	Rabbit polyclonal	1:500
α-HDAC2	Santa Cruz Biotechnology (sc-7899)	Rabbit polyclonal	1:500
α-GAL4BD	Santa Cruz Biotechnology (sc-510)	Mouse monoclonal	1:500
α-Brg1	Santa Cruz Biotechnology (sc-10786)	Rabbit polyclonal	1:500
α-H3	Upstate (05-499)	Rabbit polyclonal	NA
α-H4	Upstate (07-108)	Rabbit polyclonal	NA
α-H3K9Ac	Upstate (07-352)	Rabbit polyclonal	NA
α-H4K12Ac	Upstate (07-353)	Rabbit polyclonal	NA

Table 2.4: Secondary antibodies used in this thesis

Name	Source, catalogue number	Dilution for western blot
α-goat HRP	Sigma (G4-34)	1:2000

Name	Source, catalogue number	Dilution for western blot
α-mouse HRP	Sigma (A6782)	1:2000
α-rabbit HRP	Sigma (A0545)	1:2000

2.1.15 Chromatogrpahy

Chromatography media: All chromatography media was purchased from GE Healthcare. Use of Superose 6 and Superose 12 columns was kindly provided by Prof Adrian Bird.

Chelating Sepharose affinity chromatography (bound with Co)

Mono Q anion exchange resin (mono dispersed beads)

Superose 12 Size exclusion chromatography (24 ml analytical)

Superose 6 Size exclusion chromatography (24 ml analytical)

Chromatography solutions:

Size exclusion chromatography buffer (GF300) consisted of 20 mM HepesKOH (pH 7.9), 3mM MgCl2, 10 % glycerol and 300 mM KCl.

2.2 Methods

2.2.1 DNA manipulation

Small scale plasmid DNA preparation: Small scale preparation of plasmid DNA was carried out according to the Qiagen Miniprep Kit instruction manual. DNA was eluted in 50 µl 10mM Tris pH8.8 and stored at -20 °C.

Large scale plasmid DNA preparation: Large scale preparation of plasmid DNA was done according to the Qiagen Maxiprep Kit instruction manual. Precipitated DNA was resuspended in 10 mM Tris pH 8.0 and stored at -20 °C.

Large scale preparation of Bacmid DNA: Large scale preparation of Bacmid DNA was done according to the Qiagen Midiprep Kit. The Bacmid DNA pellet was resuspended in 10 mM Tris pH 8.0 and stored in 5 µg aliquots at - 20 °C.

Extraction of mammalian genomic DNA: For extraction of mammalian genomic DNA, 1 x 10^6 cells were resuspended in 500µl of DNA isolation buffer. The cells were treated with 100µg/ml RNase A for 15 minutes at 37°C. Following RNase treatment, SDS and Proteinase K were added to 0.5% (w/v) and 100µg/ml respectively and incubated overnight at 65°C. Digested peptides were removed by 3 extractions with phenol:chloroform:isoamyl alcohol and a single extraction with chloroform. DNA was precipitated by addition of 0.7 volumes isopropanol and centrifuged at 13000rpm for 15 minutes at 4°C. The DNA was washed twice with 70% ethanol, allowed to dry and resuspended in 100µl of TE.

Restriction endonuclease digestion of DNA: Restriction digest were generally done in 30 µl volumes in the appropriate buffers supplied with the restriction enzyme (Fermentas). Most digests were carried out at 37 °C for 1h, or at the temperature given in the product literature. Efficiency and completeness of digestion was verified by agarose gel electrophoresis in 1 X TAE and visualization by ethidium bromide staining.

Dephosphorylation of DNA fragments: The 5' phosphate of DNA molecules was removed with Antarctic alkaline phosphatase (NEB). Typically, 1μg of DNA dephosphorylated by 10 units of enzyme in 30μl volume. The reaction was allowed to proceed for 1 hr at 37°C, the reaction stopped by incubation at 65°C for 5 minutes and the DNA re-purified on a PCR purification column (Qiagen).

Methylation of DNA fragments: Methylation reactions using M.*HhaI*, or M.*SssI* methyltransferases (NEB) were carried out according to manufactures instructions.

Polymerase chain reaction (PCR): Polymerase chain reactions were carried out in 25 μl or 50 μl reaction volumes using a Biometra T3 Thermocycler. Generally 0.5 μM of each primer was used per reaction. If template DNA was plasmid 5 ng of DNA was used and if genomic DNA 100 ng was used. Cycling parameters varied depending on DNA source and primer properties. If the PCR was for analytical purposes Taq polymerase (Roche) was used to amplify the DNA. If the PCR products were subsequently to be cloned for protein expression 1u PFU polymerase (Fermentas) / 50u Taq was used to amplify the DNA. In each case the reaction buffer and salt solution provided with these commercial polymerases were used.

Purification of PCR fragments: PCR fragments were purified with the Qiagen PCR purification kit according to the manufacturer's protocol.

Molecular cloning of PCR products: PCR products were cloned using the pGEM-T Easy Vector system (Promega) according to the manufacturers instructions

DNA Sequencing: Sequencing reactions were assembled using the Big Dye terminator V3 kit (Roche). Reactions were assembled in 10 μl volumes containing 2 μl of Big Dye sequencing mix, 0.3 μM of sequencing primer, and 500 ng of template DNA. The sequencing reactions used the following program: 96 °C for 1 minute followed by 25 cycles at 96 °C for 10 seconds, 50 °C for 5 seconds, and 60 °C for 4

min. Sequences were analyzed by the sequencing facility at the University of Edinburgh.

Agarose gel extraction of DNA fragments: DNA was extracted from agarose gels using the Qiagen gel extraction kit according to the manufacturer's instructions and the DNA eluted in 30 µl of TE heated to 55 °C.

Ligations: Ligation reactions were carried out in a 20 µl reaction volume using 10 units of T4 DNA ligase (NEB) according to the manufactures recommendations.

Site directed mutagenesis: Site directed mutagenesis was carried out with the QuikchangeXL mutagenesis kit (Stratagene) according to the manufacturers recommendations.

Radioactive labelling DNA fragments: 3 µg of template DNA was end labelled in a 50 µl reaction volume with 10 units of T4 polynucleotide kinase (NEB) and 1 µl of 10 mCi / ml of gamma-^{32}P dATP for 1 hour at 37 °C. Unincorporated isotope was removed by purification of the labelled DNA using illustra Micrcospin G25 Columns (GE Healthcare).

Bisulphite treatment of DNA: Genomic DNA was isolated from transfected HCT116 cells and digested with a restriction endonuclease that cuts out with the region of DNA being analysed. The digested DNA was purified using the QIAquick PCR Purification kit (QIAGEN) according to the manufacturer's instructions. 1µg digested DNA in 25µl TE was denatured by incubation at 100°C for 5min, and the subsequent addition of 2.5µl 3M NaOH and incubation at 37°C for 20min. 270µl Bisulphite modification solution was added to the denatured DNA, and the sample was covered with a mineral oil overlay and incubated at 55°C overnight protected from light. The modified DNA was then isopropanol precipitated (in the presence of 50µg glycogen carrier) and resuspended in 25µl TE pH7.4. The DNA was then desulphonated by the addition of 2.5µl 3M NaOH and incubation at 37°C for 15min.

The bisulphite-modified DNA was purified using the QIAquick PCR Purification kit (QIAGEN) and eluted in 30µl 10mM Tris HCl pH8.5.

5µl of the bisulphite-modified DNA was used in a PCR reaction containing primers specific to the bisulphite-modified TK-promoter sequence, and the whole PCR resolved by agarose gel electrophoresis. The amplified DNA was purified from the gel using the QIAquick Gel Extraction kit (QIAGEN), and was cloned into the pGEM-T Easy plasmid using the pGEM-T Easy cloning kit (Promega) according to the manufacturer's instructions. The cloning reactions were transformed into Nova-blue competent cells and plated onto LB agar plates containing ampicillin, IPTG and X-gal.

Bacterial colonies transformed with an insert-containing plasmid were identified by blue/white selection and the plasmid DNA isolated by Qiagen Miniprep. Analytical digests with EcoRI identified true positive clones containing insert of the expected size. 500ng of insert containing plasmid was used in a BigDye sequencing reaction with T7 and SP6 sequencing primers.

2.2.2 Protein manipulation

Measuring protein concentration: Protein concentration was measured using the BCA Protein Assay (Pierce Biotechnology). Briefly, stock reagents A and B were mixed together at a 50:1 ratio and increasing amounts of protein sample added to 1 ml of the mixture. The reagent was incubated at 65 °C for 15 minutes after which absorbance at 562nm was measured. Known concentrations of BSA were added to the BCA reagent in parallel to generate a standard curve by which the concentration of the test protein sample could be determined.

SDS-PAGE: Proteins were diluted in SDS-PAGE loading buffer and boiled for 5 min at 100 °C. 0.75 mm thick gels were assembled in a Bio-Rad Mini-PROTEAN 3 apparatus. Gels consisted of a stacking gel buffered at pH 6.8 and a separating gel at pH 8.8. All gels used an acrylamide to bis-acrylamide ratio of 29:1 with 0.1% SDS. The acrylamide was polymerized by the addition of ammonium persulfate and

TEMED as a catalyst. Gels were run in TGS at 270V until the loading dye migrated off the bottom of the gel.

Coomassie blue stain: SDS PAGE gels were incubated with fresh Coomassie Brilliant Blue R-250 staining solution, heated briefly in the microwave and rocked for one hour. Excess stain was removed with H2O and the gel was immersed in de-stain solution consisting of 30 % methanol and 10% acetic acid. The de-stain was allowed to proceed for several hours until the background staining was no longer visible. Gels were then soaked overnight in H2O, dried on Whatman paper and imaged.

Silver stain: SDS-PAGE gels were silver stained using the SilverQuest Silver Staining Kit (Invitrogen), using the Basic Staining Protocol, according to the manufacturer's guidelines.

Wet transfer to nitrocellulose membrane: Wet transfer was carried out on a Bio-Rad Mini Trans-Blot Electrophoretic Transfer Cell according to the manufacturer's recommendations. Briefly, one layer of 0.3 mm Whatman paper was soaked in transfer buffer and the SDS-PAGE gel placed on it. A single layer of nitrocellulose filter paper was pre-wetted in dH_20, soaked in transfer buffer and placed on top of the gel. An additional layer of 0.3 mm Whatman paper pre-soaked in transfer buffer was placed on top on the gel. Bubbles were removed by rolling a 15 ml falcon tube several times over the assembled gel sandwich. The gel sandwich was assembled into the gel holder cassette between two fiber pads also pre-soaked in transfer buffer. The assembled cassette was inserted into the electrode module and transferred with 270V for 1 hour. After transfer the gel sandwich was disassembled and the membrane stained with Ponceau stain to ensure efficient transfer.

Western blotting: For most antibodies the following Western blot procedure was used. After a 1 hour block in a 2% milk, 0.1% Tween, 1 X TBS solution, fresh blocking solution containing a 1:500 to 1:1000 dilution of the primary antibody was applied. The primary antibody was incubated overnight at 4 °C followed by three

consecutive twenty minute washes in PBS / 0.1% Tween. The membrane was blocked again for one hour at room temperature in 2% milk, 0.1% Tween, 1 X TBS. The secondary antibody was applied at a dilution of 1:2000 in the same solution for three hours at room temperature. Four consecutive fifteen minute washes in PBS / 0.1% Tween removed unbound secondary antibody. The nitrocellulose membrane was incubated with 10ml ECL solution for 1 minute then wrapped in Saran wrap and exposed to RPNew Medical X-Ray film (CEA).

Immunoprecipitation of proteins: 100µg of nuclear extract was used for immunoprecipitations experiments. Nuclear extract was diluted to 150mM NaCl with NE1 buffer and applied to Protein G Sepharose 4 fast flow beads (GE Healthcare) for 1 hour at 4°C to remove proteins that bind to them non-specifically. 1µg of antibody was added to the pre cleared nuclear extract and incubated at 4°C for 1 hour on a spinning fly wheel. The antibody was captured for 1 hour at 4°C with 25µl of Protein G beads pre washed with NE1 buffer supplemented with 150mM NaCl. Following capture the beads were collected by centrifugation at 6000g for 1 minute at 4°C. The beads were washed 3 times with 500µl of NE1 buffer containing 150mM NaCl. After each wash the beads were collected by centrifugation at 6000g for 1 minute at 4°C. After the final wash 20 µl of 2 X SDS PAGE loading buffer was added to the beads and they were boiled at 100°C for 10 minutes to liberate the proteins. 10µl of the immunoprecipitated sample was subsequently analysed by SDS-PAGE.

Chromatin Immunoprecipitation: Chromatin immunoprecipitations experiments were performed to assess the impact of LSH binding to the TK promoter. 5 x 10^6 HEK293 cells containing a stably integrated 5x GAL4-TK-LUC construct (Ishizuka and Lazar, 2003) were transiently transfected with 10 µg of GAL4-LSH 226 or 10 µg of pCDNA3.1 GAL4BD control plasmid. At 48 h post transfection, the cells were recovered by trypsin treatment and cross-linked with 1% formaldehyde. Cross linking was quenched by addition of glycine to a final concentration of 125mM. Crosslinked cells were washed twice with ice cold PBS + complete protease inhibitors (Sigma) by centrifugation at 2000g for 10 minutes at 4°C. The cells were then resuspended in a buffer containing 10mM Tris-HCl pH8.0, 200mM NaCl, 1mM

EDTA and 0.5mM EGTA + complete protease inhibitors (Sigma) and incubated at room temperature for 10 minutes. The cells were centrifuged as above and resuspended in TE + complete protease inhibitors (Sigma) at a concentration of 5 x 10^6 per ml. 500µl aliquots of cells were sonicated in a SoniPrep sonifier six times for 20 seconds at power 5 microns. Sonication was assessed by phenol-chloroform extraction of 10 µl of sonicated material. Chromatin sonicated to ~500bp was used for subsequent immunoprecipitations. Chromatin concentration was measured at OD_{280}. 25µg of chromatin was immunoprecipitated for 2 hours at 4°C on a spinning fly wheel with 0. 25µg of the following antibodies: anti-HA control (CRUK), anti-H3 (Upstate; 05-499, lot 25198), anti-H3K9ac (Upstate; 07-352, lot 28741), anti-H4 (Upstate; 07-108, lot 25296), and anti-H4K12ac (Upstate; 07-595, lot 28885). Immunoprecipitated complexes were captured on protein G Dynabeads (Invitrogen) for 1 hour at 4°C on a spinning fly wheel. Dynabeads were captured using a magnetic strip and washed 10 times with RIPA buffer (50mM Hepes, pH8.0, 1mM EDTA, 1% NP40, 0.7% DOC, 500mM LiCl). Following the final wash all the liquid was removed and DNA eluted from the beads by overnight incubation at 65°C with 100µl TE supplemented with 100µg/ml Proteinase K (Sigma) and 0.5% SDS. The next day, the beads were extracted twice with phenol:chloroform and once with chloroform. The DNA was precipitated with 300µl isopropanol, 150µl ammonium acetate with 5µg of glycine as a carrier. The precipitated DNA was washed with 70% ethanol and resuspended in 15µl of dH2O. 1 µl of chromatin-immunoprecipitated DNA was used in 30 cycles of PCR.

Purification of 3xFlag tagged proteins from HeLa cells: For purification of 3xFlag tagged proteins 5 x 106 HeLa cells were transfected with 10 µg of expression plasmid using JetPEI transfection reagent (Section 2.2.6). 48 hour after transfection cells were harvested by trypsin/EDTA treatment and washed twice with ice cold PBS + complete protease inhibitors (Sigma). Nuclear extract was obtained with 500 µl NE2 buffer (Section 2.2.6) and applied to 100µl of anti-Flag M2 affinity gel (Sigma) that had been pre-washed with 1ml 0.1M Glycine pH3.5 and 3 x 1ml TBS. The beads were mixed with the nuclear extract for 2 hours at 4°C on a spinning fly wheel. Beads were then collected by centrifugation for 1 minute at 6000g and supernatant

removed. The beads were washed 4 times with 1ml of buffer containing 20mM Hepes pH7.0, 500mM NaCl, 10mM KCl, 1mM MgCl2, 0.1% (v/v) Triton X-100, 20% (v/v) glycerol, 0.5mM DTT and complete protease inhibitors (Sigma). Washes were carried out for 5 minutes each at 4°C on a spinning fly wheel. The beads were then further washed three times with 1ml TBS. Elution of Flag fusion proteins was carried out using an elution buffer consisting of TBS with 300 µg/ml 3xFlag peptide and complete protease inhibitors (Sigma). 30 µl elution buffer was added to the washed and left for 30 minutes at room temperature on a spinning fly wheel. The beads were collected by centrifugation for 1 minute at 6000g and supernatant removed. Elution was repeated a further 3 times. Efficiency of elution was analysed by SDS-PAGE followed by silver staining.

***In vitro* translation of peptides:** *In vitro* translation of peptide was carried out using the TnT T7 Quick for PCR DNA kit (Promega) according to the manufacturer's instructions. Typically, 500ng of plasmid was used in the reaction that was incubated at 30°C for 2 hours. Following completion of the reaction 1µl of the mix was analysed by Western blot to assess efficiency of translation. Translated peptides were stored at -20°C.

λ-phosphatase treatment of proteins: Recombinant LSH was dephosphorylated *in vitro* using λ-phosphatase (NEB). 2µg of LSH was incubated with 10 units of phosphatase for 1 hour at 37°C in the supplied buffer in a total volume of 20µl. Following incubation 1µl of treated protein was analysed by SDS-PAGE for integrity and the rest used for biochemical analysis.

***In vitro* pulldown of peptides using GST-tagged proteins:** Interactions between recombinant GST-tagged proteins and *in vitro* translated peptides were analysed by incubating 10 µl of GST-tagged protein immobilised on GST Sepharose beads with 5 µl of *in vitro* translation mix diluted in 100 µl of NE2 buffer. Incubations were performed for 1 hour at 4°C on a spinning fly wheel. The immobilised protein was pelleted by centrifugation at 6000g for 1 minute at 4°C and washed 3 times with 500 µl of NE2 buffer. Following the final wash most of the liquid was removed and the

GST Sepharose beads resuspended in 20 µl of 2X SDS-PAGE loading buffer. The beads were boiled at 100°C for 5 minutes, centrifuged at 13000g for 1 minute and 10 µl of the supernatant run on SDS-PAGE and analysed by Western blot.

2.2.3 Bacterial Methods

Bacterial growth on plates: L.B. agar was melted in a microwave and allowed to cool to ~50 °C before antibiotic was added. The L.B. agar mix was poured into round 10 cm dishes at a depth of ~ 5mm and allowed to cool at room temperature until solidified. The plates were then stored inverted at 4 °C. Prior to use plates were warmed to 37 °C. Bacteria were grown by streaking a single colony on the plate, or by diluting liquid culture and spreading evenly over the surface of the plate using sterile glass beads. Plates were inverted and incubated overnight at 37 °C.

Bacterial growth in Liquid culture: Antibiotic was added to L.B. media and mixed thoroughly. For 5ml cultures a single bacterial colony was scraped from the plate using a Gilson pipette tip and dropped into a 15 ml falcon tube. The culture was grown shaking at 250 rpm overnight at 37 °C. For larger cultures a single colony was incubated in a conical flask with a volume four times larger than the culture volume.

Preparing chemically competent strains: Bacterial cells (DH5α or Nova blue) were streaked from a glycerol stock onto LB agar plates and incubated overnight at 37 °C. A single colony was inoculated in 10 ml of LB overnight shaking at 37 °C. The 10 ml culture was then inoculated into a 200 ml flask of LB and incubated shaking at 37 °C until cell density reached an absorbance ob 0.15 at 600 nm (~1 hour). The cells were collected by centrifugation at 4000 rpm, 10 min at 4 °C and placed on ice for 10 minutes. Using sterile techniques the cells were gently resuspended in 3 ml of sterile ice cold 100 mM CaCl2, 70 mM MgCl2. The same buffer was then added to 50ml and the cells left on ice for 1 hour. The cells were next collected by centrifugation at 4000 rpm, 10 min at 4 ° and resuspended gently in 3 ml of 100 mM CaCl2, 70 mM MgCl2. 1ml of glycerol was added and mixed gently. Aliquots of competent cells (200 µl) were stored at – 80 °C. Competency was evaluated by transforming 0 ng, 10 ng, or 100 ng of super-coiled plasmid DNA into

aliquots of competent cells following 1 day at – 80 °C. Transformants per μg plasmid were derived by counting viable plasmid selected colonies arising the next day on LB agar plates. Competent cells were typically 105 to 106 colonies per μg of plasmid DNA.

Bacterial Transformation: Competent cells were defrosted on ice and DNA was added to the competent cell mix and gently mixed with a pipette tip. The transformation reaction was incubated on ice for 2 min. The cells were heat shocked for 2 min at 42 °C in a water bath, immediately followed by at 2 minute resting period on ice. 800 μl L.B. was then added to the transformation mixture and it was incubated at 37 °C shaking for 30 mins - 1 hr. For ligation reactions the cells were collected at 6000g for 1 min and resuspended in 200 μl of LB before plating on selective LB agar. If the transformed DNA was supercoiled plasmid 50 μl of the transformation reaction was directly plated on selective LB agar.

Expression of recombinant proteins in bacteria: E. coli strain ER2566 was used as a host strain to over-express target genes under the control of a T7 promoter and a lac operator. Both ER2566 and the plasmid encoding the target protein also encode the lacI repressor protein. ER2566 cells carry a chromosomal copy of the bacteriophage T7 RNA polymerase gene inserted into the endogenous lacZ gene, and thus under the control of the lac promoter and operator. In the absence of the non-cleavable lactose analog isopropyl-β-D-thiogalactoside (IPTG), the lacI repressor protein binds to the lac operators upstream of both T7 RNA polymerase and the target gene, and their expression is repressed. In the presence of IPTG, an IPTG-lacI complex forms; this complex cannot bind to the lac operators, and the T7 RNA polymerase gene is expressed. The T7 RNA polymerase protein then transcribes the target gene. The genes encoding the Ion and OmpT proteases are deleted in strain ER2566, which may increase the yields of some target proteins.

ER2566 cells were transformed with a plasmid(s) carrying a T7 promoter/lac operator-controlled target gene. A 10ml o/n bacterial culture in LB broth with appropriate antibiotic selection was prepared. The o/n culture was added to 500ml

LB broth the following morning, and the diluted culture was incubated at 37°C with shaking until an OD_{600} of 0.5 was reached. Protein expression was induced by the addition of IPTG to 1mM, and the bacterial culture was incubated at 37°C with shaking for a further 4h. Bacterial cells were pelleted by centrifugation for 30 minutes at 5000g at 4°C. Pellets were stored at -20°C until processing.

2.2.4 Purification of GST-tagged proteins from *E. coli*

Cell pellets were washed twice in ice-cold PBS and resuspended in 20ml PBS supplemented with complete protease inhibitors (Sigma). Cells were lysed by sonication using a Branson T250 Sonifier at setting 4, output 40%, for five minutes. The lysate was centrifuged at 20,000g for 30 minutes at 4°C and the supernatant further clarified by filtering it through a 0.45μm filter. The lysate was applied to 2ml of GST-sepharose beads pre-washed with PBS and incubated for 1 hour at 4°C. The protein bound beads were washed extensively with 50ml of PBS and stored at 4°C with 0.01% azide as a preservative. 5μl of beads were boiled in 2 X SDS-PAGE loading buffer and analysed by SDS-PAGE to assess yield and purity.

2.2.5 Yeast cell culture and two-hybrid screen

Yeast growth conditions: Yeast cells were recovered from frozen glycerol stocks by scraping the surface of the frozen stock with a sterile loop, and then streaking the loop across a YPDA agar plate. Plates were sealed with Parafilm to prevent desiccation, and were incubated inverted at 30°C until colonies grew to 2-3mm in diameter (2-3days). Working stock plates were then stored at 4°C for up to two months.

Liquid yeast cultures were grown in YPDA. A single yeast colony of 2-3mm in diameter was inoculated into 1ml broth. The broth was vortexed at high speed for 5min to disperse the cells, then diluted into a larger volume (up to 50ml). Liquid yeast cultures were incubated at 30°C with shaking at 200rpm o/n.

Lithium acetate-mediated transformation of yeast: Plasmid DNA was introduced into yeast cells by lithium acetate (LiAc)-mediated transformation, as described in the Yeast protocols handbook (Stratagene, 1998). A 50ml o/n yeast culture in YPDA broth was prepared, and the OD_{600} was measured the following morning. The culture was deemed to be in stationary phase when the $OD_{600} > 1.5$. The culture was diluted into 300ml YPDA broth to an OD_{600} of 0.2-0.3 (early log phase), then incubated at 30°C with shaking until an OD_{600} of 0.4-0.6 was reached (mid log phase). The yeast cells were pelleted by centrifugation at 1,000g for 5min, and were washed with 25ml TE pH7.5. The cells were then resuspended in 1.5ml TE/LiAc. 100µl resuspended yeast cells was added to a tube containing 100ng of each plasmid DNA to be transformed (i.e. 200ng plasmid DNA/co-transformation) and 100µg herring sperm ssDNA. The tube was then vortexed at high speed for 10s. 600µl PEG/LiAc was added and the transformations were mixed again by vortexing. The transformations were then incubated at 30°C with shaking for 30min. 70µl dimethylsuphoxide (DMSO) was added and the transformations were mixed gently by inversion, before being heat shocked in a 42°C water bath for 15min then incubated on ice for 2min. Finally, the transformed yeast cells were pelleted by a brief centrifugation and were resuspended in 1ml TE pH7.5. 100µl of the resuspended yeast cells was plated onto the appropriate SD agar plates, and the plates were incubated inverted at 30°C for 3-4days.

Yeast 2-hybrid screen: Yeast 2-hybrid screening of a Matchmaker Pretransformed HeLa cDNA library (Clonetech) with LSH and LSH(227-838) was carried out according to the manufacturers recommendations. Briefly, pGBTK7 plasmids expressing GAL4BD-LSH and GAL4BD-LSH(227-838) were transformed into the bait AH109 yeast strain as above. Expression of the fusion proteins was confirmed by Western blotting. For the screen, an overnight culture of the bait strain was grown overnight in 50ml SD/-Trp until the OD600 reached 0.8. The cells were pelleted by centrifugation at 1000g for 5 minutes and the supernatant discarded. The cells were resuspended in 5ml of SD/-Trp to obtain a high concentration culture. The concentrated bait strain was combined with 1ml of the pretransformed library strain

(Y187) in a sterile 2 litre flask. 45 ml of 2xYPDA liquid medium was added to the flask and mating proceeded at 30°C for 24 hours with gentle shaking (40rpm). The cells were pelleted by centrifugation at 1000g for 5 minutes and resuspended in 10ml of 0.5x YPDA liquid medium. From the mated culture 100µl of 1/10, 1/100, 1/1,000, 1/10,000 dilutions were plated on SD/-Trp, SD/-Leu and SD/-Leu/-Trp plates to assess mating efficiency. The remainder of the culture was plated, 200µl per 150mm SD/-Ade/-His/-Leu/-Trp/X-a-gal agar plate (~50 plates). The plates were incubated for 14 days and checked for positive colonies.

2.2.6 Mammalian cell culture methods

Mammalian cell culture: Mammalian cells were grown in defined growth media (DMEM or RPMI (Invitrogen)) with 10% donor calf serum, supplemented with 100µg/ml of a penicillin/streptomycin antibiotic mix and 2mM glutamine (Sigma). The cells were maintained at 37°C in 5% CO_2. Adherent cells were dislodged at 75 % confluence by washing with a dilute solution of trypsin/EDTA and replating at a 1 : 4 ratio.

Cryogenic storage of mammalian cells: Cells were collected by centrifugation at 500 g for 5 min and resuspended in 1ml of media containing 10% DMSO per 2.5 X 106 cells. Cells were frozen and stored at -80°C overnight. The following day the cells were transferred to liquid nitrogen for long term storage.

Transfection of mammalian cells: Cells were transfected with JetPEI (Polyplus Transfection) according to the manufacturer's recommendations. The next day the cells were washed with fresh media and allowed to recover for 24 hours before assaying for transfection efficiency and further manipulation.

Isolation of nuclei from tissue culture cells: Adherent tissue culture cells were dislodged from culture vessels using trypsin and collected by centrifugation at 500 g for 5 min at room temperature. The cells were washed once with PBS and resuspended in 2 ml of buffer NE2 per 107 cells. Nuclei were released by 10 plunges

of a tight fitting dounce homogeniser. Nuclei were recovered by centrifugation at 3000 rpm for 5 min at 4 °C.

Nuclear protein extracts from mammalian cells: Recovered nuclei were resuspended in 500µl buffer NE2 per 107 cells. Extraction of nuclear proteins was allowed to proceed for 1h at 4°C on a spinning fly wheel. Insoluble cellular material was pelleted at 13,000 rpm for 20 minutes at 4°C. The supernatant was taken as the nuclear extract. Nuclear extract was either used immediately or stored at –80 °C. The insoluble nuclear pellet was further extracted using the same buffer supplemented with 1M NaCl. The resuspended material was sonicated using a Branson T250 Sonifier for power setting 4, 40% output for 15 seconds. DNase and micrococcal nuclease were both added to the solution at 100u/ml along with 1mM CaCl2. Nuclease digestion was performed for 1 hour at 4°C. Centrifugation at 13,000 rpm for 20 minutes at 4°C liberated the soluble nuclear pellet fraction which was either used immediately or stored at –80°C.

Cell cycle synchronisation of mammalian cells: HeLa cells grown in DMEM were synchronised in S-phase by a double block with thymidine and mimosine. Asynchronously growing cells at ~25% confluence were blocked by addition of 2mM thymidine for 16 hours. The thymidine was washed off with PBS and fresh media added, allowing recovery from the block for 4 hours. Mimosine was then added to 40mM and the cells incubated for a further 16 hours. The cells were then washed with PBS and replenished with fresh media. Cells were allowed to grow for a further 3 hours to produce cells synchronously growing in S-phase or 8 hours to produce cells synchronously growing in G2/M-phase. At these points the cells were processed for both flow cytometric analysis to determine success of synchronisation and nuclear extraction.

Flow cytometric analysis of HeLa cells: Asynchronous or synchronised HeLa cells were isolated by treatment with trypsin/EDTA. Cells were pelleted by centrifugation at 1000g for 5 minutes and washed twice with PBS. The cell pellet was resuspended in 500 µl PBS and 5ml ice cold ethanol added dropwise. Cells were fixed at 4°C

overnight. 5 x 106 cells were pelleted as above and washed twice in 5ml PBS + 1% BSA. The washed cells were resuspended in 800 µl PBS + 1% BSA. 100 µl of 500 µg/ml propidium iodide in 38mM sodium citrate pH7.0 was added to the resuspended cells. RNase A was added to a final concentration of 1mg/ml and the cells incubated at 37°C for 30 minutes. Stained cells were analysed by flow cytometry on a FACsCalibur machine.

RNAi knock down of HDAC1 and HDAC2: SMART pool siRNA designed to target HDAC1, HDAC2, HDAC3 and GFP were purchased from Dharmacon.1.6 x 106 HCT116 cells were transfected with 5 µg SMART pool siRNA using nucleofector device and nucleofector reagents kit V (AmaxA Biosystems). 48 hours later 2.5x105 of RNA treated cells were co-transfected with reporter and effector plasmids using JetPEI reagent (Autogen Bioclear). The cells were collected for luciferase and β-galactisidase assays 24 hours after the reporter plasmids were introduced. The levels of HDAC1 and HDAC2 in nuclear extracts of RNAi treated cells were investigated 72 hours after RNAi transfection on Western blots with anti-HDAC1 and anti-HDAC2 antibodies.

2.2.7 Insect cell and baculovirus culture methods

Insect cell culture: All baculovirus stock generation and expression was done using SF9 cells (Invitrogen) which are derived from the papal ovarian tissue of the fall army worm Spodoptera frugiperda. Cells were grown in serum free SF900 II medium containing L-glutamine (Gibco) supplemented with penicillin and streptomycin (100µg/ml), and incubated at 27°C. Cells were grown as a monolayer in tissue culture flasks and split when confluent by dislodging the cells physically. Cells were diluted 1:10 to 1:20 in fresh SF900 II medium and either re-plated in new tissue culture flasks or inoculated as 50 ml cultures in a 200 ml spinner culture flask at 70 rpm. SF9 cells grown as suspension cultures in spinner flasks were maintained at densities of less than 2.5 x 10^6 cells per ml. Cells were passaged by centrifuging the cells at 500 rpm for 5 minutes to isolate the cells, and the old media was removed and replaced with fresh SF900 II media at 0.5 x 10^6 cells per ml.

Cryogenic storage of insect cells (SF9): Cells were collected by centrifugation at 500 g for 5 min and resuspended in 1ml of media containing 10 % DMSO per 2.5 X 106 cells. Cells were frozen and stored at - 80 °C overnight. The following day the cells were transferred to liquid nitrogen for long term storage.

Generation of recombinant baculovirus genome: The Fastbac system (Invitrogen) was used to generate recombinant baculovirus genome. The Invitrogen system takes advantage of site specific recombination to produce recombinant baculovirus genomes in bacteria (Luckow et al., 1993). This system relies on cloning the gene of interest into a shuttle vector followed by transformation into the bacterial strain DH10Bac which contain a recombination proficient baculoviral genome. Recombinant virus was generated according to instructions of the manufacturer and described briefly here. 10 ng of recombinant pFastbac plasmid DNA was added to 100 µl of competent DH10Bac cells which had been defrosted on ice. Cells were incubated on ice for 30 min. Cells were then heat shocked for 45 seconds at 42 °C, and immediately chilled on ice for 2 min. 900 µl of L.B. broth was added and the tubes were shaken for 4 hours at 37 °C to allow recombination into the viral genome. The cells were then plated as 100 µl of the transformation mix or 100 µl each of a 10 -1, 10 -2, and 10 -3 dilution. The recombinant viral genomes are maintained in the bacteria through selection for the co-integrated gentR marker gene Cells were grown on plates containing kanamycin, gentamicin, tetracycline, X-gal, and IPTG for 48 hours at 37 °C. White colonies should contain transposed baculovirus genome. 10 white colonies and 2 blue colonies were re-streaked onto the same selective media and incubated overnight at 37 °C. A single, confirmed white colony from the streaked plate was selected and grown in L.B. media overnight under the same selective pressure. In the morning cells were pelleted and the baculovirus genome recovered by Midi prep (Qiagen). To verify the transposition of the protein coding sequence into the baculovirus genome, diagnostic PCR was carried out on 5ng of bacmid DNA using M13 forward and reverse primers that flank the insertion site, and additional internal gene specific primers. Bacmid DNA isolated from a single blue colony was used as a negative control. PCR amplification was carried out with an initial denaturation of 96 °C for 5 min followed by 24 cycles of 95 °C for 30

second, 55 °C for 30 seconds, and 72°C for 7 min was carried out. A final extension of 72 °C for 10 min was used to extend PCR products.

The recombinant viral genomes are maintained in the bacteria through selection for the co-integrated gentR marker gene (Figure 3.1). The viral DNA is purified from the bacteria and used to transfect Sf9 insect cells which express the recombinant protein and also generate infectious recombinant viral particles. Recombinant baculovirus was isolated from the transfected Sf9 cells, amplified and used for large scale infection and protein expression. Typically, for protein expression and purification ~5 x 108 Sf9 cells were infected with virus at a multiplicity of infection of ~7 for 48h. Nuclei were prepared from these cells and a high salt buffer used to extract soluble protein. Both proteins were purified from the extract using a Cobalt affinity resin and analysed by SDS-PAGE to determine purity

Transfection of SF9 cells and isolation of P1 baculovirus stocks: Infectious P1 baculovirus stocks were generated according to the manufacturer's instructions (Invitrogen). Briefly, SF9 cells were seeded at a density of 9 X 10 5 in six well plates. 1 µg of baculovirus genome was diluted in 100 µl with un-supplemented Graces Medium (Invitrogen) and mixed with 6 µl of Cellfectin (Invitrogen) diluted in 100 µl Graces Medium. The mixture was left at room temperature for 45. During this incubation period the cells were washed once with 2ml Graces Medium which was replaced with 800 µl of the same media. The DNA / Cellfectin mixture was added to the plates and incubated at 27 °C for 5 hours. After 5 hours the transfection mix was removed and 2 ml of SF900 II medium was added. Cells were incubated at 27 °C for 72 hours and the supernatant collected as the viral P1 stock. The P1 stock was used immediately for subsequence amplification of virus. For storage, sterile FBS was added to the P1 stock to a final concentration of 2% and the virus stored at 4 °C for no longer than 1 week. The cells from the transfection were lysed on the plate with 200 µl of 2 X SDS loading buffer. The lysate was analyzed by western blotting for the expressed protein as a verification of baculovirus genome transfection.

Amplification of viral stock: To amplify the P1 viral stock 2 X 10 6 cells were plated in a 6 well dish in 2ml of SF900 II medium and 0.2ml of P1 stock added. The cells were incubated at 27 °C for at least 48 hours until cells began to detach. The media was collected from the cells and centrifuged at 500g for 5 minutes to remove cellular debris. The supernatant was saved as P2 stock, FBS added to 2% and stored at 4 °C protected from light. For long term storage aliquots of the P2 stock were frozen at - 70 °C.

Infection of SF9 cells for protein production: Prior to large scale protein expression and purification of protein using Baculovirus, P3 viral stocks were prepared. 5 x 107 Sf9 cells growing in suspension cultures at a density of 1 X 106 cells per ml were infected with 2ml of P2 stock for 72 hours at 27 °C. P3 viral stocks were isolated by centrifuging the cell culture at 500g for 5 minutes at room temperature. 35ml of P3 stock was used to infect 5 X 108 cells growing in suspension cultures at a density of 2 X 106 cells per ml for 48 hours 27 °C. Cells were collected by centrifugation at 500g for 5 min and the supernatant removed and discarded. Cell pellets were processed immediately as described below.

2.2.8 Protein purification from insect cells

Preparation of Cobalt affinity Sepharose: For purification of 6xHIS tagged proteins chelating Sepharose resin prebound with $CoCl_2$ was used. To generate the Cobalt affinity resin 1ml of chelating Sepharose was washed with 30 column volumes of dH_2O. One column volume of 100mM $CoCl_2$ was then added to the Sepharose and mixed by vortexing. Unbound Cobalt was removed by washing with 30 column volumes of dH_2O. Cobalt affinity resin was prepared fresh prior to protein purification. For protein purification a batch method was typically utilised, using 1ml of Cobalt affinity resin in a 15 ml falcon tube. Following binding and two washes the affinity resin was transferred to a 10 ml poly-prep column (Bio-Rad) for two further washes and protein elution.

Purification of recombinant DDM1: HIS-DDM1 was expressed in 5 x 10^8 Sf9 cells for 48 hours at 27°C. Cells were collected by centrifugation at 500g for 5 minutes at

4°C and the cell pellet washed twice with ice cold PBS + complete protease inhibitors (Sigma). After this point all procedures were carried out at 4°C or on ice. The cell pellet was resuspended in 10ml of extraction buffer (50mM Tris HCl pH 7.8, 300mM NaCl, 10% glycerol, 0.1% NP40 (v/v), complete protease inhibitors (Sigma)) and incubated on ice for 30 minutes. During incubation the cells were dounce homogenised with a tight dounce for 30 strokes. Cellular debris was pelleted by centrifugation at 13,000g for 30 minutes at 4°C. The supernatant was recovered and added to a 1ml MonoQ column pre-equilibrated with extraction buffer. The flow-through was recovered, imidazole added to 20mM and applied to a 1ml Cobalt affinity Sepharose column pre-equilibrated with extraction buffer + 20mM imidazole After 1 hour binding, the supernatant was removed and the column washed four times with 10ml extraction buffer + 20mM imidazole. HIS-DDM1 was eluted five times with 1ml extraction buffer supplemented with 250mM imidazole. The eluted protein was filtered through a 0.45μm filter and DTT added to 1mM. The protein was aliquoted and stored at -80°C.

Purification of recombinant DNMT1: HIS-DNMT1 was expressed in 1 x 10^9 Sf9 cells for 62 hours at 27°C. Cells were collected by centrifugation at 500g for 5 minutes at 4°C and the cell pellet washed twice with ice cold PBS + complete protease inhibitors (Sigma). After this point all procedures were carried out at 4°C or on ice. The cell pellet was resuspended in 10ml of homogenisation buffer (20mM Pipes pH 6.2, 10% (w/v) Sucrose, 3mM $MgCl_2$, 0.1% NP40, complete protease inhibitors (Sigma)) and dounce homogenised on ice for 15 strokes with a tight dounce. Following homogenisation, 5M NaCl was added to a final concentration of 250mM and homogenisation repeated as before and the cells incubated on ice for 30 minutes. The cell debris was pelleted by centrifugation at 13,000g for 30 minutes at 4°C. The supernatant was recovered and applied to a 1ml Cobalt affinity Sepharose column pre-equilibrated with buffer A (20mM Pipes pH 6.2, 250mM NaCl, 10% (w/v) Sucrose, 3mM $MgCl_2$, 0.1% (v/v) NP40, complete protease inhibitors (Sigma)). The column was washed four times with 10ml of buffer A and then eluted five times with 1ml buffer A supplemented with 250mM imidazole. The fractions

were pooled, filtered through a 0.45µm and DTT added to 1mM. The protein was aliquoted and stored at -80°C.

Purification of recombinant DNMT3B: HIS-DNMT3B was expressed in 5 x 10^8 Sf9 cells for 60 hours at 27°C. Cells were collected by centrifugation at 500g for 5 minutes at 4°C and the cell pellet washed twice with ice cold PBS + complete protease inhibitors (Sigma). After this point all procedures were carried out at 4°C or on ice. The cell pellet was resuspended in 10ml of buffer A (20mM Tris HCl pH 7.4, 1M NaCl, 1M Sucrose, 3mM $MgCl_2$, 0.3% Triton X 100 (v/v), complete protease inhibitors (Sigma)) and incubated on ice for 20 minutes. Following incubation the cells were dounce homogenised with a tight dounce for 30 strokes and left on ice for a further 20 minutes. The cell debris was pelleted by centrifugation at 13,000g for 30 minutes at 4°C. The supernatant was recovered and added to a 1ml MonoQ column equilibrated with buffer A. The flow-through was recovered and applied to a 1ml Cobalt affinity Sepharose column pre-equilibrated with buffer B (70mM Tris HCl pH 7.6, 1M NaCl, 0.29M Sucrose, 3mM $MgCl_2$, 0.1% Triton X 100 (v/v), complete protease inhibitors (Sigma)). After 1 hour binding, the supernatant was removed and the column washed four times with 10ml buffer B. HIS-DNMT3B was eluted five times with 1ml buffer B supplemented with 250mM imidazole. The eluted protein was filtered through a 0.45µm filter and DTT added to 1mM. The protein was aliquoted and stored at -80°C.

Purification of recombinant LSH: HIS-LSH was expressed in 5 x 10^8 Sf9 cells for 48 hours at 27°C. Cells were collected by centrifugation at 500g for 5 minutes at 4°C and the cell pellet washed twice with ice cold PBS + complete protease inhibitors (Sigma). After this point all procedures were carried out at 4°C or on ice. The cell pellet was resuspended in 10ml of extraction buffer (50mM Tris HCl pH 8.0, 300mM NaCl, 10% glycerol, 0.1% NP40 (v/v), complete protease inhibitors (Sigma)) and incubated on ice for 30 minutes. During incubation the cells were dounce homogenised with a tight dounce for 30 strokes. Following homogenisation, the nuclei were sonicated on ice using a Branson T250 Sonifier at setting 4, output 40%, for one minute. Cellular debris was pelleted by centrifugation at 13,000g for 30

minutes at 4°C. The supernatant was recovered and added to a 1ml MonoQ column pre-equilibrated with extraction buffer. The flow-through was recovered, imidazole added to 20mM and applied to a 1ml Cobalt affinity Sepharose column pre-equilibrated with extraction buffer + 20mM imidazole After 1 hour binding, the supernatant was removed and the column washed four times with 10ml extraction buffer + 20mM imidazole. HIS-LSH was eluted five times with 1ml extraction buffer supplemented with 250mM imidazole. The eluted protein was filtered through a 0.45μm filter and DTT added to 1mM. The protein was aliquoted and stored at -80°C.

2.2.9 Generating DNA probes for binding experiments

Generation of probes using PCR: DNA probes based on the natural MMTV nucleosome A (NucA) or unnatural 601.3 (W) sequences used for binding experiments with LSH were generated by PCR. The MMTV plasmid contains 485bp of MMTV promoter DNA cloned into pDONOR201 (Invitrogen) and was supplied by the Owen-Hughes laboratory (Flaus and Richmond, 1998) (Appendix I). The 601.3 template is cloned into the SmaI site of pGEM-3Z (Promega) and was supplied by the Widom laboratory (Anderson et al., 2002; Lowary and Widom, 1998). The primer sequences for amplifying the different fragments are listed in Table 2.1. For binding experiments using 36W36 nucleosomes and nucleosome mobilisation assays the primers were labelled with Cy5 dye at the 5' end for visualisation.

PCRs were carried out in 96 50μl reactions using a Biometra T3 Thermocycler in order to generate large amounts of product. Reaction included 1μM of each primer, 0.4ng/μl template plasmid, 120μM each dNTP, 2mM MgCl2 and 5 units / reaction of Taq polymerase (Roche). The reaction conditions were 94°C for 2 minutes, followed by 30 cycles of 94°C for 30 seconds, 50°C for 30 seconds and 72°C for 1 minute; after this the reactions were incubated at 72°C for 5 minutes. Following amplification the DNA in the reaction was precipitated by adding three volumes of 100% ethanol and one tenth volume of 3M sodium acetate pH 5.2. The DNA was pelleted by centrifugation at 13,000 rpm for 15 minutes and the DNA resuspended in 500μl of TE. The concentrated DNA was subsequently purified using ten PCR

purification columns (Qiagen) and eluted twice with 50µl TE. The purified DNA was then precipitated again as before and following two washes with 70% ethanol resuspended in 20µl TE.

Generation of probes using single stranded oligonucleotides: Equal molar amounts of a complementary single stranded oligonucleotides (X base pairs) were incubated in 10 mM Tris pH 8.0 with 100mM NaCl for 10 min on a 100°C hot block. The hot block was removed from the heat source and allowed to cool to room temperature. The product was analysed by agarose gel electrophoresis and ethidium bromide staining to assay for integrity.

2.2.10 Electrophoretic Mobility Shift Assay (EMSA)

Probes used for EMSA were generated either by PCR amplification from plasmid DNA or by annealing single stranded oligonucleotides (Section 2.2.10). The probes were either labelled with ^{32}P (Section 2.2.1) or contained a Cy5 labelled primer (Section 2.2.9). EMSA reactions were assembled at room temperature by combining protein and specific probe in binding buffer (25mM Tris-HCl pH7.5, 50mM KCl, 0.5mM MgCl2, 10% glycerol, 100µg/ml BSA and 1mM DTT) with or without non specific competitor DNA. The reaction was incubated at 37°C for 30 minutes to reach equilibrium and bromophenol blue added to 0.01%. The binding reaction was loaded on to a 5 % non denaturing polyacrylamide gel made with TGE that had been pre run for 1 hour at 200 V in TGE. The samples were electrophoresed for 2 hours at 200 V. Gels using 32P labelled probes were transferred to 2 layers of 0.2 mm Whatman paper and dried at 80 °C for 1 hour. Dried gels were exposed to a phosphor-imager screen overnight and developed the next day. Gels using Cy5 labelled primers were imaged immediately following electrophoresis using a Fuji Phosphoimager FLA-5100.

2.2.11 Biochemical Assays

Histone octamer preparation:

Histone octamers were a kind gift from the Owen-Hughes laboratory. They were prepared as follows. *Xenopus laevis* histones H2A, H2B, H3 and H4 were expressed

and purified from *E. coli* inclusion bodies individually and stored as powders at -20°C as described (Luger et al., 1999). The pure histones were dissolved in unfolding buffer (20mM Tris-HCl pH7.5, 7M guanidinium hydrochloride, 10mM DTT) and mixed together in equimolar ratios. The mixture of histones was diluted with unfolding buffer to a concentration of 1mg/ml. The mixture was dialysed at 4°C against 2 litres refolding buffer (2M NaCl, 10mM Tris-HCl pH7.5, 1mM EDTA, 5mM mercaptoethanol) three times using 6000-8000 molecular weight dialysis tubing (Spectrum). Precipitate was removed by centrifugation at 12000g for 30 minutes at 4°C. Using Superdex 200 16/60 gel filtration chromatography, refolded octamers were separated from other arrangements of histone proteins. Refolded octamers eluted from the column were concentrated to 15 – 30μM using Amicon Ultra 4ml 5000 molecular weight cut off concentrators (Millipore) spun at 4°C at 4000g. Octamers were stored at 4°C.

Nucleosome reconstitution: To assemble nucleosomes a stepwise dialysis method was used. Histone octamer and DNA were added in equimolar ratios at 2μM to a buffer containing 2M NaCl and 10mM Tris HCl pH 7.5 in a volume of 25μl. The mixture was added to a 25μl dialysis button (Hampton Research) and covered with dialysis tubing. The samples were dialysed against 2.5ml 0.85M NaCl, 10mM Tris HCl pH 7.5 for two hours with gentle agitation. This was repeated with buffers containing 0.65M NaCl and 0.5M NaCl. The final dialysis step, into 10mM Tris HCl pH 7.5 was carried out once for two hours and then again overnight. Reconstitutions were carried out at 4°C. Following reconstitution a small amount of the sample was run on a 1X TBE 5% acrylamide gel for 1 hour at 100V alongside free DNA. The gel was stained with ethidium bromide and assembly of nucleosomes determined by shift in migration of the DNA.

ATPase assay: ATPase assays used ATP ^{32}P labelled at the γ phosphate position. Hydrolysis of ATP therefore produced radioactive phosphate and unlabelled ADP. To allow calculation of the amount of hydrolysis ATP and Pi were separated using thin layer chromatography (TLC). TLC plates were 20cmx20cm PEI cellulose plastic plates (Merck). These were pre run with water and allowed to dry before use.

Reactions contained the indicated concentration of ^{32}P- γ-ATP and were carried out in a buffer of 20mM Tris-HCl, pH8.0, 50mM NaCl, 1.5mM $MgCl_2$, 1mM mercaptoethanol and 100μg/ml BSA. For stimulated reactions 200nM of stimulus was used. Following incubation at 37°C for the appropriate time, reactions were stopped by incubation on ice. 0.3μl of each reaction was spotted onto plates and allowed to dry. Separation was achieved using a buffer of 0.5M lithium chloride and 1M formic acid. The plates were dried, exposed to a phosphoimaging screen and ATP hydrolysis calculated using ImageJ software.

Nucleosome mobilisation assay: Nucleosome mobilisation assays were carried out in the Owen-Hughes laboratory. Nucleosomes mobilisation assays take advantage of the different mobilities through native polyacrylamide gels of nucleosomes at different positions on the same DNA fragment (Meersseman et al., 1992). Reactions contained 100nM of nucleosomes assembled on Cy5 labelled DNA or 1nM of nucleosomes assembled on 32P labelled DNA. Reactions were carried out in a buffer of 50mM Tris-HCl, pH8.0, 50mM NaCl, 3mM $MgCl_2$, 100μg/ml BSA and 1mM ATP and concentrations of LSH and DDM1 as indicated in the figures. Reactions were incubated 37°C for 1 hour then stopped by addition of NaCl to 200mM and pUC18 DNA to a final concentration of 0.5μg/μl and further incubation at 37°C for 10 minutes. Sucrose was added to a final concentration of 2% as a loading buffer and reactions run for 3.5 hours on a 0.2X TBE 5% acrylamide gel at 4°C and 300V with 0.2X TBE as running buffer.

SssI methyltransferase assay: SssI methylation of CpG sites protected by nucleosome position in 67A0 nucleosomes was used as an assay of nucleosome disruption by LSH. Reaction were carried out in a buffer of 10mM Tris-HCl pH7.9, 50mM NaCl, 10mM $MgCl_2$, 1mM DTT, 1mM ATP, 160μM SAM and 100μg/ml BSA with the amount of SssI and LSH indicated in the figure. 100nM 67A0 nucleosome was used as a substrate for SssI methylation. Reactions were incubated for 15 minutes at 37°C and stopped by addition of equal volume of phenol-chloroform-isoamyl alcohol. Two phenol extractions and one chloroform extraction liberated the nucleosomal DNA which was precipitated with ethanol as previously

described and resuspended in 20µl of TE. The resuspended DNA was digested with the methylsensitive restriction enzyme AciI (NEB) in the recommended buffer for 1 hour at 37°C and the digested DNA run on a 1X TGE 7% acrylamide gel at 100V for 1 hour. DNA was stained with SYBR gold (Invitrogen) and protection from AciI digestion due to SssI methylation analysed.

DNA methyltransferase assay: DNA methyltransferase assays were carried out in a buffer consisting of 10mM Tris-HCl pH7.9, 50mM NaCl, 10mM $MgCl_2$, 1mM DTT, 1mM ATP, 1µCi of ^{3}H-SAM and 100µg/ml BSA. 200ng dGdC, 100nM nucleosomal DNA or 100nM nucleosome was used as a substrate in a total volume of 25µl. The amount of enzyme added was indicated in the figure. Reactions were incubated at 37°C for 2 hours after which the entire reaction was spotted onto DE81 paper (Whatman) and left to dry. The paper was then washed three times with 200mM ammonium bicarbonate, once with dH2O and once with 100% ethanol. Each wash lasted five minutes. Following the final was the paper was dried at 65°C and each reaction added to a scintillation vial. 5ml of ecoscint (National Diagnostics) was added and radioactivity incorporation measured by scintillation counting on a TriCarb 2100TR machine.

2.2.12 Analysis of protein complexes

Gel filtration chromatography: Proteins were applied directly to a Superose 6 or Superose 12 gel filtration column equilibrated according to the manufactures recommendations. Molecular weight size standards (Sigma) were run over the column and their elution volumes monitored by absorbance at 280 nm (Table 2.5). Protein samples were applied to the column with volumes not exceeding that recommended by the manufacturer and eluted with a constant buffer flow rate of 0.2 ml / minute collecting fractions within the inclusion limit. These columns were run on a ACTA purifier FPLC machine. Samples were stored at –80 °C or used immediately.

Marker	Size (kDa)	Elution volume (ml)	Elution fraction	Stokes radius [R_s] (nm)
Thyroglobulin	669	12.4	12	8.5
Apoferritin	443	14.1	15.5	6.1
Alcohol dehydrogenase	150	15.8	18.5	4.55
Bovine serum albumin	66	16.4	22	3.55

Table 2.5. Elution of size standards from Superose 6 column.
The molecular size, elution volume, elution fraction and Stokes radius of the protein size standards used to equilibrate the Superose 6 column are listed.

Sucrose gradient sedimentation: Sucrose gradients were manually formed at room temperature in 2 ml 1-3/8 UC tubes (Beckman). 500µl of NE2 buffer without glycerol containing 5% stepwise increases in sucrose concentration from 20 – 5% were carefully pippeted on top of each other. Following addition of the final layer the gradient was allowed to sit at room temperature for one hour. 100µl of nuclear protein extracted with NE2 buffer without glycerol was added to the top of the gradient. An identical gradient was loaded with 50µg of standard proteins BSA, appoferritin, β-amylase and ADH. The gradients were centrifuged at 55,000 rpm in a TLS-55 rotor for 4 hours at 4 °C. 100µl fractions were collected manually from the top of the gradient using wide bore pipette tip. 10µl of each fraction was used for an SDS PAGE gel followed by Coomassie stain to identify marker proteins or a western blot to identify nuclear protein of interest.

Calculating the native molecular mass of a protein: Calculations to determine the molecular weight of native LSH were applied as described previously (Siegel and Monty, 1965) using the equation $M_r = 6\pi\eta_{20,w} \cdot s_{20,w} \cdot R_s \cdot N/(1 - \nu_{20,w}\upsilon)$, where R_s is the Stoke's radius (nm), $s_{20,w}$ is the sedimentation velocity (S x 10^{-13}), $\eta_{20,w}$ is the viscosity of water at 20°C (0.01002 $g \cdot s^{-1}$ cm^{-1}), N is Avogadro's number (6.022 x

10^{23}·molecules^{-1}), $\upsilon_{20,w}$ is the density of water at 20°C (0.9981 g·cm^3), and $\upsilon$is the partial specific volume of protein (used 0.725 cm^3/g).

2.2.13 Analysing the effect of LSH on transcription

Luciferase Reporter assay: In reporter assays, GAL4-TK-LUC (500 ng) and pact-βgeo (500 ng) plasmids were cotransfected with 100ng or 500ng of the indicated effector plasmids into 2.5 x 10^5 HCT116 cells or DNMT knockout (KO), using JetPEI reagent (Autogen Bioclear). To test whether LSH-mediated repression was sensitive to HDAC inhibitors, trichostatin A (TSA) was added to the tissue culture medium after transfection to a 100 nM final concentration 24 h prior to lysing the cells. For rescue experiments with HCT116 KO cell lines, 3000 ng of DNMT1-green fluorescent protein (GFP) or DNMT3B-GFP plasmid was cotransfected with the reporter plasmids in the presence or absence of GAL4-LSH or GAL4-LSH(1-226). All transfection experiments were performed in triplicate. At 48 h post transfection, cells were harvested using reporter lysis buffer (Promega), and detection of luciferase activity was carried out according to the manufacturer's guidelines. Briefly, cells were washed twice on the plate with PBS then incubated with 1X reporter lysis buffer and subjected to a single freeze thaw cycle. Lysed cells were transferred to eppendorf tubes and centrifuged at 13,000 rpm for 1 minute to remove cellular debris. 20µl of cell lysate was added to 100µl of Luciferase Assay Reagent (Promega) and luminescence was measured in a TD20/20 luminometer (Turner Designs).

B-gal assay: Cell lysate was prepared from transfected cells using 1X reporter lysis buffer (Promega). 150 µl of lysate was added to 150 µl of a buffer containing 200mM sodium phosphate pH7.3, 2mM MgCl2, 100mM mercaptoethanol, 1.33mg/ml ortho-Nitrophenyl-β-galactoside. The reactions were incubated for 30 minutes at 37°C until a faint yellow colour developed. Reactions were stopped by addition of 500 µl 1M sodium carbonate and absorbance determined at 420nm.

Chapter three - ATPase activity and chromatin remodelling by LSH

3.1 Introduction

The SNF2 family of enzymes was first identified in genetic screens for genes involved in regulating mating type switching (SWI) and sucrose fermentation (Sucrose Non-Fermenting (SNF)) in yeast. Subsequent cloning of one of the mutants involved in this process identified it as the protein Snf2p which is the catalytic core of the multi-subunit SWI/SNF complex (Abrams et al., 1986). Several other proteins homologous to yeast Snf2p were cloned from a number of organisms and were shown to contain the seven conserved sequence motifs common to the SF2 superfamily of helicases (Davis et al., 1992; Laurent et al., 1992; Okabe et al., 1992). Sequence analysis placed these proteins in a distinct family of helicases which was termed SNF2 (Bork and Koonin, 1993). Laurent and colleagues were the first to demonstrate the catalytic activity of a SNF2 family member and showed that similar to other helicases Snf2p is an ATPase stimulated by DNA *in vitro*. Unlike other helicases however, the purified SWI/SNF complex showed no detectable helicase activity (Cote et al., 1994). This study and others that followed outlined another important function for SNF2 proteins, ATP dependant disruption of chromatin structure (Cairns et al., 1996; Owen-Hughes et al., 1996).

3.1.1 The enzymatic activity of SNF2 enzymes are believed to determine their function *in vivo*

Following these pioneering experiments ATPase activity has been demonstrated for members of all four SNF2 classes of protein *in vitro*. This activity is found in protein complexes purified biochemically from native sources and in recombinant SNF2 proteins purified from *E. coli* or insect cells. Depending on the protein, the ATPase activity is stimulated either by DNA, nucleosomes or both and is required for its chromatin remodelling role. The biochemistry of SNF2 enzymes has been

extensively studied and a wide range of activities have been detected including nucleosome repositioning, DNA translocation and histone dimer exchange (Narlikar et al., 2002). The specific *in vitro* activities of different SNF2 enzymes are intrinsically linked to their function *in vivo*. For example, nucleosome repositioning at the promoters of rDNA genes is catalysed by SNF2h and mutants deficient for ATP hydrolysis are deficient in remodelling, indicating that nucleosome repositioning *in vivo* is an ATP dependant process (Li et al., 2006b). Also, the SNF2 enzyme Mot1, a conserved transcriptional regulator has been shown to displace the TATA binding protein (TBP) from TATA Box sequence DNA in an ATP dependant manner (Auble et al., 1994). The mechanism of this action has been investigated and appears to involve relocation of TBP to a position upstream of the TATA box (Sprouse et al., 2006). This data suggests a mechanism by which Mot1 utilises ATP hydrolysis to displace TBP from TATA Boxes by translocating DNA. Thus, the DNA dependant ATPase and suggested, but currently unconfirmed translocase activity of Mot1 is critical to its *in vivo* function (Auble and Steggerda, 1999; Sprouse et al., 2006).

3.1.2 The enzymatic activities of SNF2 enzymes involved in DNA methylation

How the biochemical activities of several SNF2 enzymes involved in DNA methylation relate to their *in vivo* function has been investigated. Recombinant SNF2h for example has nucleosome stimulated ATPase activity and can reposition nucleosomes *in vitro* (Aalfs et al., 2001; Fan et al., 2005a). Also, as outlined above, an attractive model for its *in vivo* function has been developed based on evidence obtained at the promoters of rDNA genes. This model proposes ATP dependant nucleosome sliding of a specific nucleosome positioned at −157 to −2 relative to the transcription start site to a location 25bp upstream. This repositioning leads to exposure of a single CpG site that subsequently becomes methylated, leading to transcriptional silencing of the gene (Li et al., 2006b). Another SNF2 protein involved in DNA methylation, ATRX has also been studied biochemically. When purified from HeLa cells, native ATRX-Daxx complex is an efficient DNA stimulated ATPase that shows DNA translocase activity (Xue et al., 2003). The purified complex also has the ability to disrupt DNA:histone contacts at the entry site

of nucleosomes but does not appear able to reposition nucleosomes (Xue et al., 2003). How this correlates with its *in vivo* function and role in DNA methylation is unclear but a model where ATRX peels DNA off the surface of the nucleosome at the entry site to allow DNMTs access is plausible. The roles of SNF2h and ATRX in DNA methylation appear to be very specific and relate only to a limited number of loci (Gibbons et al., 2000; Santoro et al., 2002). The other SNF2 enzymes with a known role in DNA methylation are plant DDM1 and it mammalian homologue Lsh. These enzymes have been shown to have a role in both global and locus specific DNA methylation indicating that they may be more general regulators of DNA methylation (Dennis et al., 2001; Jeddeloh et al., 1999; Vongs et al., 1993). To date, only DDM1 has been characterised biochemically and the data available is limited. Using recombinant protein expressed in insect cells, Brzeski and Jerzmanowski demonstrated that DDM1 is an efficient DNA and nucleosome dependant ATPase that hydrolyses ATP at a maximum rate (~400 ATP/min/DDM1) comparable to other SNF2 enzymes. They also show that DDM1 binds DNA and nucleosomes and can reposition mononucleosomes from the end to the centre of a DNA fragment (Brzeski and Jerzmanowski, 2003). They also show that the methylation status of the DNA used for the ATPase assays does not affect the activity of DDM1. These experiments determined that DDM1 is an active enzyme that has nucleosome sliding activity but gave no indication of how this relates to the critical role of DDM1 in DNA methylation. They do however indicate that DNA methylation in plants requires active chromatin remodelling. The mammalian homologue of DDM1, Lsh/LSH has not been characterised biochemically. As outlined above the enzymatic function of SNF2 enzymes is intrinsically linked to their *in vivo* roles and can give indications as to their function. As the molecular function of LSH is largely unknown I attempted to biochemically characterise it. Hopefully, *in vitro* kinetic data from these experiments would allude to the mechanism by which LSH facilitates DNA methylation *in vivo*.

3.2 Results - Recombinant LSH binds DNA

3.2.1 Cloning, baculovirus expression and purification of LSH

The activities of several SNF2 ATPases from several different species have been determined. In most cases the activity determined for the native remodelling complex purified biochemically is very similar to that of recombinant protein expressed and purified in *E. coli* or insect cells. The notable exception is the previously discussed modulation of ISWI activity by its cofactor Acf1 (Langst et al., 1999). I chose to initially test whether LSH displays ATPase and chromatin remodelling activities using recombinant protein. I chose to derive this from insect cells using a baculovirus expression system. This system is advantageous over bacterial systems for expressing mammalian proteins as the insect cells used for expression have a post translational modification system more closely related to mammals.

Full length human LSH cDNA and LSH carrying a point mutation at a conserved lysine critical for ATP binding (K245Q) were obtained from the Arceci laboratory (Raabe et al., 2001). Other SNF2 enzymes carrying mutations at this residue are catalytically inactive so LSH^{K254Q} served as a negative control in enzymatic experiments (Fitzgerald et al., 2004). These plasmids were sequenced to confirm the cDNAs were not carrying mutations and then cloned into the Invitrogen baculovirus shuttle vector pFAST-BAC. Following this cloning step both plasmids were sequenced to ensure PCR amplification had not introduced any mutations (Appendix II). The Invitrogen system takes advantage of site specific recombination to produce recombinant baculovirus genomes in bacteria (Luckow et al., 1993). This system relies on cloning the gene of interest into a shuttle vector followed by transformation into the bacterial strain DH10Bac which contain a recombination proficient baculoviral genome (Figure 3.1B). The recombinant viral genomes are maintained in the bacteria through selection for the co-integrated gentR marker gene (Figure 3.1B). The viral DNA is purified from the bacteria and used to transfect Sf9 insect cells which express the recombinant protein and also generate infectious recombinant viral particles. Recombinant baculovirus was isolated from the transfected Sf9 cells,

amplified and used for large scale infection and protein expression. Typically, for protein expression and purification ~5 x 10^8 Sf9 cells were infected with virus at a multiplicity of infection of ~7 for 48h. Nuclei were prepared from these cells and a high salt buffer used to extract soluble protein. Both proteins were purified from the extract using a Cobalt affinity resin and analysed by SDS-PAGE to determine purity (Figure 3.2A). In order to remove DNA and other contaminants that may interfere with downstream applications I applied purified HIS-LSH proteins to a MonoQ anion exchange column. HIS-LSH did not bind this column but DNA as analysed by agarose gel electrophoresis was removed (Figure 3.2B and C). The concentration of the purified proteins was determined by BCA protein assay in relation to BSA standards. I estimated the concentration of my purified recombinant proteins to be 100ng/µl and the purity ~90% (Figure 3.2B).

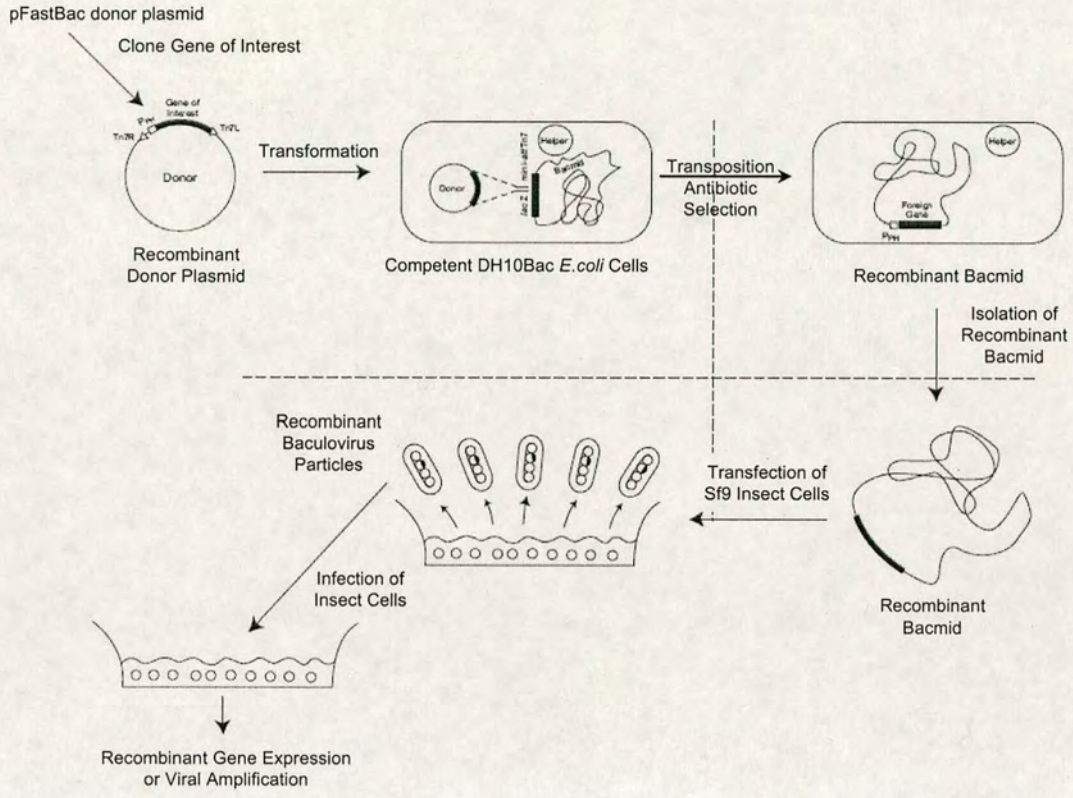


Figure 3.1. Bac-2-bac baculovirus expression system .
Schematic of the bac-2-bac baculovirus expression system is shown. Gene of interest is cloned into pFAST-BAC shuttle vector, which is transformed into DH10Bac *E.coli* cells. Transposition into the baculoviral genome occurs in this strain. The recombinant baculovirus genome is isolated from these cells and used to transfect Sf9 insect cells. Recombinant baculovirus particles are formed in the insect cells and released into the media. The virus is amplified and used for large scale expression experiments.

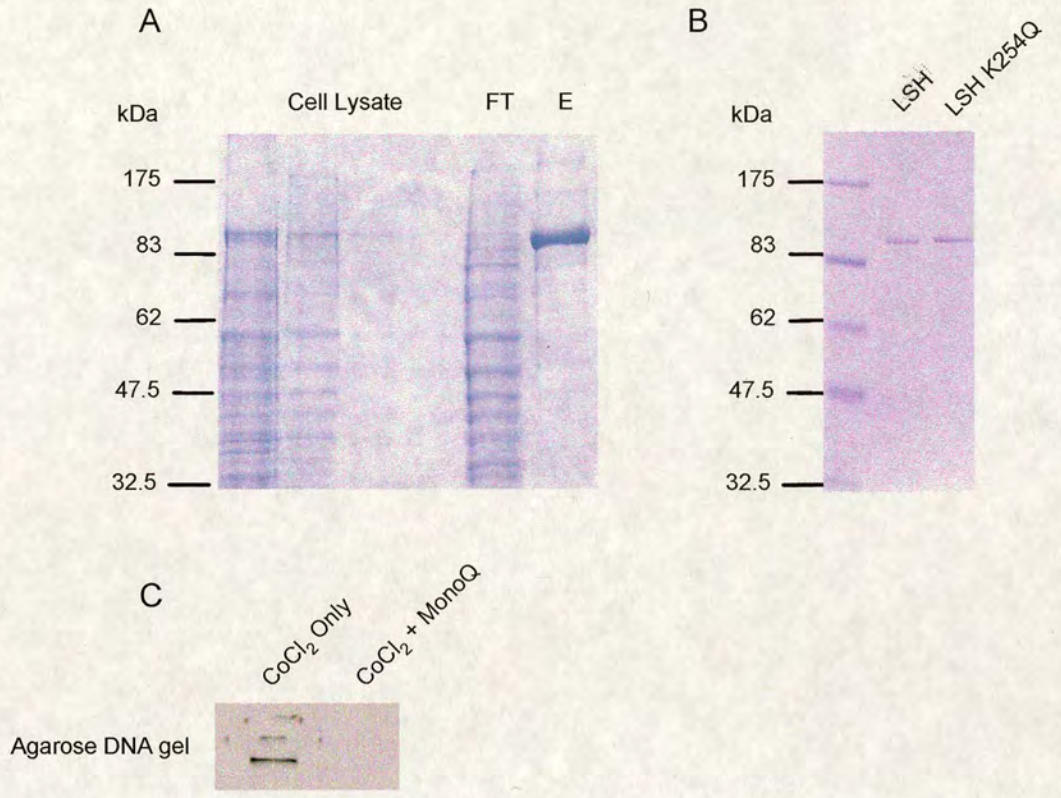


Figure 3.2. Purification of recombinant LSH from insect cells.
A. Coomassie stained gel outlining the purification of HIS-LSH via Cobalt affinity chromatography. 5µl, 2µl and 1µl of cell lysate, 5µl of flow-through (FT) from the Cobalt column and 5µl of elution (E) from the Cobalt column loaded. Size markers in kDa listed to the left.
B. Coomassie stained gel of 1µl of purified recombinant LSH and LSH^{K254Q} following Cobalt affinity and anion exchange (MonoQ) chromatography. Size markers in kDa listed to the left.
C. Ethidium bromide stained agarose gel of 5µl of recombinant LSH purified via Cobalt affinity or Cobalt affinity and anion exchange chromatography. High molecular weight DNA present following Cobalt affinity chromatography is removed by anion exchange (MonoQ) chromatography.

3.2.2 LSH binds DNA and mononucleosomes with linker DNA

3.2.2.1 LSH binds DNA

To characterise LSH biochemically I used a number of controlled reagents. Several DNA sequences that direct nucleosome assembly to defined locations have been isolated and are commonly used for studies involving SNF2 enzymes (Lowary and Widom, 1998). I used two such sequences in this study: the artificial 601.3 sequence

and the MMTV nucleosome A; also known as NucA. The nomenclature for the fragments used is xAy (for NucA) and xWy (for 601.3); x and y indicate the length of DNA extension in base pairs on either side of the nucleosome assembly site. These DNA sequences were initially used to test the DNA binding ability of LSH. I also used these sequences to reconstitute nucleosomes with purified recombinant histones for binding studies. I first performed an electrophoretic mobility shift assay (EMSA) with recombinant LSH and radio labelled NucA DNA probe. I found that increasing concentrations of LSH shifted the mobility of increasing amounts of the probe to a higher molecular weight and depleted free probe indicating that LSH can bind DNA (Figure 3.3A; bottom arrow). I noted that the LSH:DNA complex did not appear as a single discreet band in the gel but rather formed a smear with a proportion of it not entering the gel at the highest concentrations of LSH (Figure 3.3A; top arrow). This led me to believe that multiple LSH molecules may be binding to the 147bp DNA fragment or the conditions may not be optimised for detecting a soluble LSH:DNA complex. I therefore sought to confirm DNA binding by LSH using a shorter probe with a single central CpG site and optimised conditions. I repeated the experiment using a 68bp DNA probe with or without unlabelled competitor DNA (Figure 3.3B). This experiment gave a clearer result with a soluble LSH:DNA complex clearly identifiable. Addition of an excess of the same unlabelled, non methylated DNA successfully depleted the LSH:DNA complex indicating formation of a specific complex. I therefore conclude that LSH is able to bind DNA in EMSA experiments. Due to its role in DNA methylation I thought it pertinent to ask whether the DNA methylation status of the probe affected the ability of LSH to bind it. In parallel, I performed the EMSA experiment using the same probe that had been *in vitro* methylated by the bacterial DNMT Sss1. I found that the ability of LSH to shift the probes was not significantly altered by methylation of the single CpG site (Figure 3.3B). There did appear to be a very slight difference in the efficiency of competition of the two probes by unlabelled non methylated competitor (Figure 3.3B; compare lanes 6+7 to 13+14). However, this difference is very minor and could be due to the slightly unequal loading of the probes (Figure 3.3B; compare lanes 1 and 8). Thus, the binding of LSH to this probe is not significantly affected by methylation of a single CpG site.

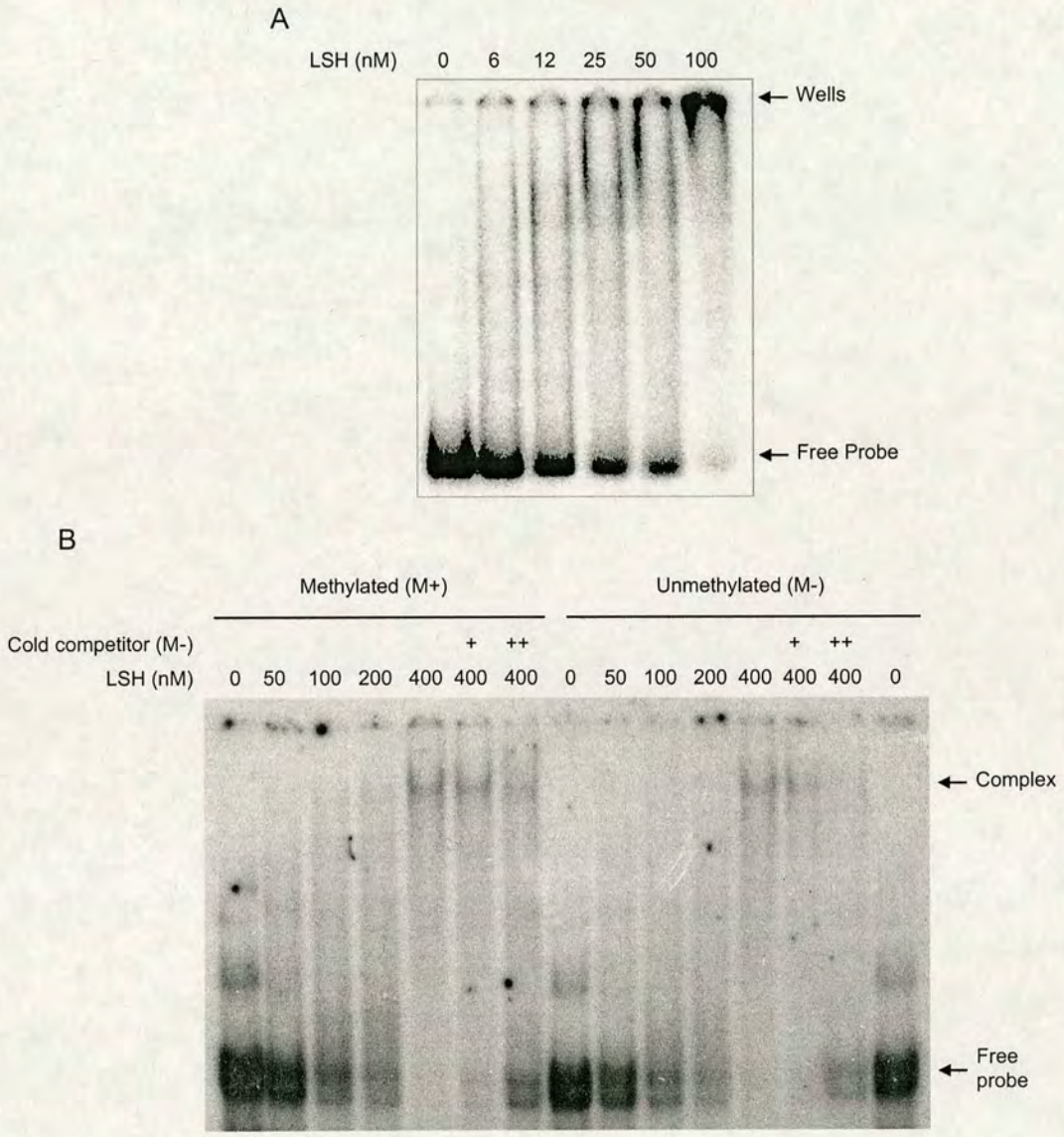
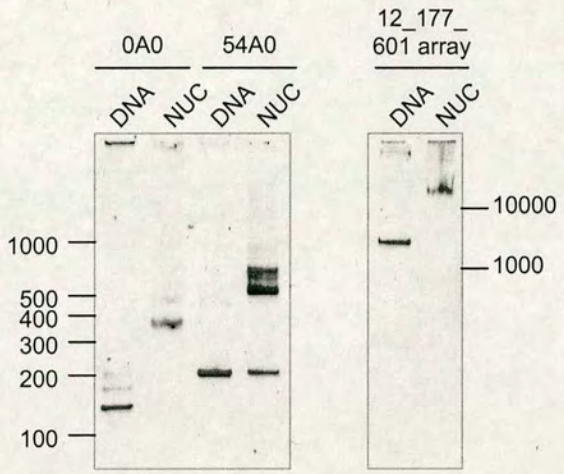


Figure 3.3. LSH binds DNA in electrophoretic mobility shift assays.
A. This experiment was carried out with the help of Dr Chris Stockdale. Electrophoretic mobility shift assay (EMSA) with recombinant LSH and radio labelled 0A0 DNA probe. Increasing concentrations of LSH (indicated on the top of the panel) mixed with 10nM probe for 30 minutes at 37°C. Free probe separated from LSH bound DNA by native polyacrylamide gel electrophoresis. Following electrophoresis, the gel was dried, exposed to a phosphor imager screen and developed using a Storm PhosphorImager (GE Healthcare). Free probe and LSH:DNA complex (Wells) indicated to the right.
B. EMSA with recombinant LSH and radio labelled 68bp DNA probe with a single central CpG site. Increasing concentrations of LSH (indicated on the top of the panel) mixed with 10nM of methylated (M+, left panel) or unmethylated (M-, right panel) probe for 30 minutes at 37°C. 10nM (+) or 20nM (++) of the same cold unmethylated probe included as competitor where indicated. Free probe separated from LSH bound DNA by native polyacrylamide gel electrophoresis. Following electrophoresis, the gel was dried and exposed to photographic film. Free probe and LSH:DNA complex indicated to the right.

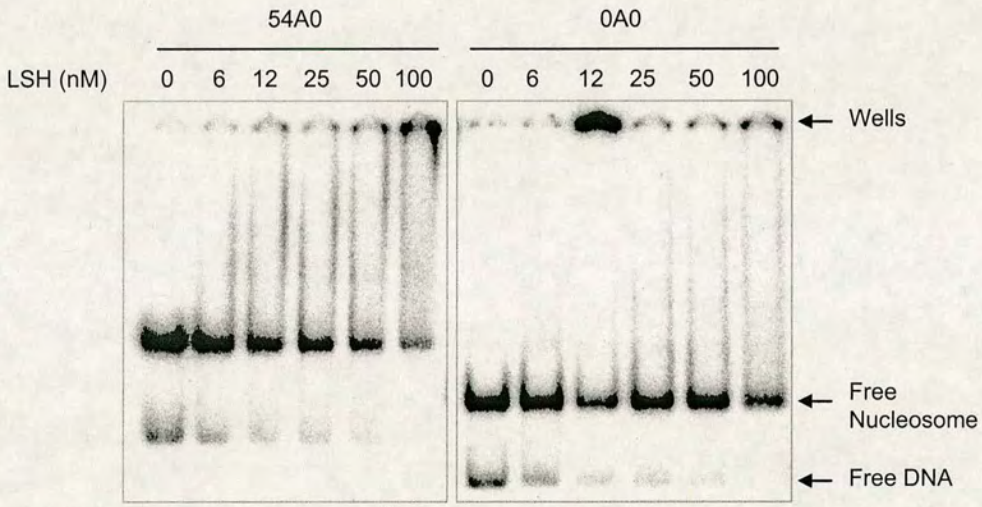
3.2.2.2 LSH bind nucleosomes with linker DNA

I next tested whether LSH can bind DNA assembled into nucleosomes and whether this interaction requires linker DNA. For this I used the two positioning sequences to assemble nucleosomes with recombinant core histone octamers. As mentioned above the DNA sequence flanking the nucleosome assembly site determines the amount of linker DNA present and is designated as the number of bp 5' or 3' of it. For example 54A0 is a 3' end positioned NucA nucleosome with 54bp of DNA 5' to the nucleosome and 0bp 3' of it. A variety of nucleosomes with different linker lengths were assembled and run on native polyacrylamide gels to determine assembly of the constitute parts into nucleosomes. A representative gel of a typical nucleosome reconstitution shows that the migration of DNA is retarded following reconstitution, indicative of nucleosome assembly (Figure 3.4A). EMSA experiments using reconstituted 54A0 nucleosomes gave similar results to those using free NucA DNA. That is, increasing amounts of a slower migrating population was generated upon addition of increasing amounts of LSH and free probe was depleted (Figure 3.4B; left panel). The shifted probe did not form a clear band and again, a proportion of it did not enter the gel at all. I was unable to optimise these experiments using unlabelled competitor DNA or nucleosomes so had to assess LSH binding by depletion of free probe. The reconstitution of nucleosomes is not 100% efficient and a small proportion of free, non nucleosomal DNA can be observed in these reactions. Depletion of this was not considered when analysing nucleosome binding (Figure 3.4B; bottom arrow) Interestingly, LSH was able to deplete the free nucleosome with linker DNA but was much less efficient at depleting 0A0 nucleosomes that lack linker DNA (Figure 3.4B; right panel). This indicates that LSH is able to bind nucleosomes but binds with greater affinity if linker DNA is present. Depletion of free 0A0 DNA indicates that LSH was active for DNA binding in this experiment and acts as an internal control (Figure 3.4B; right panel). The role of LSH in DNA methylation is largely uncharacterised but the role of the plant protein DDM1 appears to be correlated with H3K9 methylation (Gendrel et al., 2002). I therefore determined whether LSH bound nucleosomes containing H3K9me3 with higher

A


0A0
54A0
12_177_
601 array
DNA
NUC
DNA
NUC
DNA
NUC
1000
500
400
300
200
100
10000
1000

B


54A0
0A0
LSH (nM)
0 6 12 25 50 100
0 6 12 25 50 100
Wells
Free
Nucleosome
Free DNA

C

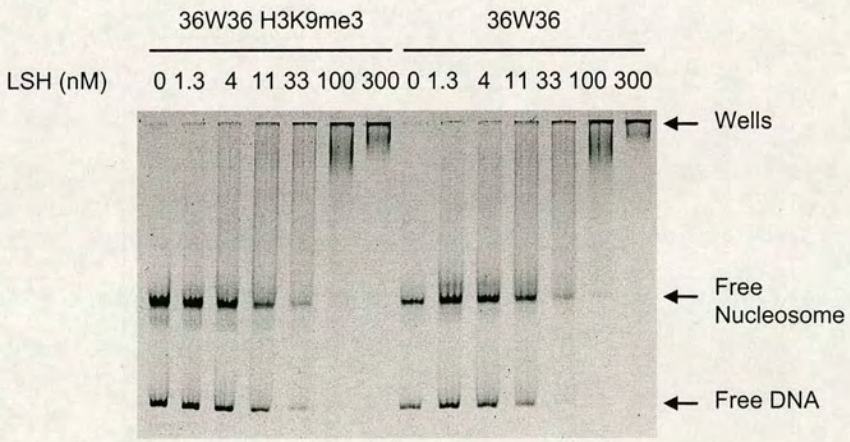

36W36 H3K9me3
36W36
LSH (nM)
0 1.3 4 11 33 100 300 0 1.3 4 11 33 100 300
Wells
Free
Nucleosome
Free DNA

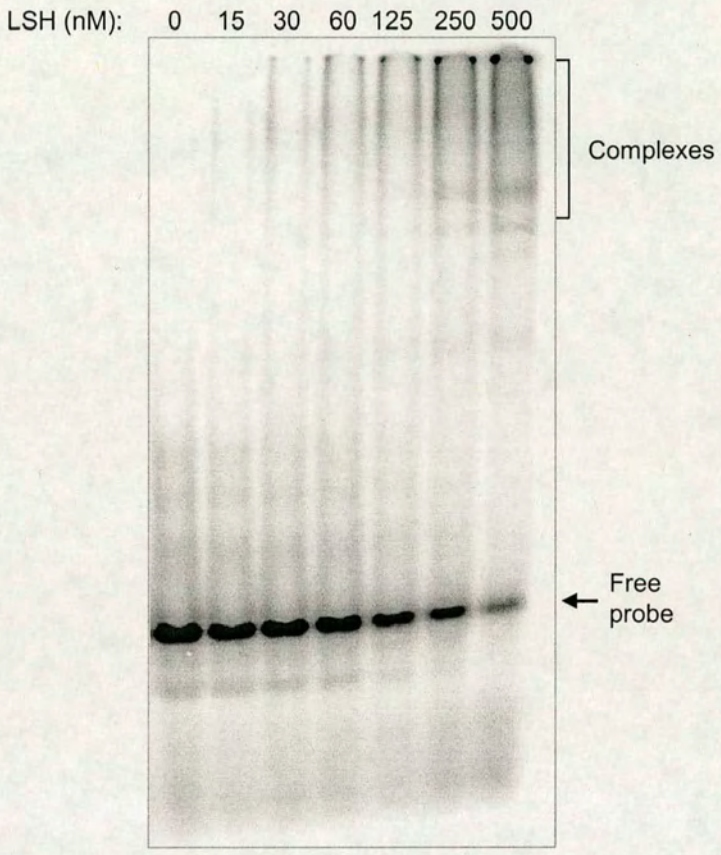
Figure 3.4. LSH binds nucleosomes with linker DNA in EMSA.
A. Native polyacrylamide gel of nucleosome reconstitutions using recombinant core histone octamers and various DNA templates (indicated on the top of the panels). Success of the reconstitutions was judged by depletion of the faster migrating DNA into a slower migrating species. This was judged by running free DNA template on the same gel and comparing its migration to that of reconstituted nucleosomes.
B. This experiment was carried out with the help of Dr Chris Stockdale. EMSA with recombinant LSH and radio labelled 54A0 (left panel) and 0A0 (right panel) nucleosomes. Increasing concentrations of LSH (indicated on the top of the panel) mixed with 10nM nucleosomes for 30 minutes at 4°C. Nucleosome substrate indicated above the gels. Free nucleosomes separated from those bound by LSH by native polyacrylamide gel electrophoresis. Following electrophoresis, the gel was dried, exposed to a phosphor imager screen and developed using a Storm PhosphorImager (GE Healthcare). Free DNA, free nucleosome and LSH:DNA/nucleosome complexes (Wells) indicated to the right.
C. This experiment was carried out with the help of Dr Helder Ferreira. EMSA with recombinant LSH and fluorescently labelled 36W36 H3K9me3 (left panel) and 36W36 (right panel) nucleosomes. Increasing concentrations of LSH (indicated on the top of the panel) mixed with 10nM nucleosomes for 30 minutes at 4°C. Nucleosome substrate indicated above the gels. Free nucleosomes separated from those bound by LSH by native polyacrylamide gel electrophoresis. Following electrophoresis, the wet gel was developed using a Fuji Phosphoimager FLA-5100. Free DNA, free nucleosome and LSH:DNA/nucleosome complexes (Wells) are indicated to the right.

affinity than unmodified nucleosomes. Recombinant H3 tri-methylated at K9 generated by peptide ligation was used to reconstitute nucleosomes for use in this experiment. I found that the ability of LSH to deplete free probe as a measure of binding ability was unchanged by the presence of H3K9me3 indicating this histone modification does not impact LSH binding to nucleosomes. These binding studies show that LSH has the ability to bind DNA and bind nucleosomes efficiently only when linker DNA is present. Together, this would imply that LSH is primarily a DNA binding protein (Figures 3.3 and 3.4).

3.2.2.3 LSH binds DNA with a K_d of ~100nM

I next determined the dissociation constant (K_d) of the LSH:DNA interaction. I performed EMSA with 10nM of probe and a range of LSH concentrations from 8nM to 500nM. The titration was carried out in the absence of competitor DNA to improve the accuracy of the experiment. Because of this, at higher LSH concentrations, large DNA:protein complexes are visible that have difficulty entering the gel. I included these large complexes in calculations of DNA binding by LSH

A


LSH (nM): 0 15 30 60 125 250 500
Complexes
Free
probe

B

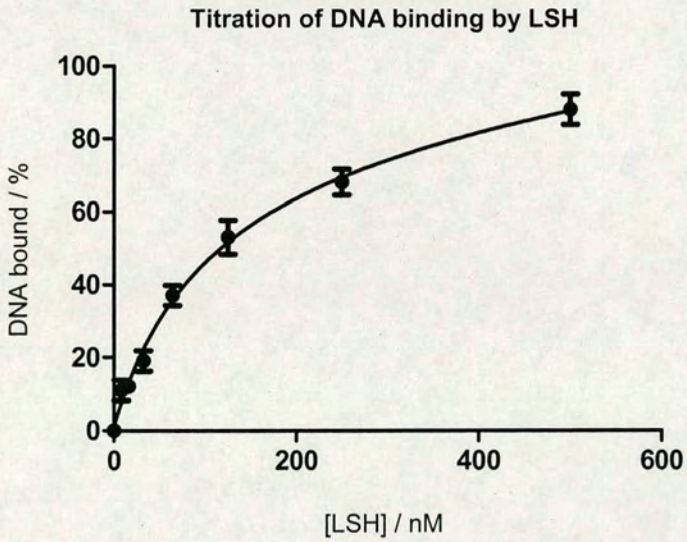

Titration of DNA binding by LSH
DNA bound / %
100
80
60
40
20
0
0
200
400
600
[LSH] / nM

Figure 3.5. Measuring the dissociation constant of the LSH – DNA interaction.
A. EMSA with recombinant LSH and radio labelled 4+5Gs (36bp) DNA probe. Two-fold increasing concentrations of LSH (indicated on the top of the panel) mixed with 10nM probe for 30 minutes at 37°C. Free probe separated from LSH bound DNA by native polyacrylamide gel electrophoresis. Following electrophoresis, the gel was dried, exposed to a phosphor imager screen and developed using a Storm PhosphorImager (GE Healthcare). Band intensity calculated using Image Quant TM software (GE Healthcare). Free probe and LSH:DNA complexes indicated to the right.
B. Saturation curve of LSH binding to 4+5Gs (36bp) DNA probe. Percentage probe bound plotted against LSH concentration. The percentage probe bound was calculated using the following equation with intensities derived using Image Quant TM software (GE Healthcare): (B / (U + B)) x 100. Where B is bound and U is unbound DNA probe. Curve fitting and calculation of K_d performed by non-linear regression using GraphPad Prism software (GraphPad Software). Error bars are the standard deviation from the mean of three independent experiments.

(Figure 3.5A). DNA binding by LSH was calculated as the percentage of total probe shifted (shifted probe / (shifted probe + free probe)) and plotted against LSH concentration. Nonlinear regression was used to calculate the K_d for the LSH:DNA interaction to be 105 ± 45 nM (Figure 3.5A + B). The dissociation constants of a number of SNF2 enzymes interactions with DNA have been published and range from 0.1nM to ~300nM. DNA dependant ATPase A has a calculated K_d of ~0.1nM, SWI/SNF 1nM, ISWI 1-10nM, yeast Rad54 ~100nM and SsoRad54 ~300nM (Fitzgerald et al., 2004; Lewis et al., 2008; Muthuswami et al., 2000; Quinn et al., 1996; Raschle et al., 2004). Thus the strength of the interaction between LSH and DNA is at the lower end of the scale recorded for other SNF2 enzymes but comparable to a number of them. This is not entirely unexpected as recombinant LSH is a monomeric protein whereas SWI/SNF and ISWI are present within multi protein complexes. It is possible that other members of these complexes provide additional DNA binding motifs that strengthen their interaction with DNA.

3.3 LSH is a weak DNA stimulated ATPase *in vitro*

3.3.1 LSH hydrolyses ATP and is stimulated by DNA and nucleosomes with linker DNA

As outlined above the enzymatic activity of SNF2 enzymes *in vitro* gives an indication to their *in vivo* function. I therefore sought to determine whether LSH is an active ATPase and how it is stimulated *in vitro*. To do this I used a highly sensitive ATPase assay that measures the hydrolysis of ATP to ADP + Pi via thin layer chromatography (Whitehouse et al., 2003). I performed the ATPase assay with recombinant LSH in the presence of a range of stimuli and used LSH^{K254Q} in each case as a control (Figure 3.6A + B). I found that LSH showed a low level of ATPase activity, that was stimulated ~2 fold by the addition of DNA or nucleosomes with linker DNA. In a result correlated to the lack of LSH binding to nucleosomes with no linker DNA, LSH activity was not stimulated by 0A0 nucleosomes. Thus LSH appears to be a DNA stimulated ATPase. The nucleosomes used in the ATPase experiments are derived from recombinant core histones and thus lack any post-translational modification. It has been reported that the catalytic activity of the ISWI containing ACF complex is dependant on residues 16-19 of the histone H4 tail and is inhibited by acetylation of H4K16 in particular (Hamiche, et al., 2001; Shogren-Knaak et al., 2006). I thus tested the ability of LSH to hydrolyse ATP in the presence of native nucleosomes purified from chicken erythrocytes (Figure 3.6B). As before, LSH activity was stimulated ~2 fold by these nucleosomes indicating that post-translational modifications of histones do not affect LSH activity. I purified recombinant DDM1 to use as a positive control (Figure 3.6C) and found that it showed higher levels of ATPase activity and was more strongly stimulated by DNA (Figure 3.6D) (Brzeski and Jerzmanowski, 2003). Thus, I concluded that the ATPase assays were working and that LSH has a low level of ATPase activity.

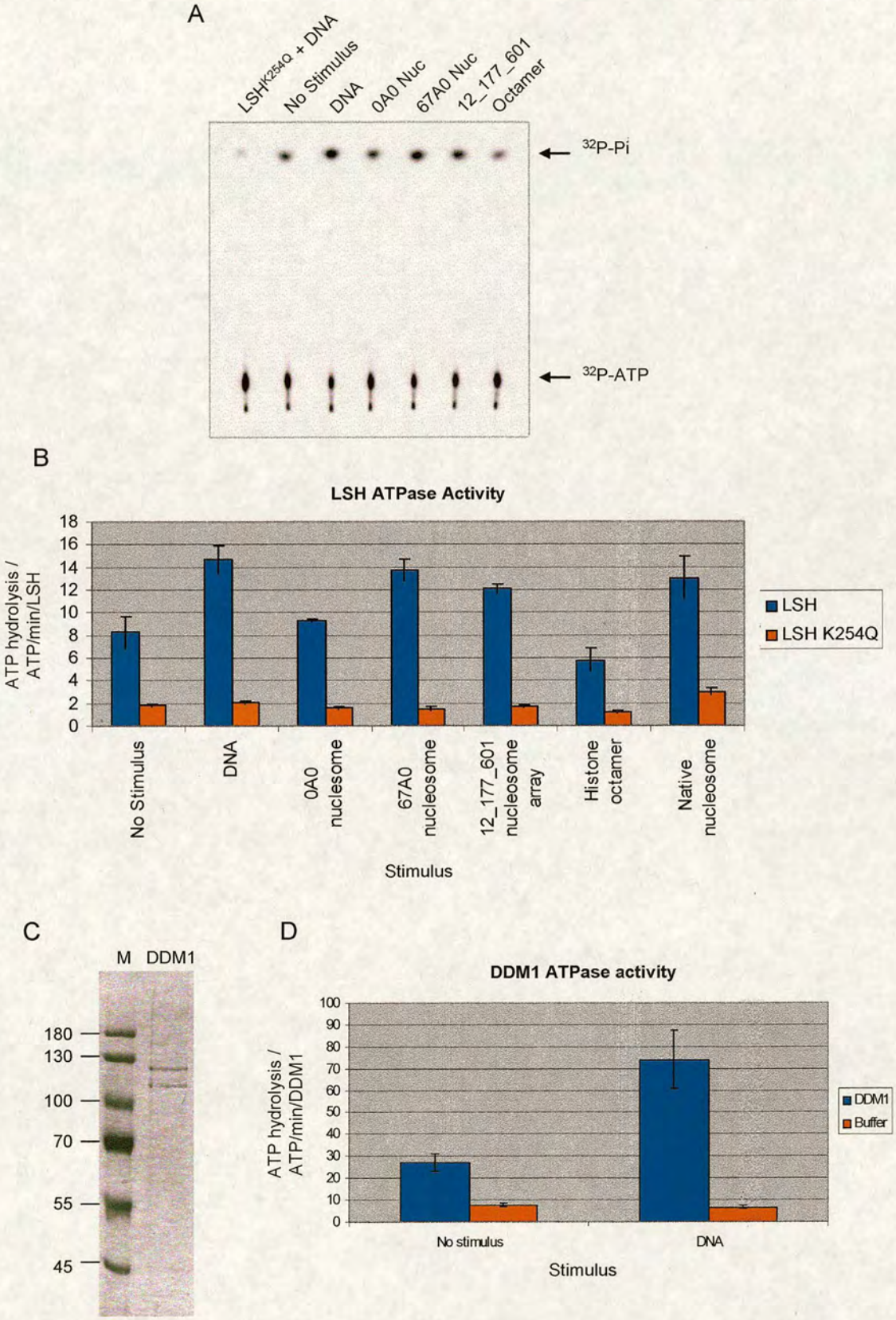
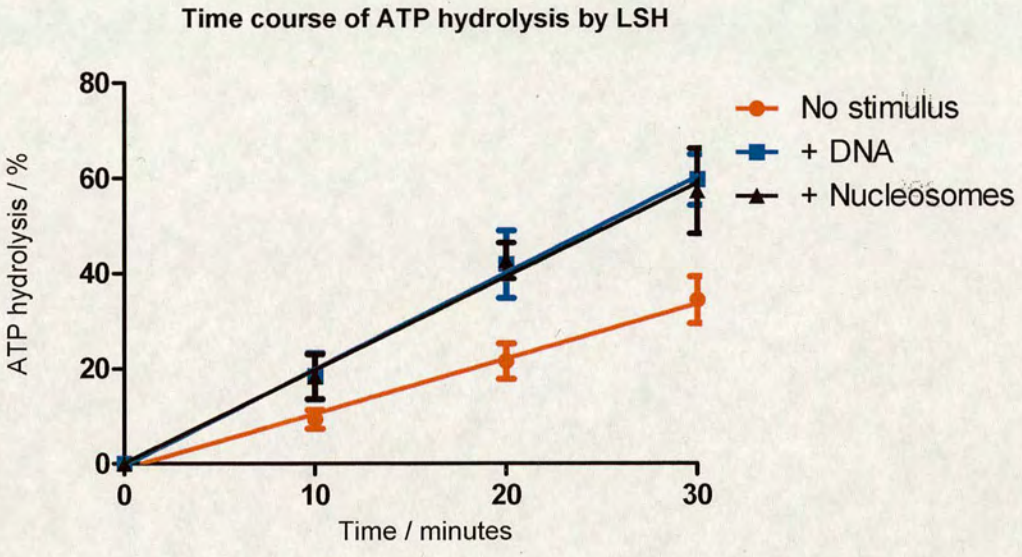

A
LSHK254Q + DNA
No Stimulus
DNA
0A0 Nuc
67A0 Nuc
12_177_601
Octamer
32P-Pi
32P-ATP
B
LSH ATPase Activity
ATP hydrolysis / ATP/min/LSH
18
16
14
12
10
8
6
4
2
0
LSH
LSH K254Q
No Stimulus
DNA
0A0 nuclesome
67A0 nucleosome
12_177_601 nucleosome array
Histone octamer
Native nucleosome
Stimulus
C
M
DDM1
180
130
100
70
55
45
D
DDM1 ATPase activity
ATP hydrolysis / ATP/min/DDM1
100
90
80
70
60
50
40
30
20
10
0
DDM1
Buffer
No stimulus
DNA
Stimulus

Figure 3.6. LSH is a DNA dependent ATPase.
A. Thin layer chromatography plate showing ATP hydrolysis by recombinant LSH or LSH^{K254Q} and a range of stimuli. Reactions (10µl) contained 100nM LSH, 200nM stimulus, 50µM ATP and were incubated for 20 minutes at 37°C. Plates were dried, exposed to Phosphoimager screens and developed using a Storm PhosphorImager (GE Healthcare). Intensity of the spots was calculated using ImageQuant TM software (GE Healthcare). Free phosphate (^{32}P-Pi) and unhydrolysed ATP (^{32}P -ATP) are indicated to the right.
B. Graphical representation of (A). Recombinant LSH (100nM) was used in ATPase assays in the presence of a range of stimuli (indicated on the X axis). The number of ATP molecules hydrolysed per minute per LSH calculated by analysis of thin layer chromatography plates (Y axis). ATP hydrolysis rate observed for LSH (blue bars) and LSH^{K254Q} (red bars) for each stimulus is displayed. Error bars denote standard deviation from the mean of three independent experiments.
C. Coomassie stained gel of recombinant DDM1 purified from insect cells. Size markers (M) indicated to the left.
D. Recombinant DDM1 (100nM) was used in ATPase assays in the absence or presence 200nM DNA (indicated on the X axis). Rate of ATP hydrolysis as ATP molecules hydrolysed per minute per DDM1 indicated on the Y axis. ATP hydrolysis rates observed for DDM1 (blue bars) and buffer only control (red bars) are displayed. Error bars denote standard deviation from the mean of three independent experiments.

I next determined some kinetic parameters of ATP hydrolysis by recombinant LSH. I first determined that ATP hydrolysis was linear over time by performing a time-course experiment (Figure 3.7A). This showed that ATP hydrolysis both in the presence of DNA and without stimulus was linear for at least 30 minutes. Subsequent experiments were carried out for 15 minutes to ensure linearity. I next determined the velocities of the ATP hydrolysis by recombinant LSH in the absence or presence of DNA by titrating ATP (Figure 3.7B). The maximal velocity (V_{max}) and the Michealis-Menton constant (K_m) were determined by non linear regression of the plotted data (Figure 3.7B, GraphPad Prism 5). The V_{max} value for LSH was found to be 12.2 ± 0.25 ATP/min/LSH without DNA and 22.4 ± 1.5 ATP/min/LSH with DNA. The K_m values were found to be 4.6µM and 8.7µM, in the absence and presence of DNA respectively. Thus, the presence of DNA appears to increase the affinity of LSH for ATP and approximately double its rate of turnover. Both the V_{max} and K_m are at least an order of magnitude lower than those calculated for other SNF2 enzymes such as plant DDM1 and mammalian Mi2α indicating that under the conditions tested, LSH is a relatively inefficient ATPase (Brzeski and Jerzmanowski, 2003; Wang and Zhang, 2001b) (Table 3.1).

A


Time course of ATP hydrolysis by LSH
ATP hydrolysis / %
80
60
40
20
0
0
10
20
30
Time / minutes
No stimulus
+ DNA
+ Nucleosomes

B

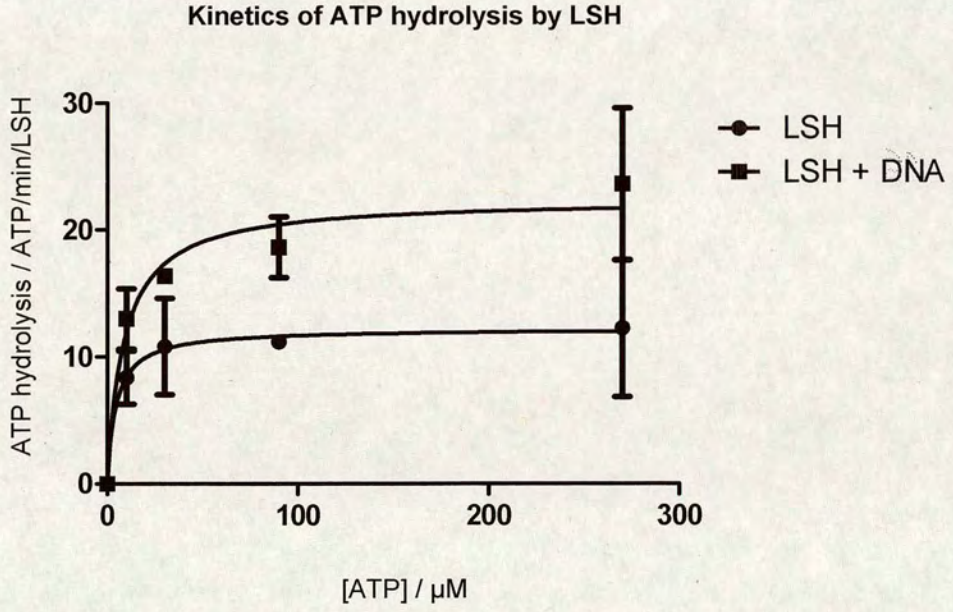

Kinetics of ATP hydrolysis by LSH
ATP hydrolysis / ATP/min/LSH
30
20
10
0
0
100
200
300
[ATP] / µM
LSH
LSH + DNA

Figure 3.7. Kinetic analysis of LSH ATPase activity.
A. TLC plate showing a time course of ATP hydrolysis with no stimulus, DNA and nucleosomes. Reactions (10 µl) contained 100nM LSH, 200nM stimulus and 50µM ATP were incubated at 37°C. Error bars denote standard deviation from the mean of three independent experiments.
B. Recombinant LSH (100nM) was used in ATPase assays in the absence of any cofactor (circles) or in the presence of 200nM free DNA (squares). The assays were performed with increasing concentrations of ATP ranging from 10 µM to 270 µM and the rate of hydrolysis at each concentration plotted. Curve fitting and calculation of V_{max} and K_m performed by non-linear regression using GraphPad Prism software (GraphPad Software).

Protein	K_d / nM (DNA)	Stimulated V_{max} / ATP/min/mol	Translocase distance / bp	References
Mi2α	ND	~200	ND	(Wang and Zhang, 2001a)
NuRD	ND	~4000	ND	(Wang and Zhang, 2001a)
DDM1	ND	~400 This study >70	ND	(Brzeski and Jerzmanowski, 2003) and this study
SWI/SNF	1	~1000	~100	(Logie et al., 1999; Zhang et al., 2006a)
ISWI	1-10	>75	40	(Corona et al., 1999; Whitehouse et al., 2003)
RAD54	100	~1000	80-300	(Kiianitsa et al., 2002; Li et al., 2007b)
LSH	100	~20	ND	This study

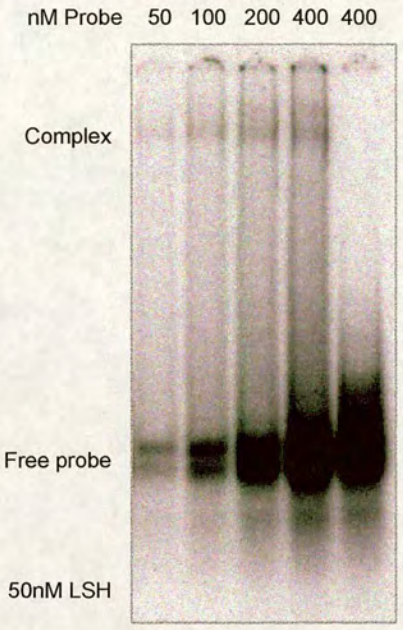
Table 3.1 Catalytic properties of selected SNF2 enzymes.
The catalytic properties of a range of SNF2 enzymes are shown. These include, the K_d of binding to DNA, V_{max} and approximate distance of DNA translocation where known.

3.3.2 Recombinant LSH is ~50% active for DNA binding

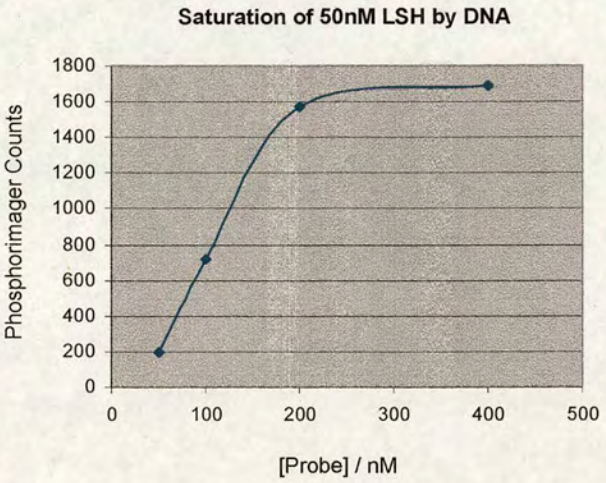
The low level of ATP hydrolysis measured for LSH could be due to a number of reasons. It is possible that LSH is simply not a particularly efficient ATPase and this

level of activity is sufficient for its *in vivo* function. Another alternative is the recombinant protein I have produced is not as active as the native protein. A further possibility is that a co factor required for full ATPase activity is missing. I first asked whether recombinant LSH was active and chose to assess this by determining the proportion of my protein that can bind DNA. To achieve this I first determined the amount of DNA required to fully saturate 50nM LSH (Figure 3.8A). This initial step is required because in the second step, the concentrations of both LSH and DNA need to be above the affinity constant of their interaction (Chmiel et al., 2006). By quantifying the shifted probe at the different concentrations it was determined that 200nM of DNA was sufficient to saturate LSH binding (Figure 3.8B). Using this concentration of probe a titration of LSH concentration was used to saturate DNA binding and thus determine the active fraction of the purified protein. The vast majority of probe was depleted with ~400nM LSH (Figure 3.8C and D). This indicates that ~50% of recombinant LSH is active for DNA binding or LSH is binding DNA as a dimer. In chapter four I describe experiments that did not detect native or recombinant LSH as a dimer biochemically. I therefore think it unlikely that LSH is binding DNA as a dimer and that ~50% of the recombinant protein is inactive. The implications of this experiment are twofold. Firstly, the rate of ATP hydrolysis by LSH should be recalculated accordingly to take into account the inactive fraction, but this level of ATPase activity is still relatively low. Secondly, it is important to consider the possibility that the potentially inactive fraction of LSH is acting as a dominant negative mutant thus explaining the relatively low level of ATPase activity. Future work to determine if inactive LSH acts as a dominant negative mutant could be carried out by spiking the ATPase assays with increasing concentrations of LSH^{K254Q} and assessing inhibition of ATPase activity.

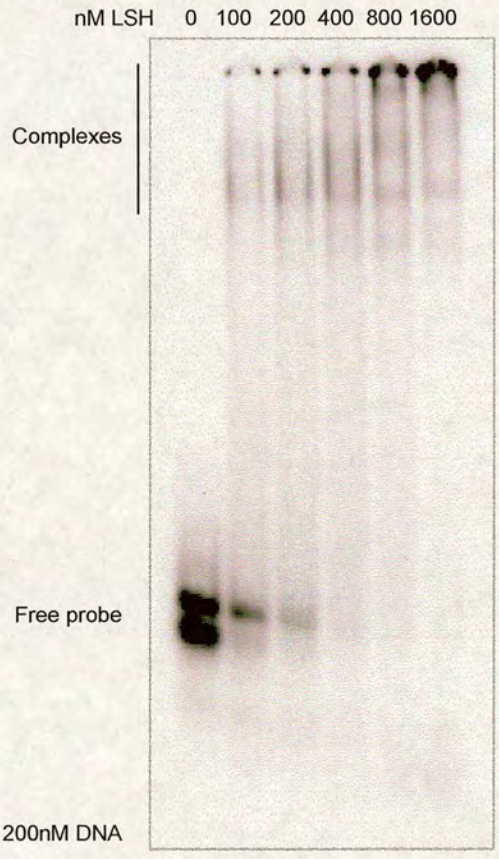
A


nM Probe 50 100 200 400 400
Complex
Free probe
50nM LSH

B


Saturation of 50nM LSH by DNA
Phosphorimager Counts
1800
1600
1400
1200
1000
800
600
400
200
0
0
100
200
300
400
500
[Probe] / nM

C


nM LSH 0 100 200 400 800 1600
Complexes
Free probe
200nM DNA

D

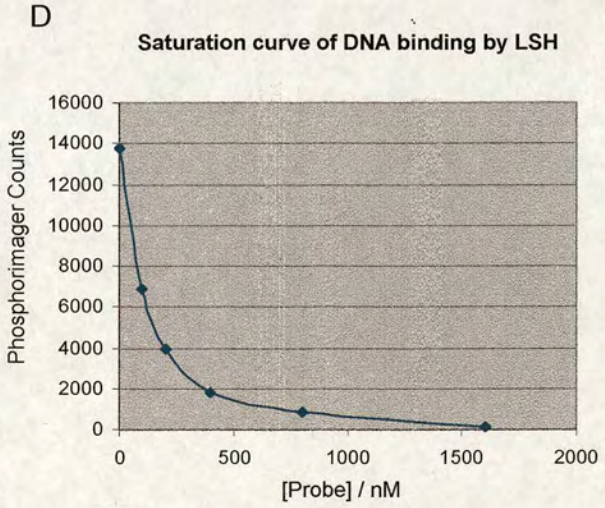

Saturation curve of DNA binding by LSH
Phosphorimager Counts
16000
14000
12000
10000
8000
6000
4000
2000
0
0
500
1000
1500
2000
[Probe] / nM

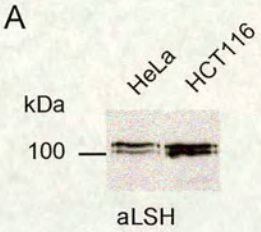
Figure 3.8. At least half of purified recombinant LSH is active for DNA binding.
A. EMSA with recombinant LSH and radio labelled 4+5Gs (36bp) DNA probe. 50nM LSH was mixed with increasing concentrations of radio labelled DNA (indicated on the top of the panel) and incubated for 30 minutes at 37°C. Free probe was separated from LSH bound DNA by native polyacrylamide gel electrophoresis. Following electrophoresis, the gel was dried, exposed to a phosphor imager screen and developed using a Storm PhosphorImager (GE Healthcare). Band intensity calculated using ImageQuant TM software (GE Healthcare).
B. Saturation curve of LSH binding to DNA. PhosphorImager counts of shifted LSH:DNA complex plotted against probe concentration.
C. EMSA with recombinant LSH and radio labelled 4+5Gs (36bp) DNA probe. Increasing concentrations LSH (indicated on the top of the panel) were mixed with 200nM of radio labelled DNA and incubated for 30 minutes at 37°C. Free probe was separated from LSH bound DNA by native polyacrylamide gel electrophoresis. Following electrophoresis, the gel was dried, exposed to a phosphor imager screen and developed using a Storm PhosphorImager (GE Healthcare). Band intensity calculated using ImageQuant TM software (GE Healthcare).
D. Saturation curve of LSH binding to DNA. PhosphorImager counts of free probe plotted against LSH concentration

3.3.3 Mass spectrometry of recombinant LSH detects phosphorylation at serines 115 and 503.

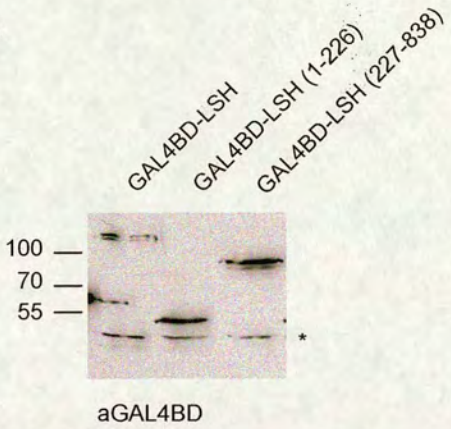
Western blot experiments show native LSH is detected by Western blot from a number of cell lines as a doublet (Figure 3.9A). Interestingly, when over expressing different GAL4BD tagged LSH deletion constructs I also observed the appearance of doublet bands that appear to map to amino acids 227-838 (Figure 3.9B). As this doublet is observed following overexpression of LSH cDNA this indicates that it could be due to a post translational modification of LSH. I was interested to test whether this potential modification could have some impact on activity of LSH so in collaboration with Dr Juri Rappsilber had recombinant LSH analysed by mass spectrometry. Interestingly, mass spectrometry identified two phosphorylated peptides: ETIELSPTGRPK, phosphorylated on S6 and GKNSIDASEEKPVMR, phosphorylated on S4. These phosphorylation sites map to serines 115 and serine 503 respectively. The phosphorylation at S503 was of particular interest as this residue resides in the linker region between helicase motifs III and IV (Figure 3.9C). This linker region is extended in SNF2 enzymes compared to other SF2 helicases and is one of the defining characteristics of the SNF2 family. Interestingly, the part of this region in which S503 resides is longer in LSH subfamily members than other SNF2 ATPases and S503 is found only in certain members of this subfamily (Figure 3.9D;

red box) (Flaus et al., 2006). An attractive hypothesis would be that phosphorylation of this serine could modulate the activity of LSH *in vivo*. I thus attempted to probe whether S503 is a phosphorylation site in mammalian cells also. To achieve this I expressed GAL4BD-LSH mutants where this serine is mutated to alanine (S503A) or aspartic acid (S503D) and thus no longer able to be phosphorylated in HCT116 cells. I made NE from these cells and probed the extracts with antibodies specific to GAL4BD and LSH. Unfortunately, GAL4BD-LSH^{S503A} and GAL4BD-LSH^{S503D} still showed a doublet migration pattern identical to GAL4BD-LSH (Figure 3.9E). This indicated that either this serine is not subject to phosphorylation in these cells or the change in charge due to its loss is not sufficient to alter the migration of LSH. It is also possible that this site is phosphorylated *in vivo* but S115 is also phosphorylated and both would need to be mutated to remove the doublet pattern. Mutation of the second site proved problematic so this issue remains unresolved.

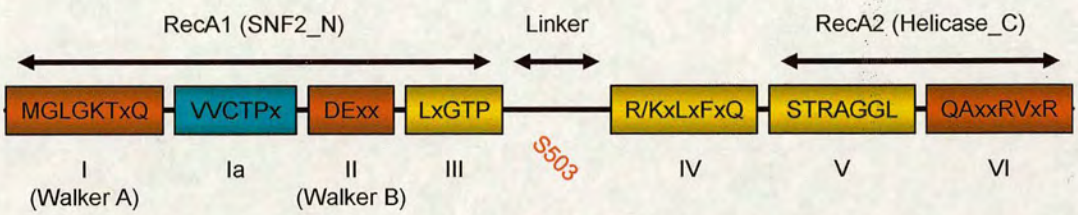
A



B



C



D

	Q–I	I–Ia	Ia–II	II–III	III–B	B–C	C–K	K–IV	IV–V	V–VI
Snf2	17 (0.0)	24 (0.1)	66 (0.7)	25 (0.3)	63 (1.8)	65 (3.2)	26 (4.0)	18 (0.0)	46 (0.2)	21 (0.0)
Iswi	17 (0.0)	24 (0.0)	68 (7.2)	24 (0.1)	53 (1.7)	61 (1.2)	21 (0.3)	18 (0.0)	46 (2.3)	21 (0.0)
Lsh	17 (0.0)	23 (0.3)	70 (7.7)	24 (0.0)	65 (11.7)	110 (33.6)	23 (4.4)	18 (1.8)	46 (1.2)	21 (0.0)
ALC1	17 (0.0)	24 (0.0)	69 (4.2)	24 (0.4)	53 (2.9)	61 (2.9)	19 (2.5)	18 (0.0)	47 (5.3)	21 (0.0)
Chd1	17 (0.5)	24 (0.0)	74 (8.3)	24 (0.0)	49 (4.2)	61 (0.4)	24 (2.4)	18 (0.0)	46 (0.0)	21 (0.0)
Mi-2	17 (0.0)	24 (0.5)	85 (3.7)	24 (0.0)	49 (0.8)	61 (2.8)	26 (1.9)	18 (0.0)	46 (0.1)	21 (0.0)
CHD7	17 (1.8)	23 (0.3)	78 (6.0)	24 (0.2)	49 (0.2)	62 (1.7)	33 (5.2)	18 (0.1)	46 (0.3)	21 (0.0)
Swr1	17 (0.0)	24 (0.0)	67 (0.0)	24 (0.0)	64 (8.0)	60 (0.0)	484 (306.5)	18 (0.0)	45 (0.0)	21 (0.0)
EP400	17 (0.0)	40 (0.0)	51 (0.0)	24 (0.0)	53 (0.0)	60 (0.0)	469 (37.9)	18 (0.0)	45 (0.0)	24 (0.0)
Ino80	17 (0.0)	24 (0.0)	73 (2.4)	24 (0.0)	59 (0.2)	69 (4.3)	280 (27.2)	18 (0.0)	45 (0.2)	21 (0.0)
Etl1	17 (0.2)	24 (2.4)	71 (4.7)	24 (0.4)	60 (3.4)	67 (18.5)	68 (2.5)	18 (0.7)	45 (1.4)	21 (0.0)
Rad54	26 (15.0)	29 (3.0)	71 (4.3)	24 (0.0)	64 (2.4)	60 (2.4)	36 (10.6)	19 (1.8)	47 (1.5)	20 (0.0)
ATRX	25 (3.8)	25 (1.9)	87 (10.4)	24 (0.0)	49 (0.6)	75 (8.6)	105 (37.1)	18 (1.5)	65 (8.4)	21 (0.0)
Arip4	25 (2.0)	25 (5.9)	118 (16.3)	24 (0.0)	64 (0.0)	59 (1.5)	101 (36.8)	18 (0.0)	61 (7.1)	21 (0.0)
DRD1	26 (6.0)	42 (0.5)	76 (13.9)	24 (0.0)	67 (11.6)	59 (3.7)	32 (2.6)	18 (0.4)	51 (0.7)	21 (0.0)
JBP2	23 (11.8)	31 (15.5)	86 (13.1)	24 (1.0)	22 (1.0)	92 (3.6)	75 (20.7)	18 (0.5)	46 (3.0)	21 (0.0)
Rad5/16	49 (24.1)	68 (58.2)	79 (16.9)	24 (0.0)	62 (15.7)	70 (1.8)	123 (28.1)	20 (1.9)	48 (5.2)	20 (0.0)
Ris1	22 (8.4)	50 (31.7)	97 (20.8)	24 (0.3)	54 (3.1)	72 (9.8)	152 (58.6)	30 (17.9)	45 (1.9)	21 (0.0)
Lodestar	19 (5.7)	44 (19.8)	78 (10.9)	24 (0.4)	47 (3.5)	97 (17.4)	60 (25.8)	19 (3.2)	46 (0.7)	21 (0.0)
SHPRH	60 (17.7)	185 (134.5)	92 (11.8)	24 (1.6)	59 (22.9)	74 (4.3)	475 (166.6)	21 (6.5)	43 (3.9)	21 (0.0)
Mot1	17 (0.0)	33 (7.6)	65 (3.7)	24 (0.0)	62 (6.6)	70 (6.1)	29 (5.6)	33 (7.0)	48 (0.6)	21 (1.6)
ERCC6	17 (0.4)	32 (10.0)	74 (15.1)	24 (0.0)	63 (6.5)	72 (14.6)	35 (13.2)	18 (2.4)	45 (4.9)	21 (0.1)
SSO1653	17 (0.0)	25 (10.8)	63 (2.3)	24 (1.0)	51 (5.5)	65 (2.5)	12 (2.9)	18 (1.0)	44 (1.6)	21 (0.0)
SMARCAL1	15 (0.7)	19 (0.1)	61 (7.6)	26 (1.3)	59 (3.1)	65 (8.8)		21 (3.8)	46 (5.5)	21 (0.4)

E

Figure 3.9. Analysis of potential post translational modification of LSH.
A. Nuclear extracts from HeLa and HCT116 cells were separated by SDS-PAGE and LSH detected with α-LSH antibodies. Size marker in kDa indicated to the left.
B. Nuclear extracts from HCT116 cells over expressing various GAL4BD-LSH constructs (indicated above the panel) were separated by SDS-PAGE and protein detected with α-GAL4BD antibodies. Size marker in kDa indicated to the left. Non-specific band that controls for equal loading indicated (*).
C. Schematic representation of the seven conserved helicase motifs defined by Gorbalenya and Koonin, indicated by roman numerals. The letters within the boxes indicate the conserved amino acids found in these motifs. The two RecA like domains are labelled alongside their ENSEMBL classification. Boxes coloured red denote motifs involved in ATP binding and hydrolysis, the blue box denotes the motif involved in binding DNA. The position of serine 503 within the extended linker region that links the two RecA like domains is indicated (red).
D. Schematic outlining the spacing between the different defined helicase domains in different SNF2 enzymes. Serine 503 of LSH resides in the spacer region between motifs III-B and C-IV. This region is specifically extended in LSH members compared to other SNF2 enzymes (red box). Adapted from (Flaus et al., 2006).
E. Nuclear extracts from HCT116 cells over expressing GAL4BD-LSH serine 503 mutants (indicated above the panel) were separated by SDS-PAGE and protein detected with α-LSH antibodies. Size marker in kDa indicated to the left. Endogenous and GAL4BD-LSH fusions indicated to the right.

Despite being unable to confirm that S503 is a phosphorylation site *in vivo* I reasoned that due to its location within the catalytic domain of LSH it may still have an effect on the activity of recombinant LSH. I thus expressed and purified two LSH mutants where serine 503 is mutated to either alanine or aspartic acid (Figure 3.10A). These mutations are commonly used in studies on the effects of phosphorylation as aspartic acid mimics the negative charge of phosphorylation (Wassmann et al., 2003). I tested the ATPase activity of both of these mutants using DNA as a stimulus. Surprisingly, both of these mutants displayed no ATPase activity above that of the control LSH^{K254Q} (Figure 3.10B). Also, no stimulation of activity upon the addition of DNA was observed. Together, these results indicate that S503 itself is critical to the ATPase function of LSH. Thus, the potential role phosphorylation of this residue plays in modulating LSH activity *in vitro* cannot be investigated using these mutant proteins. Thus I attempted to investigate the role of S503 phosphorylation by dephosphorylating recombinant LSH using λ phosphatase. Analysis of LSH by SDS-PAGE following λ phosphatase treatment did not show a noticeable shift in LSH migration so the effectiveness of this treatment could not be

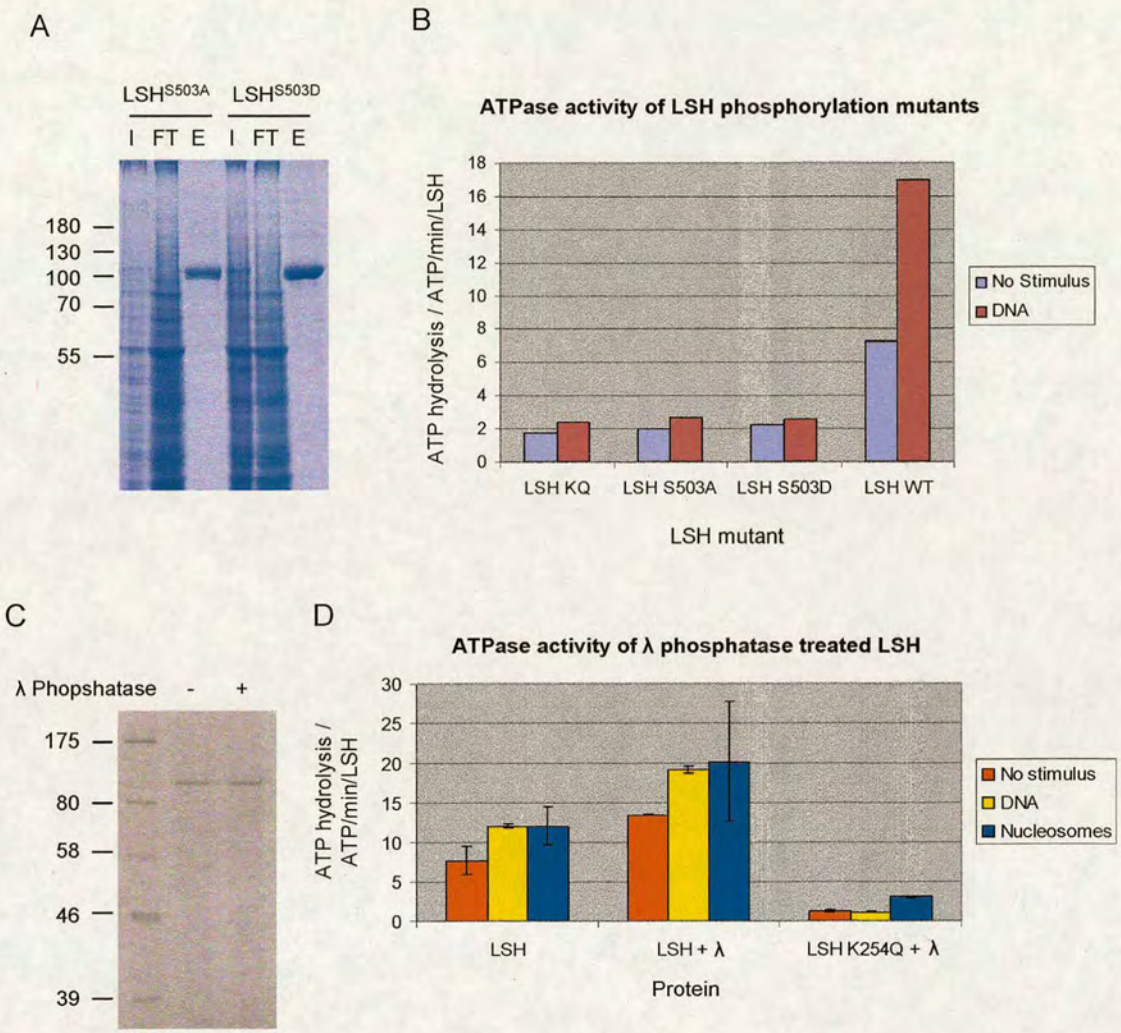


Figure 3.10. Phosphorylation of serine 503 does not inhibit LSH ATPase activity.
A. Coomassie stained gel outlining the purification of HIS-LSH^{S503A} and HIS-LSH^{S503D} via Cobalt affinity chromatography. 5µl cell lysate input (I), 5µl of flow-through (FT) from the Cobalt column and 5µl of elution (E) from the Cobalt column loaded. Size markers in kDa listed to the left.
B. 100nM of a number of mutated, recombinant LSH proteins (indicated on the X-axis) were used in ATPase assays in the absence (blue bar) or presence (purple bar) of 200nM DNA. The number of ATP molecules hydrolysed per minute per LSH calculated by analysis of thin layer chromatography plates (Y axis).
C. Coomassie stained gel of recombinant LSH incubate with λ-phosphatase (+) or mock treated (-) for 1 hour at 37°C. Size markers in kDa listed to the left
D. 100nM of λ-phosphatase or mock treated LSH proteins (indicated on the X-axis) were used in ATPase assays with non stimulus (red bar), 200nM DNA (yellow bar) or 200nM nucleosomes (blue bar). The number of ATP molecules hydrolysed per minute per LSH calculated by analysis of thin layer chromatography plates (Y axis).

assayed by this method (Figure 3.10C). ATPase assays using λ phosphatase treated LSH showed only a modest (~1.5 fold) increase in ATPase activity compared to untreated LSH indicating that the DNA stimulated ATPase activity of LSH is not inhibited by phosphorylation of S503 (Figure 3.10D).

3.3.4 Flag-LSH purified from HeLa cells does not have ATPase activity

As recombinant LSH purified from insect cells has very low activity I wanted to investigate whether LSH purified from human cells was more active. I decided to over express and purify 3xFlag-LSH, 3xFlag LSH K254Q and 3xFlag from HeLa cells as purification using α-Flag M2 antibody is an efficient method of obtaining pure protein. Following overexpression and extraction of nuclear proteins I performed Western blots to confirm presence of 3xFlag tagged proteins (Figure 3.11A). As the proteins were expressed I subjected the nuclear extract to affinity purification using α-Flag M2 affinity gel. I washed the beads extensively, eluted bound protein with 3xFlag peptide and subjected a portion of the eluate to SDS-PAGE. I then silver stained the gel to assess yield and purity. 3xFlag-LSH and 3xFlag-LSH K254Q were clearly visible by silver stain although a number of other bands were also present (Figure 3.11B). Some of these bands were unique to the Flag-LSH sample and were not seen in the Flag control (Figure 3.11C). As these bands could correspond to potential interaction partners of LSH the samples were analysed by mass spectrometry. This analysis identified the protein WD repeat domain 76 as the only unique protein in the Flag-LSH and Flag-LSH K254Q samples. This ~70kDa protein is largely uncharacterised and contains 4 WD40 repeat domains. WD40 repeats are believed to mediate protein – protein interactions and are found in proteins involved in a variety of processes including transcriptional regulation, signal transduction and cell cycle progression (Li and Roberts, 2001). Further investigation of this potential interaction would be an interesting future project. Interestingly, mass spectrometric analysis did not detect any serine/threonine/tyrosine phosphorylation or Arginine methylation sites within purified Flag-LSH. Additionally I did not detect purified Flag-LSH as a dimer perhaps indicating that it is not subject to the same post-translational modifications as native LSH.

I carried out ATPase assays with 3xFlag-LSH using DNA as a stimulus (Figure 3.11D). Surprisingly, 3xFlag-LSH showed no ATPase activity above the two control proteins and its activity was not stimulated by DNA (Figure 3.11D). Thus, overexpressed 3xFlag-LSH purified from HeLa cells is not an active ATPase or has an activity too low to be detected in this assay. This possibility must be recognised as I had to use much lower concentrations of LSH (~2.5nM) in these assays. However, I can conclude that Flag-LSH does not exhibit ATPase activity comparable to that of other SNF2 enzymes as that level of activity would be well within the detection limits of this assay. A further possibility is that purification and elution of 3xFlag-LSH has resulted in a loss of the low level of ATPase activity observed for HIS-LSH. Thus over expressing and purifying HIS-LSH from HeLa cells would perhaps be a viable option in the future.

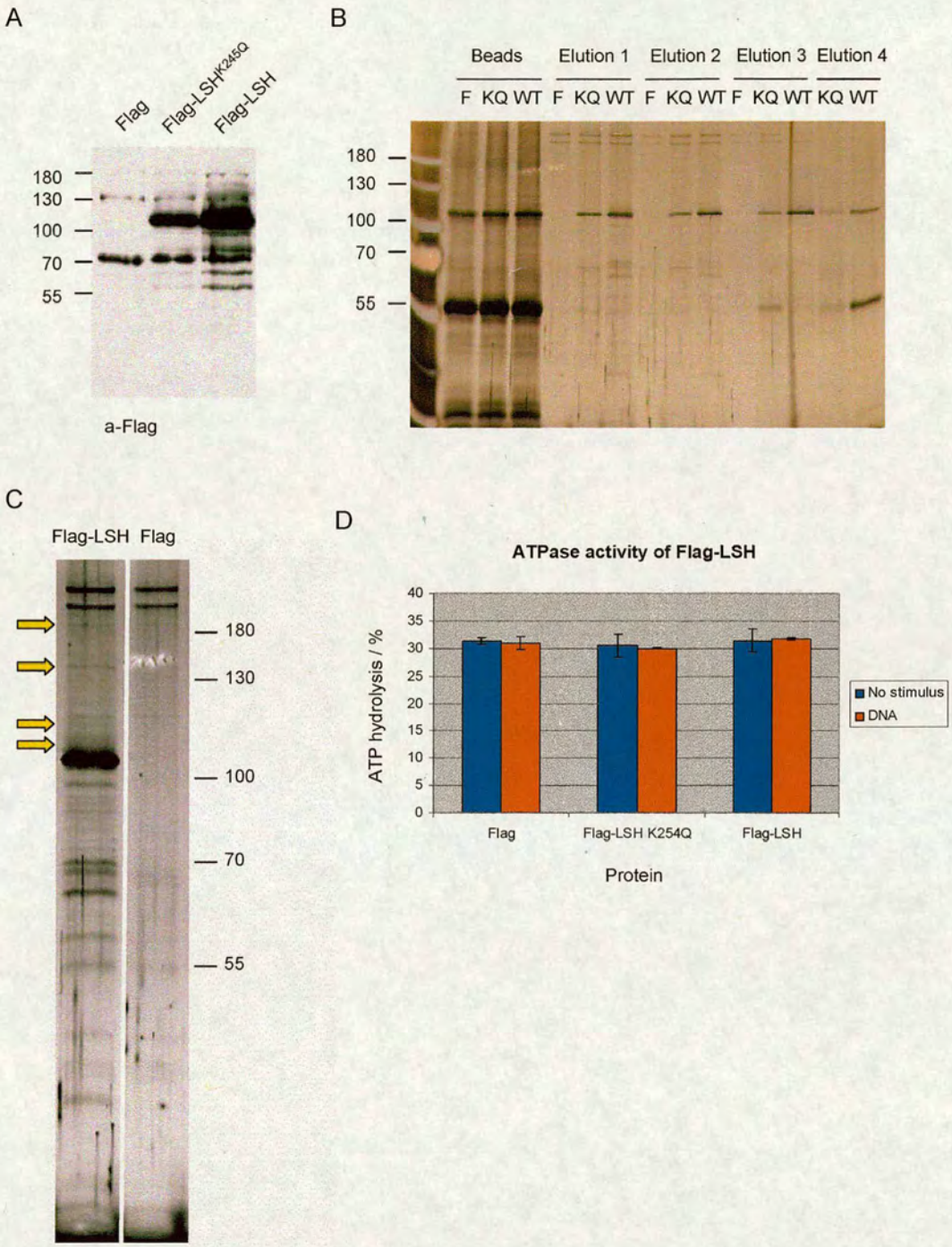


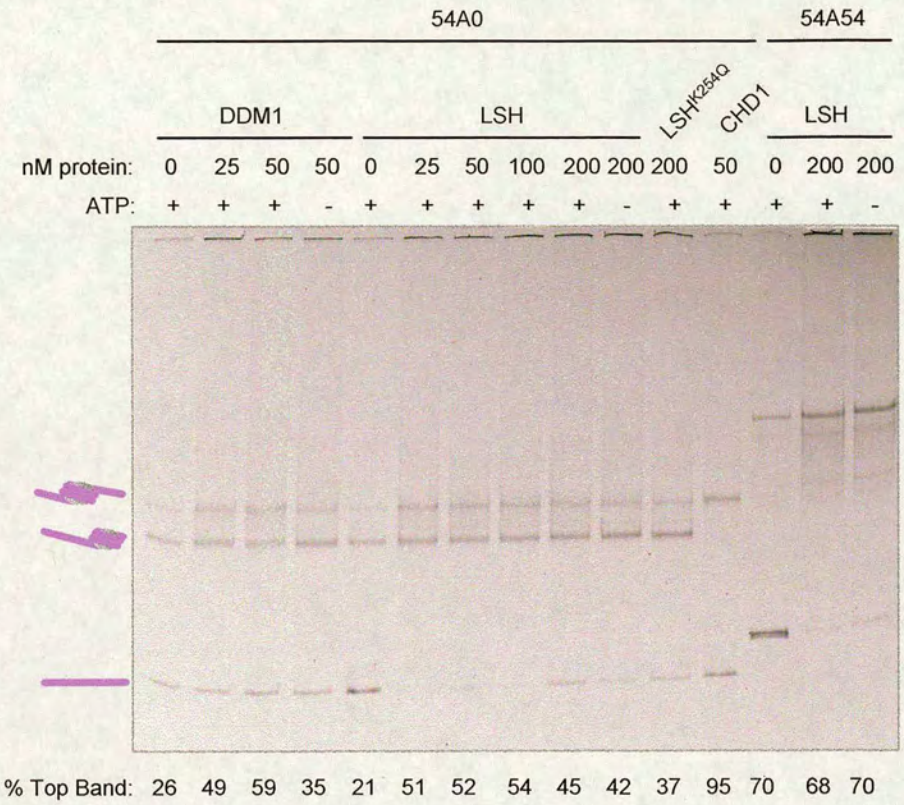
Figure 3.11. Flag-LSH purified from HeLa cells displays no ATPase activity.
A. Nuclear extracts from HCT116 cells transfected with constructs encoding 3xFlag, 3xFlag-LSH or 3xFlag-LSH^{K254Q} (indicated above the panel) were separated by SDS-PAGE and protein detected with α-Flag antibodies. Size markers in kDa indicated to the left.
B. Silver stained gel of M2-Flag Sepharose purified 3xFlag (F), 3xFlag-LSH (WT) or 3xFlag-LSH^{K254Q} (KQ). 10µl of each elution or the beads following 4 elution's was loaded onto the gel.

C. Close up of the silver stained gel of purification from 3xFlag-LSH and 3xFlag expressing HCT116 cells. A number of bands, specific to the 3xFlag-LSH purification are indicated (yellow arrows). Size markers in kDa indicated to the right.
D. 8µl of the eluted α-Flag purified proteins (indicated on the X-axis) were used in ATPase assays in the absence (red bar) or presence (blue bar) of 200nM DNA with 5µM ATP in 10 µl reactions for 15 minutes. The percentage ATP hydrolysed is indicated (Y axis). Error bars denote standard deviation from the mean of three independent experiments.

3.4 LSH does not show detectable nucleosome remodelling activity *in vitro*

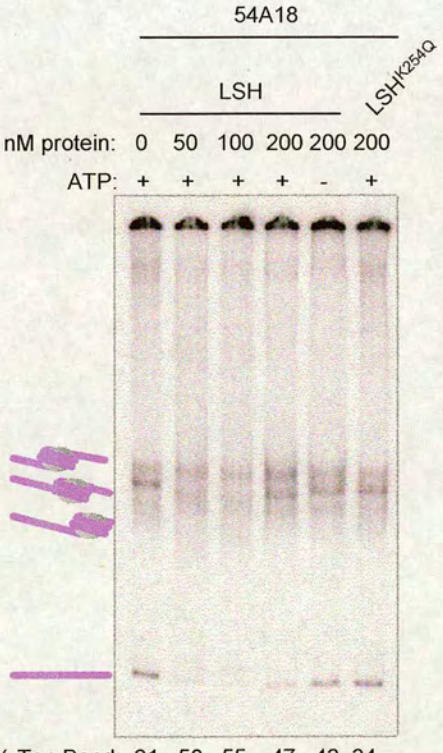
In parallel with the experiments described above I tested the ability of recombinant LSH to remodel chromatin. Nucleosomes at different positions on the same DNA fragment have different mobilities on native polyacrylamide gels (Meersseman et al., 1992). I took advantage of this mobility shift to test for the ability of LSH to reposition nucleosomes. I incubated 10-200nM of LSH with 200nM of 54A0 or 54A54 nucleosomes in the presence or absence of ATP. Following a 1h reaction at 37°C I analysed nucleosome repositioning by native PAGE. I initially chose these nucleosomal templates as they would allow me to assess whether LSH repositions nucleosomes to the end or centre of the DNA template. Addition of LSH led to a slight increase in centrally positioned nucleosomes above the no protein control. This increase was not dependant on LSH concentration or the presence of ATP indicating that LSH does not exhibit ATP dependant nucleosome sliding activity (Figure 3.12A; lanes 5-11). LSH was also unable to slide centrally positioned nucleosomes (Figure 3.12A; lanes 13-15) The positive control protein CHD1 efficiently repositioned the end positioned nucleosomes to the centre of the DNA fragment indicating that the assay had worked (Figure 3.12A; lane 12). I wanted to be sure the low level of ATPase activity exhibited by LSH and its inability to remodel chromatin was not due to problems with my protein purification. I therefore sought to reproduce the nucleosome remodelling activity previously demonstrated for DDM1 using protein I had purified (Brzeski and Jerzmanowski, 2003) (Figure 3.6C). I used this recombinant protein in the same sliding assays with 54A0 nucleosomes and detected a low level of repositioning by DDM1 that is dependant on the presence of ATP and increases with higher amounts of DDM1. The level of repositioning with DDM1 is very low however, particularly compared to that observed with CHD1. Thus, DDM1 appears to be weakly active, indicating my protein purification protocol is working. However, it is important that this experiment is repeated to be sure that this low level of activity is significant and reproducible (Figure 3.12A; lanes 1-4). I reasoned that the lack of remodelling activity detected in the assays with LSH could be explained

A



% Top Band: 26 49 59 35 21 51 52 54 45 42 37 95 70 68 70

B



% Top Band: 31 50 55 47 42 34

Figure 3.12. LSH does not reposition nucleosomes
A. Nucleosome sliding assay with nucleosomes positioned at the end (54A0) or the centre (54A54) of a DNA fragment. Increasing concentrations of DDM1, LSH or LSH^{K254Q} (indicated above the gel) were incubated with 200nM fluorescently labelled nucleosomes in the presence (+) or absence (-) of ATP for 1 hour at 37°C. 5nM CHD1 incubated at 30°C was used as a positive control. Following incubation, reactions were stopped by addition of plasmid DNA and reaction products separated by native polyacrylamide gel electrophoresis. Following electrophoresis, the wet gel was developed using a Fuji Phosphoimager FLA-5100. Band intensity was calculated using Aida software (FujiFilm) and the proportion of the slower migrating (centrally positioned nucleosome) band in each reaction indicated below the gel. The migration of central and end positioned 54A0 nucleosomes are indicated to the left. Non nucleosomal free DNA is also indicated (purple line).
B. Nucleosome sliding assay with 54A18 nucleosomes. Increasing concentrations of LSH or LSH^{K254Q} (indicated above the gel) were incubated with 0.5nM radiolabel nucleosomes in the presence (+) or absence (-) of ATP for 1 hour at 37°C. Following incubation, reactions were stopped by addition of plasmid DNA and reaction products separated by native polyacrylamide gel electrophoresis. Following electrophoresis, the gel was dried, exposed to a phosphor imager screen and developed using a Storm PhosphorImager (GE Healthcare). Band intensity calculated using ImageQuant TM software (GE Healthcare) and the proportion of the slower migrating (centrally positioned nucleosome) band in each reaction indicated below the gel. The migration of off centre, central and end positioned 54A18 nucleosomes are indicated to the left. Non nucleosomal free DNA is also indicated (purple line).

by its relatively low ATPase activity. I therefore used a more sensitive sliding assay utilising radio labelled nucleosomes to assay for low levels of remodelling activity. This assay is more sensitive as it allows the use of much lower concentrations of nucleosomes. However, using 0.5nM of 54A18 nucleosomes I was still unable to detect nucleosome sliding in this assay that was dependant on the presence of ATP (Figure 3.12C). This assay showed a similar result as the previous one with no ATP dependant nucleosome sliding activity by LSH. Together, these data indicate that although able to hydrolyse ATP, LSH does not show detectable levels of nucleosome remodelling in classical sliding assays *in vitro*.

It is possible that LSH uses ATP hydrolysis to remodel histone:DNA interactions in a manner that does not result in nucleosome sliding. I tested this hypothesis by assessing the accessibility of nucleosomal DNA to the bacterial DNMT SssI in the presence of LSH (Figure 3.13A). This assay uses a methyl sensitive restriction digestion step by AciI to assess methylation by SssI. The digestion pattern of the isolated 67A0 nucleosomal DNA following SssI incubation should be indicative of the exposure of naked DNA as SssI is inefficient at methylating nucleosomal DNA

(Gal-Yam et al., 2006) (Figure 3.13A and B). This assay should be able to detect chromatin remodelling and the generation of DNA loops on the nucleosome surface even in the absence of repositioning (Figure 3.13A). I first assessed the protection offered by the nucleosome against SssI methylation by comparing the activity of SssI towards 54A18 DNA or nucleosomes. I first assessed the protection offered by the nucleosome against SssI methylation by comparing the activity of SssI towards 67A0 DNA or nucleosomes. A restriction map of the nucleosomal template is shown to aid interpretation of this experiment (Figure 3.13B). I found that assembly of this sequence into nucleosomes offered protection from the methyltransferase as judged by the increased AciI digestion of the DNA isolated from nucleosomes (Figure 3.13C). However, I did note that there appeared to be less DNA in some of the samples recovered from the nucleosome reactions (Figure 3.13C). The reason for this is unclear but may be due to inefficiencies in the phenol extraction methods used to isolate this DNA. As the nucleosome appeared to be conferring protection to the DNA from SssI activity I next tested whether LSH could expose the DNA. However, I was unable to detect an increase in SssI methylation in the presence of LSH indicating that it is not generating accessible DNA on the surface of nucleosomes (Figure 3.13C).

The experiments described above show that LSH is an active ATPase that is stimulated by DNA. The level of ATPase activity is low compared to other SNF2 enzymes, as is the relative stimulation by DNA. I have not been able to detect nucleosome remodelling activity by LSH using two distinct assays suggesting that recombinant LSH does not possess such an activity.

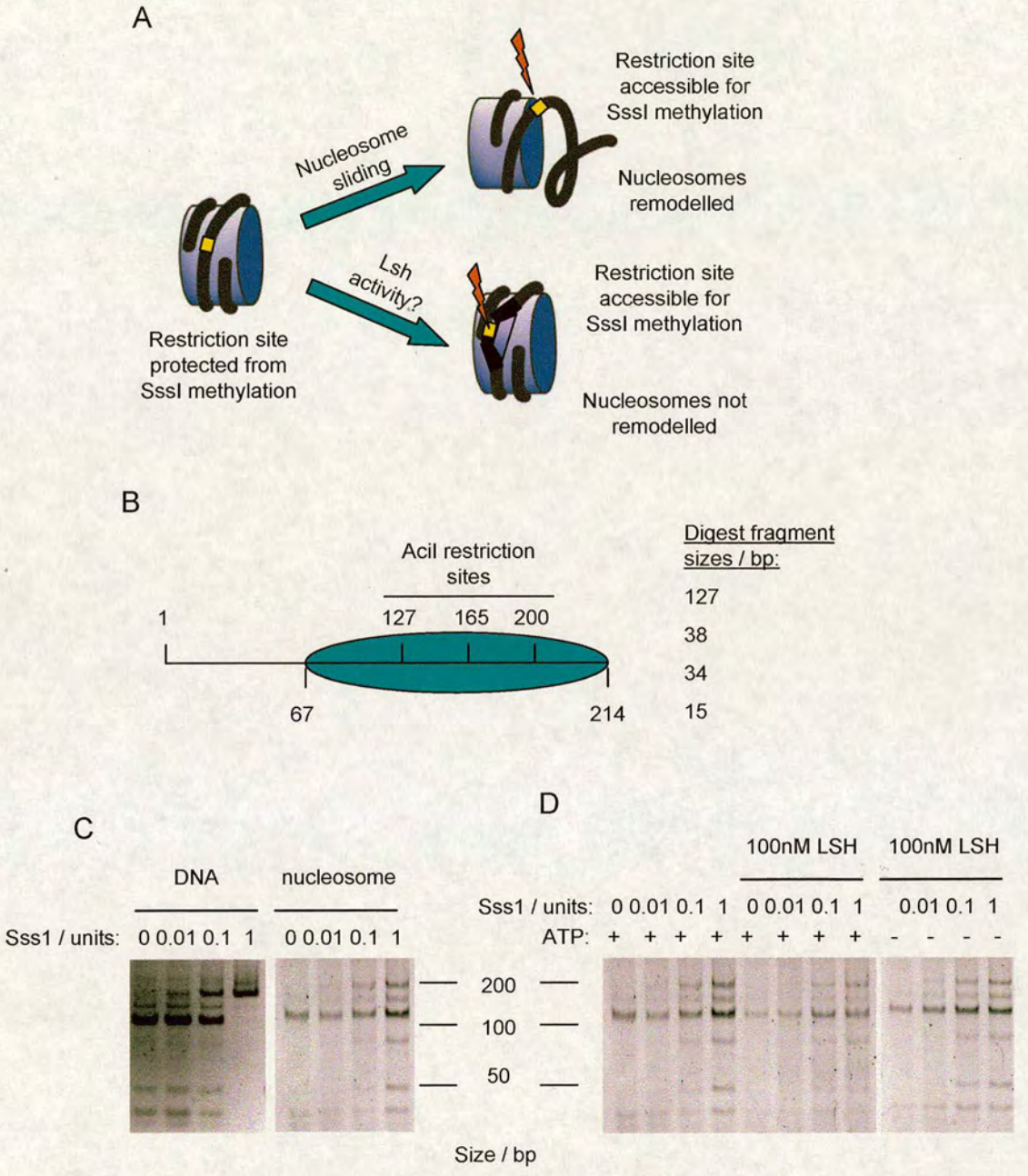


Figure 3.13 LSH does not increase the accessibility of nucleosomal DNA to SssI methyltransferase.

A. Schematic outlining the SssI methyltransferase assay. Chromatin remodelling by nucleosome sliding *or* disruption of DNA:histone contacts may lead to exposure to CpG sites. If methylated the site becomes resistant to restriction enzyme digestion.

B. Diagram of the 67A0 nucleosome used for this assay. AciI digestion sites are indicated on the DNA. The positioned nucleosome, which should protect the AciI sites from methylation is

indicated in green. The size of the expected digestion products following AciI digestion are indicated.

C. 100nM 67A0 DNA or 67A0 nucleosomes (indicated) were incubated with increasing amounts of SssI methyltransferase for 15 minutes at 37°C. Following incubation, the reaction was stopped and DNA liberated by two phenol:chloroform and a single chloroform extraction step. DNA was ethanol precipitated, washed and resuspended in 20µl of TE. The DNA was digested with 10 units AciI in a reaction volume of 30 µl and 10 µl separated by native polyacrylamide gel electrophoresis. The gel was stained with SYBR-gold (Invitrogen) and imaged using a transilluminator.

D. 100nM 67A0 nucleosomes were incubated with increasing amounts of SssI methyltransferase with or without 100nM LSH for 15 minutes at 37°C. Following incubation, the reaction was stopped and DNA liberated by two phenol:chloroform and a single chloroform extraction step. DNA was ethanol precipitated, washed and resuspended in 20µl of TE. The DNA was digested with 10 units AciI in a reaction volume of 30 µl and 10 µl separated by native polyacrylamide gel electrophoresis. The gel was stained with SYBR-gold (Invitrogen) and imaged using a transilluminator.

3.5 Summary

The plant SNF2 family protein DDM1 and its mammalian homologue Lsh were originally identified as proteins essential for high levels of DNA methylation *in vivo* (Dennis et al., 2001; Jeddeloh et al., 1999; Vongs et al., 1993). The roles of these proteins in DNA methylation has been characterised genetically using plants lacking DDM1 and mice deficient for Lsh. Further experiments using cell lines lacking Lsh has shown that Lsh is required for *de novo* but not maintenance methylation of replicating episomal plasmids and can interact with Dnmt3a and Dnmt3b (Zhu et al., 2006). This indicates that Lsh has a direct role in *de novo* methylation of DNA (Dennis et al., 2001; Zhu et al., 2006). The biochemical role of LSH and how this relates to its function *in vivo* are currently unknown. As a member of the SNF2 family of chromatin remodelling enzymes an attractive hypothesis is that the enzymatic activity of LSH is related to its role *in vivo*. SNF2 enzymes use ATP hydrolysis to disrupt chromatin structure in a manner that alters the accessibility of DNA (Narlikar et al., 2002). It could be envisaged that if LSH also has this function it could alter the accessibility of nucleosomal DNA to DNMT enzymes. In agreement with this hypothesis, data on the enzymatic activity of DDM1 indicates that it has the ability to slide nucleosomes *in vitro*. As the information gained from biochemical studies would provide useful insights into the role of LSH, I attempted to elucidate its function *in vitro*.

I was able to obtain highly pure, recombinant LSH and LSH^{K245Q} from insect cells using a baculovirus expression system. I used the recombinant proteins to show that LSH binds DNA and mononucleosomes with linker DNA. This binding was not seen with mononucleosomes without linker DNA suggesting that LSH is predominantly a DNA binding protein. I determined the K_d for the interaction between LSH and DNA to be ~100nM. In comparison with the K_ds of other SNF2 enzymes and DNA, such as SWI/SNF and ISWI, this interaction appears rather weak. As mentioned above this may not be entirely surprising as the K_ds for these proteins was determined for the protein complexes these enzymes reside in (Fitzgerald et al., 2004; Lewis et al.,

2008; Muthuswami et al., 2000; Quinn et al., 1996; Raschle et al., 2004). Thus, different members of these complexes may contribute towards DNA binding. How the DNA binding efficiencies of SNF2 enzymes relate to other catalytic functions or *in vivo* roles is unknown. Something that does not appear to be directly linked to efficiency with which these enzymes bind DNA is their ability to hydrolyse ATP. The DNA binding constants and maximal rates of ATP hydrolysis for a number of SNF2 enzymes shown above display no correlation between K_d and V_{max} (Table 3.1). Thus, the affinity of SNF2 enzyme for DNA does not appear to be directly linked to their ATPase activity. As some of these enzymes function as DNA translocases it is possible that the affinity of the SNF2 protein for DNA is linked to their ability to translocate DNA. Limited data is available on the kinetics of DNA translocation by SNF2 enzymes but no correlation has been detected between DNA translocation and affinity for DNA (Table 3.1). It is possible that the ability of SNF2 enzymes to bind DNA has a role in their targeting to genomic loci or their ability to remodel chromatin but this is currently unknown. A further caveat to this data is that it is not clear how relevant binding constants derived *in vitro* from EMSAs are to *in vivo* function.

The core SNF2 ATPase domain consists of seven conserved sequence motifs common to the SF2 superfamily of helicases. Together, these motifs are involved in binding to DNA and binding and hydrolysing ATP. A common characteristic of SNF2 enzymes is the ability to hydrolyse ATP when exposed to different substrates such as DNA or chromatin. SNF2 enzymes appear to form two distinct groups based on their ATPase activities. Enzymes such as SWI/SNF are maximally stimulated by DNA, while enzymes such as ISWI and SNF2h require nucleosomes for maximal stimulation (Aalfs et al., 2001; Tsukiyama and Wu, 1995). I tested the ability of LSH to hydrolyse ATP and found that it exhibited a low level of activity that was stimulated ~2 fold by DNA. The same level of stimulation was also seen when nucleosomes with linker DNA were used as a substrate but not when nucleosomes lacking linker DNA were used. Together these data indicate that LSH is a weak DNA stimulated ATPase and correlates well with ability of LSH to bind DNA *in vitro*. These observations also correlate with data on the enzyme kinetics of DDM1 which

has been shown to be a DNA stimulated ATPase (Brzeski and Jerzmanowski, 2003). Thus DDM1 and LSH belong to the DNA stimulated group of SNF2 ATPases. However, the activity of LSH differs from that recorded for other SNF2 enzymes in two ways. The first is the maximal rate of ATP hydrolysis which I calculated by titrating ATP concentration. I found the V_{max} of unstimulated LSH to be ~10 ATP/min/LSH and DNA stimulated LSH to be ~20 ATP/min/LSH. These levels are low compared to other SNF2 enzymes including DDM1 which has a V_{max} of ~400 ATP/min/DDM1. The second difference is the low level of stimulation of LSH by DNA which is only ~2 fold. Other SNF2 enzymes tend to be stimulated to a higher degree by their substrate. For example, SWI/SNF and ISWI are stimulated ~5 and ~10 fold by their substrates, respectively (Corona et al., 1999; Laurent et al., 1993). However, the ATPase activity of DDM1 is only ~3 fold higher in the presence of DNA. Thus, although the level of stimulation shown by LSH is low compared to SWI/SNF and ISWI, it is comparable to that of DDM1. This may indicate a functional difference between the LSH subfamily of SNF2 enzymes compared to other subfamilies. On the other hand, it may indicate that the full stimulation of these enzymes has not yet been observed and another co-factor or different substrate may be required.

To ensure that the low level of activity shown by recombinant LSH was not due to the purified protein being unable to bind substrate the fraction of LSH that actively binds DNA was determined using EMSAs. These assays showed that using saturating concentrations of DNA, ~800nM of LSH can completely shift ~400nM of DNA. Thus either ~50% of LSH is active for DNA binding or LSH bind DNA as a dimer. I have not been able to detect LSH as a dimer *in vivo* or *in vitro* using biochemical methods so this possibility seems unlikely (Chapter 4). I cannot completely rule out this possibility but it seems reasonable to conclude that *at least* 50% of recombinant LSH is active for DNA binding. The consequence of having a 50% inactive fraction of protein is that the dissociation constant of LSH should be adjusted accordingly. Thus the Kd of LSH:DNA is ~50nM and the V_{max} of stimulated LSH is ~40ATP/min/LSH. These modest adjustments do no alter the conclusions made previously regarding the weak DNA binding and low level of

ATPase activity observed for LSH. These adjustments may be more significant if the inactive LSH was acting as a dominant negative mutant. However, this is unlikely as the inactive fraction of protein that cannot bind DNA should not interfere with the activity of the active protein unless they dimerise. As mentioned above, I have not seen evidence of dimerisation and if it were the case, then it would appear that 100% of LSH dimers are active for DNA binding. Also, it is unlikely that inactive LSH could prevent DNA binding by the active protein by saturating it as ~20 fold molar excess of DNA was used for stimulating the ATPase activity.

The observation of native and overexpressed LSH running as a doublet band in human cells by Western blot indicates that it may be subject to post translational modification. Mass spectroscopy analysis of purified recombinant LSH detected phosphorylation at serines 115 and 503. The phosphorylation of S503 is particularly intriguing as this serine resides within a region of the SNF2 catalytic domain that is specifically extended in the LSH subfamily (Flaus et al., 2006). S503 is present in LSH homologues in mouse and Xenopus but not in *S. cerevisiae* or DDM1. Therefore if LSH activity is regulated by phosphorylation of this residue, its function is not conserved in plants. Despite being unable to determine if S503 is a phosphorylation target in human cells, I asked whether it may be inhibiting the activity of LSH *in vitro*. Unfortunately, the LSH mutants generated to test this were completely inactive and dephosphorylation of LSH by λ-phosphatase had little effect on its activity. Thus, I was not able to conclusively determine whether phosphorylation of S503 or S115 have an impact on LSH activity *in vitro*. If it were possible to determine whether LSH is phosphorylated at either of these sites *in vivo*, then further work on the effect of this modification in regulating LSH activity may prove fruitful.

Two distinct assays were used to analyse the ability of LSH to disrupt histone:DNA contacts *in vitro*. The first assay was a classical sliding assay that can be used to observe repositioning of nucleosomes on a stretch of DNA. In these experiments no ATP dependant nucleosome sliding activity by LSH was observed. Thus it would appear that although able to hydrolyse ATP, LSH is not able to reposition

nucleosomes *in vitro*. However, it should be cautioned that as the ATPase activity of LSH is much lower than other SNF2 enzymes that do exhibit this activity it may be the case that LSH requires some other stimulus for full ATPase and nucleosome sliding activities. Despite being unable to reposition nucleosomes *in vitro*, it is possible that LSH can alter chromatin structure in some other way. This is the case for ATRX, which has not been reported to slide nucleosomes but can increase the accessibility of nucleosomal DNA *in vitro* (Xue et al., 2003). It could be envisaged that the increase in accessibility is due to formation of DNA loops on the nucleosome surface, as has been demonstrated for ACF (Strohner et al., 2005). The ability of LSH to generate regions of increased DNA accessibility on nucleosomal templates was assayed indirectly using SssI methyltransferase. The activity of this bacterial DNMT is strongly inhibited by nucleosomes (Gal-Yam et al., 2006). Formation of a DNA loop on the surface of the nucleosome would lead to exposed DNA that can be methylated by the SssI. Methylation sensitive restriction digests are then used to assess the levels of DNA methylation. This assay did not detect any changes in DNA methylation in the presence of LSH indicating that LSH was not causing changes in nucleosome structure detectable by this assay. Looking for changes in nucleosome structure using this assay is limited in several ways. Firstly, it can only detect generation of DNA loops if they contain CpG sites. Thus, if LSH is producing a specific alteration in nucleosome structure at a particular region of the nucleosome surface not containing a CpG, this will not be detected. Also, if the DNA loops generated are small and/or very transient, these may not be prone to SssI methylation. Therefore, it would be useful in the future to confirm this result using a higher resolution technique that is not biased to particular DNA sequences. DNase I digestion of nucleosomes is often used to assess changes in nucleosomal DNA accessibility in response to SNF2 enzyme activity and may be informative of LSH activity (Saha et al., 2005; Strohner et al., 2005).

The ability to translocate DNA has been observed for several SNF2 enzymes and complexes including ISWI, SWI/SNF, RSC and ATRX (Saha et al., 2005; Whitehouse et al., 2003; Xue et al., 2003; Zhang et al., 2006b; Zofall et al., 2006). As LSH shows a low level of DNA stimulated ATPase activity but cannot remodel

chromatin it is possible that it may exhibit translocase activity. During the biochemical characterisation of LSH I attempted to test whether LSH exhibits translocase activity by performing a triple helix displacement assay. This involves generating a DNA triple helix between a double helical d(GA) · d(TC) tract and a third strand, consisting of a single stranded homopyrimidine repeat dTC_{20}. The triple stranded DNA forms via Hoogsteen hydrogen bonds between the pyrimidine bases of the single stranded DNA and the purine bases of the double stranded DNA. Displacement of the single stranded DNA from the triple helix is indicative of DNA translocation (Saha et al., 2002). Interestingly, ATRX is able to efficiently disrupt triple helical DNA but is not able to reposition nucleosomes (Xue et al., 2003). I was not able to perform this assay as reconstitution of the triple helix was unsuccessful and so the question of whether LSH displays DNA translocase activity remains to be answered.

The inability of recombinant LSH to remodel chromatin allied to its low level of ATPase activity raises the question of whether recombinant protein is fully active. It is possible that LSH requires a different stimulus or additional co-factors to display its maximal activity. It is worth noting that most SNF2 enzymes reside in multi-subunit complexes and it has been reported that these subunits can modulate enzymatic activity (Langst et al., 1999). Thus, I next sought to determine whether LSH is associated with other proteins *in vivo* and if so, determine their functional significance.

Chapter four - LSH cooperates with DNMTs to repress transcription

4.1 Introduction

SWI2/SNF2 chromatin remodelling enzymes are commonly found associated with at least one other stably bound protein (Narlikar et al., 2002). The specific subunits found within these multisubunit complexes can modulate the activity of the remodelling enzyme and are often involved in their targeting. The number of different associated subunits varies considerably between different classes of remodeler. SWI2/SNF2 family members tend to be found in huge mega Dalton (MDa) sized complexes with many different associated subunits. The yeast RSC complex for example contains at least 14 additional subunits in addition to its SWI2/SNF2 remodeler core (Cairns et al., 1996). ISWI complexes tend to be smaller and contain up to four subunits in addition to the SWI2/SNF2 enzyme (Figure 1.3C). The human ACF complex for example contains just one stably bound protein hACF, whereas huCHRAC contains ISWI, hACF, p17 and p15 (He et al., 2006; Poot et al., 2000). The Mi-2/CHD family is slightly different in that different subfamily members show different protein complex preferences. On the one hand members of the Mi-2/CHD subfamily stably associate with the large NuRD complex (Tong et al., 1998). Conversely members of the CHD1 subfamily do not associate with other proteins and are found predominantly as monomers (Lusser et al., 2005; Tong et al., 1998).

There are three distinct roles for proteins that stably interact with SWI2/SNF2 chromatin remodelling enzymes, modulation of remodeler activity, genomic targeting or concurrent covalent modification of chromatin. The most striking example of a protein cofactor influencing remodeler activity is that of the Acf1 interaction with ISWI in the ACF and CHRAC complexes (Langst et al., 1999). In this example, Acf1 increases both the rate of remodelling and the outcome of the

remodelling reaction by ISWI (Langst et al., 1999). Genomic targeting of remodelling complexes via specific subunits can occur via direct or indirect mechanisms. The BAF200 subunit of the PBAF complex for example is believed to directly target it to genomic loci (Yan et al., 2005). Indirect targeting involves the interaction of stable subunits with transcription factors as is the case for targeting yeast and mammalian SWI/SNF complexes (Lickert et al., 2004; Neely et al., 2002). The subunits of several SWI2/SNF2 complexes also act as covalent modifiers of chromatin. The NuRD complex for example contains histone deacetylase (HDAC) enzymes and forms a large transcriptionally repressive complex (Tong et al., 1998).

The proteins that interact with SWI2/SNF2 enzymes have a critical role in their function. Thus their identification is a crucial step in understanding the role of particular remodelers *in vivo*. As the molecular function of LSH is largely unknown and as it displayed very low ATPase activity *in vitro* I have attempted to determine whether it is present in a stable multiprotein complex. If it were then identification of the associated factors would potentially provide crucial insight into the regulation of its activity, its targeting to chromatin and its role in DNA methylation.

4.2 Results - LSH does not form a stable multisubunit complex

4.2.1 Native LSH is not part of a large stable complex

Before testing whether LSH is part of a stable complex I first characterised a commercial LSH antibody (Santa Cruz: sc-46665). I used the antibody to perform Western blots on nuclear extract (NE) from the colon cancer cell line HCT116 and 6xHIS tagged recombinant LSH produced using the baculovirus system (Figure 4.1A). The antibody detected a doublet band in HCT116 NE at around 100kDa and the single band of the tagged purified protein slightly larger than this indicating the antibody is specific to LSH. To provide further confirmation of the specificity of the antibody I used it to Western blot NE from mouse embryonic fibroblasts (MEFs) derived from wild-type and $LSH^{-/-}$ embryos (Figure 4.1B; top panel, lanes 1 and 2). A single band was detected in wild-type MEFs but not in $LSH^{-/-}$ MEFs. Together these data strongly indicate that this LSH antibody is highly specific to LSH in Western blots. I next assessed the expression of LSH in a number of commonly utilised cell lines; mouse embryonic stem (ES) cells, HCT116, HeLa, H520, H226 and MRC5 (Figure 4.1B; top panel, lanes 3-8). As can be seen from the blot, LSH is expressed in all the tested cell lines but appears to be most strongly expressed in ES cells and least expressed in H226 cells. HDAC2 served as a loading control for this experiment (Figure 4.1B; bottom panel). As LSH was ubiquitously expressed, I decided to initially investigate whether an LSH complex can be detected in the colon cancer cell line HCT116.

In order to investigate whether LSH forms a stable multisubunit complex I used size exclusion chromatography to separate proteins on the basis of size. I used a Superose 6 column from which large proteins and protein complexes elute from the column early and small proteins elute later. The Superose 6 column was first equilibrated using a range of size markers (Materials and Methods).

I initially applied HCT116 nuclear extract to the column and assayed for the presence of LSH in the eluted fractions by Western blot (Figure 4.2A; top panel). As can be

seen, the majority of LSH eluted in fractions 18-19 – with a peak at approximate mass of 150kDa. This is similar to the estimated molecular mass of LSH, 97kDa. I next assayed the integrity of the nuclear extract in 420mM NaCl to make sure this low size was not due to dissociation of protein complexes. To achieve this I reprobed the Western blot with an antibody that specifically recognises the SWI/SNF component Brg1 (Figure 4.2A; middle panel). This experiment revealed that Brg1 eluted in fractions 5-7 corresponding to a molecular mass greater than 669kDa. Brg1 has an estimated molecular mass of (200kDa) and the SWI/SNF complex has an estimated molecular mass of ~2MDa. Thus I conclude that Brg1 is most likely still associated with the intact SWI/SNF complex and lack of NE integrity is unlikely to have caused dissociation of a stable LSH complex. The elution of native LSH at a relatively small size does not rule out the possibility that it stably associates with another protein of ~50kDa. In an attempt to rule out this possibility I compared purified recombinant LSH produced by the baculovirus expression system to the native LSH protein. Western blotting of the eluted fractions showed elution of recombinant LSH from the Superose 6 column in the same fractions as native LSH, indicating that the majority of LSH does not stably associate with other proteins in HCT116 NE (Figure 4.2A; bottom panel).

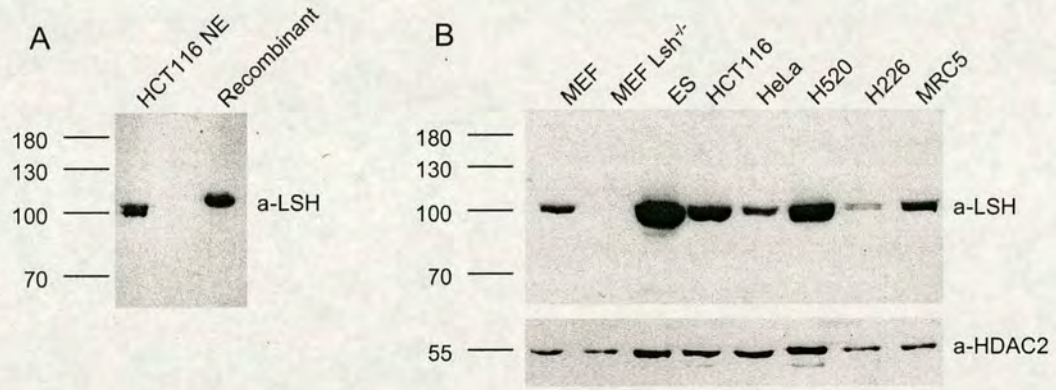


Figure 4.1. Expression of LSH in various cell lines .
A. Nuclear extract from HCT116 cells and purified recombinant protein (indicated above the gel) separated by SDS-PAGE. LSH detected by Western blot using commercial α-LSH antibody (sc-46665). Size in kDa indicated to the left of the gel.
B. Nuclear extract from the cell lines indicated above the gel separated by SDS-PAGE. LSH expression (top panel) detected by Western blot and compared to that of the loading control HDAC2 (bottom panel). Size in kDa indicated to the left.

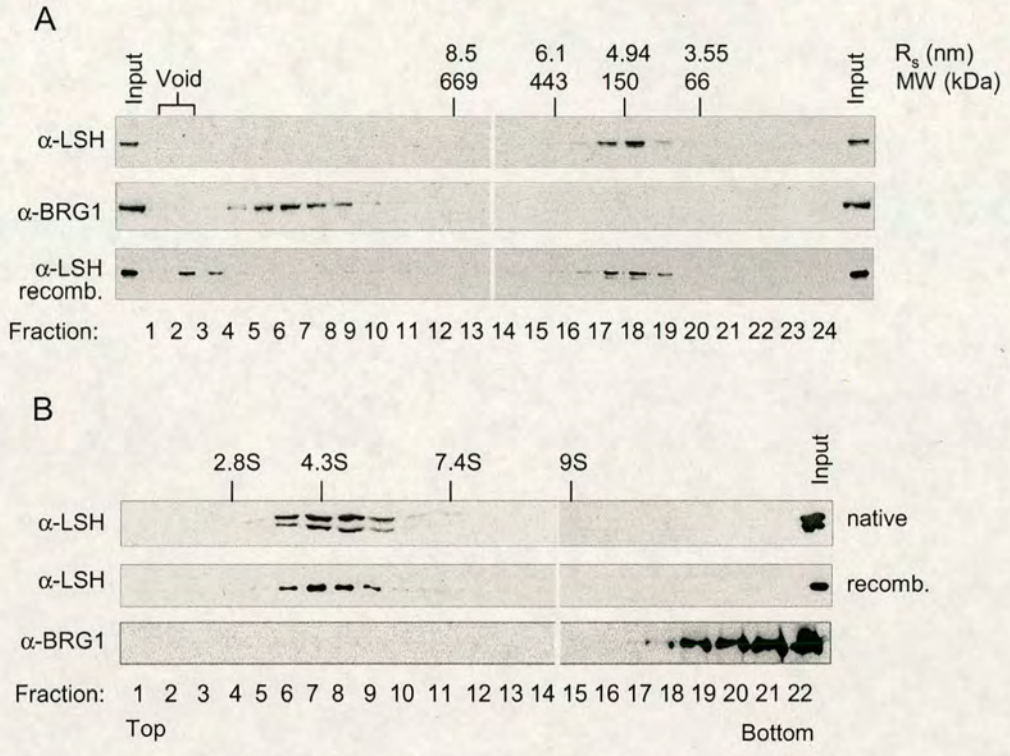


Figure 4.2. Native LSH is a monomer in nuclear extracts of human cells.
A. HCT116 nuclear extracts and recombinant LSH purified from insect cells were fractionated on a Superose 6 size exclusion column and run on sodium dodecyl sulfate-polyacrylamide gels. LSH and BRG1 proteins were detected in the collected fractions by Western blotting with appropriate antibodies. The molecular masses and Stokes radii (*Rs*) of marker proteins used to calibrate the column are indicated at the top.
B. HCT116 nuclear extracts and recombinant LSH were fractionated in 5 to 20% sucrose gradients. LSH and BRG1 proteins were detected in gradient fractions on Western blots with appropriate antibodies. Sedimentation coefficients of marker proteins are indicated above the blot.

The lack of an obvious large stable LSH complex in HCT116 NE does not completely exclude the possibility that LSH associates with other proteins. It is possible this complex exists in other cell lines, at specific points during the cell cycle or that the nuclear extraction methods used had disrupted it. To test these possibilities I repeated the gel filtration experiment using NE derived from a number of different cell lines including HeLa, a VA13 cell line stably expressing Lsh-GFP and mouse embryonic stem (ES) cells (Figure 4.3A). Of particular interest were the ES cells as Lsh is highly expressed and has been shown to co-immunoprecipitate with Dnmt3a and Dnmt3b from nuclear extract made from these cells (Zhu et al.,

2006). As can be seen from this experiment, similar to HCT116 nuclear extract the majority of LSH/Lsh or Lsh-GFP does not appear to form a large stable complex in HeLa, VA13 or ES cells (Figure 4.3A). The majority of native LSH is not part of a stable complex in asynchronously growing cells. It is possible however that LSH may form a low abundance cell cycle stage specific complex that is below the detection limit of the experiment. To test this hypothesis I synchronised HeLa cells in S phase and G2/M and assayed nuclear extract derived from these cells for an LSH complex. Using a double S-phase block and release synchronisation protocol, I was able to highly enrich my HeLa sample for cells in S-phase (~90%) and G2/M (~80%) compared to asynchronously growing cells (S-phase - ~15%; G2/M - ~10%) (Figure 4.3B; compare top left with bottom left and bottom right panels). I was unable to enrich for cells in G1 using this protocol but, as ~75% of asynchronously growing cells are in G1, I assumed I would be able to see a G1 specific complex in these cells. Using nuclear extract from the synchronised cells I did not detect a cell cycle stage specific LSH complex (Figure 4.3C). The preparation of nuclear extract for these experiments followed the Dignam protocol that uses a relatively high salt concentration (420mM NaCl) to elute nuclear proteins from prepared nuclei (Dignam et al., 1983). To ensure that the salt extraction method had not disrupted an unstable LSH complex I performed the same experiment using a lower salt concentration (210mM). The low salt extracted LSH was also not present in a large complex (Figure 4.3D; top panel). Alternatively a stable LSH complex may be tightly bound to chromatin and 420mM NaCl may not be high enough to elute it. To determine if a fraction of LSH associates with the insoluble nuclear pellet fraction that remains following nuclear extraction I attempted to solubilise it. I found that high salt extraction (1M NaCl), in conjunction with sonication, DNase and MNase treatment successfully solubilised the nuclear pellet. LSH was present in the HCT116 nuclear pellet as can be seen from the Western blot of the gel filtration (Figure 4.3D; bottom panel). When run through the gel filtration column however, the nuclear pellet fraction of LSH eluted as before, at ~150kDa (Figure 4.3D; bottom panel).

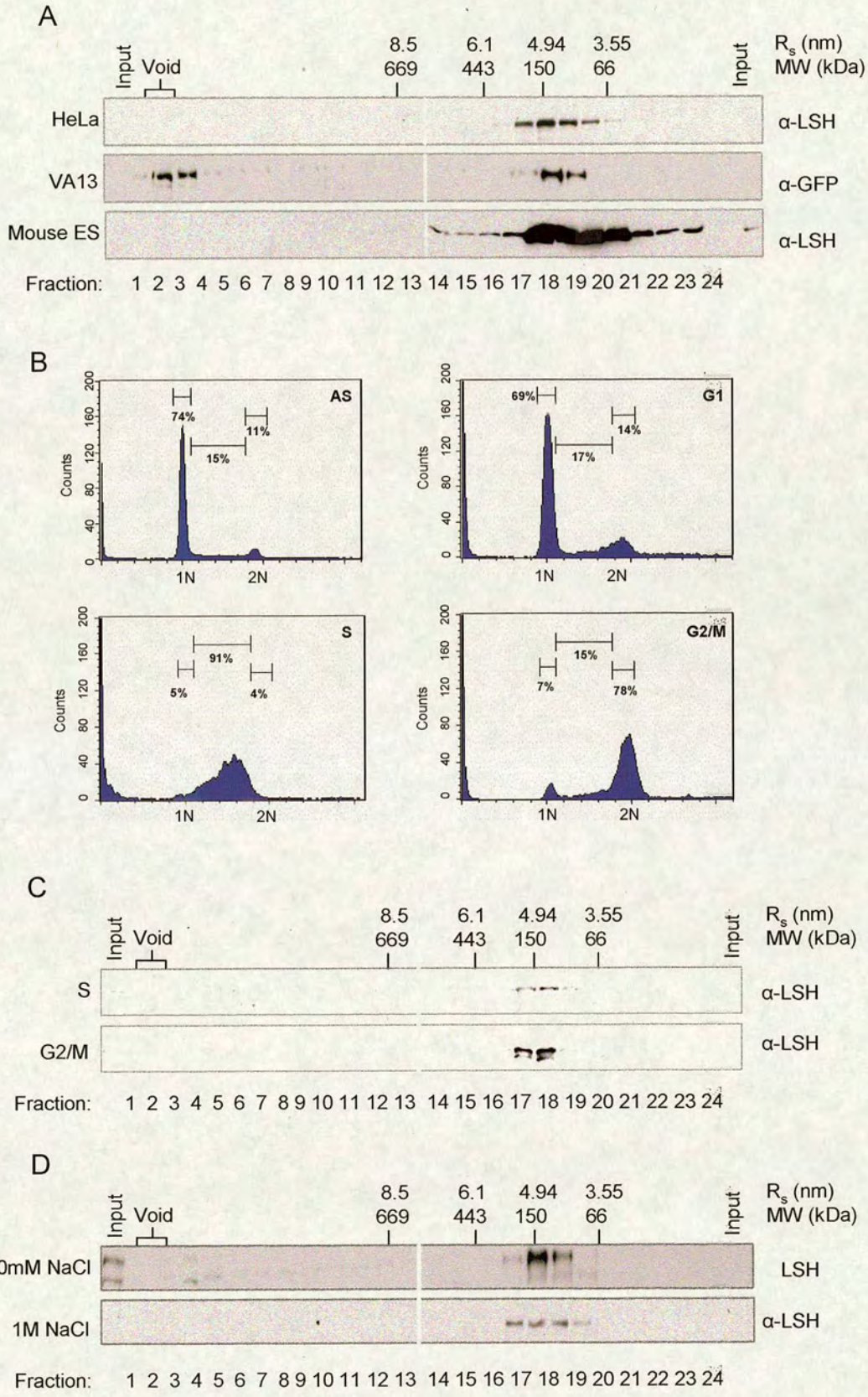

A
Input
Void
8.5
669
6.1
443
4.94
150
3.55
66
Input
Rs (nm)
MW (kDa)
HeLa
α-LSH
VA13
α-GFP
Mouse ES
α-LSH
Fraction: 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24
B
AS
74%
11%
15%
G1
69%
14%
17%
S
91%
5%
4%
G2/M
15%
7%
78%
Counts
1N
2N
C
Input
Void
8.5
669
6.1
443
4.94
150
3.55
66
Input
Rs (nm)
MW (kDa)
S
α-LSH
G2/M
α-LSH
Fraction: 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24
D
Input
Void
8.5
669
6.1
443
4.94
150
3.55
66
Input
Rs (nm)
MW (kDa)
220mM NaCl
LSH
1M NaCl
α-LSH
Fraction: 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24

Figure 4.3. Native LSH is not a member of a large stable complex.
A. Gel filtration with mouse ES cell extract was performed by Melanie Lawrence. Nuclear extract from HeLa, VA13 and mouse ES cells was fractionated on a Superose 6 size exclusion column and run on sodium dodecyl sulfate-polyacrylamide gels. LSH/Lsh or Lsh-GFP protein was detected in the collected fractions by Western blotting with appropriate antibodies. The molecular masses and Stokes radii (*Rs*) of marker proteins used to calibrate the column are indicated at the top.
B. FACS analyses are shown of asynchronously growing HeLa cells (AS) and cells in G1, S and G2/M phases of the cell cycle after synchronization by thymidine block followed by mimosine arrest and release into S-phase. The numbers below the vertical bars represent % of cell in G1, S and G2/M in each of the four samples. 1N and 2N indicate DNA content.
C. S and G2/M extracts were fractionated on Superose 6 size exclusion column and individual fractions were analyzed on Western blots with anti-LSH antibodies. In all cases, LSH eluted at about 150 kDa suggesting that the majority of endogenous LSH is not involved in a protein complex that assembles at some specific time of the cell cycle
D. Low (220mM) and high (1M) NaCl extracts were fractionated on Superose 6 size exclusion column and individual fractions were analyzed on Western blots with anti-LSH antibodies. In both cases, LSH eluted at about 150 kDa suggesting that the majority of endogenous LSH is not involved in a protein complex that is unstable in high salt or tightly associated with chromatin.

4.2.2 Native LSH is present as a monomer

The molecular mass of a native protein or a protein complex can be determined accurately by an equation derived by Siegel and Monty which combines the Stokes radius (a hydrodynamic radius of a molecule freely tumbling in solution) calculated from size exclusion chromatography with the sedimentation coefficient determined by separation in sucrose gradient (Siegel and Monty, 1965). As the size exclusion chromatography does not rule out the possibility that LSH associates with itself, I sought to determine its native molecular mass using this method. Based on the elution profile of protein standards from size exclusion chromatography I calculated the Stokes radius of LSH to be ~4.94 nm. To establish the sedimentation coefficient of LSH, I fractionated HCT116 and mouse ES nuclear extracts and recombinant purified LSH on 5 to 20% sucrose gradients and detected LSH in the gradient fractions by Western blots. In these experiments the sedimentation coefficient of native and recombinant human LSH was calculated to be 4.5S, relative to protein standards (Figure 4.2B; top and middle panels). As in the gel filtration experiments Brg1 was used to control for NE integrity (Figure 4.2B, bottom panel). To determine the molecular weight of LSH the stokes radius and sedimentation co-efficient were applied to the Siegel and Monty formula as follows:

$M_r = 6\pi\eta_{20,\omega} \cdot s_{20,\omega} \cdot R_s \cdot N/(1 - \rho_{20,\varpi}\upsilon)$

M_r=91,500 kDa

Where:

R_s is the Stoke's radius (nm) determined by Superose 6 gel filtration chromatography,

$R_{s(LSH)} = 4.94$nm

$s_{20,\omega}$ is the sedimentation velocity (S x 10^{-13}) determined from the sucrose gradient,

$s_{20,\omega\ (LSH)} = 4.5$ S

$\eta_{20,\varpi}$ is the viscosity of water at 20°C (0.01002 $g \cdot s^{-1}\ cm^{-1}$),

N is Avogadro's number (6.022 x 10^{23}·$molecules^{-1}$),

$_{20,\varpi}$ is the density of water at 20°C (0.9981 $g \cdot cm^3$),

υ is the partial specific volume of protein (used 0.725 cm^3/g).

Using this equation, the native molecular mass of LSH was determined to be 91,500 kDa, which is very close to the predicted mass of monomeric LSH (97,000 kDa). Thus, the majority of LSH protein in nuclear extracts of mouse and human cells is present as a free monomeric peptide.

Theoretical mass (Da)	**Sedimentation coefficient (S)**	**Stokes radius [R_s] (nm)**	**Derived molecular mass (Da)**
97000	4.5	4.94	91500

Table 4.1 Hydrodynamic properties of native LSH.
Hydrodynamic analysis of LSH was carried out using the Siegel and Monty equation (Siegel and Monty, 1965). The sedimentation coefficient (determined by sucrose gradient) and Stokes radius (determined by gel filtration) were used to derive the molecular mass of native LSH.

4.3 Yeast 2-hybrid screen for LSH interacting proteins

4.3.1 Yeast 2-hybrid screen with GAL4BD-LSH did not uncover any positive clones

In addition to the stably bound proteins in SWI2/SNF2 complexes a wide range of less stable interactions have been reported. Of particular interest in the case of the DNA methylation phenotype reported for LSH is the reported interaction between SNF2H and DNMT3B (Geiman et al., 2004b). This interaction can be detected in coimmunoprecipitation experiments but not via gel filtration chromatography (Strohner et al., 2001). In order to identify proteins that interact with LSH less stably I employed a yeast 2-hybrid (Y2H) screen. The Y2H system I used was the Clonetech Matchmaker GAL4 Two-Hybrid System. In this system expression of four selectable markers is dependent on an interaction between a bait protein fused to the GAL4 binding domain (BD) and a prey protein fused to the GAL4 activation domain (AD). The bait protein is expressed in the MATa yeast strain AH109 and is present on a plasmid containing a tryptophan (TRP) marker. The library of prey proteins are expressed in the MATα strain Y187 from a plasmid containing a leucine (LEU) marker. An interaction between the bait protein and any of the screened prey proteins leads to expression of adenine (ADE) and histidine (HIS) selectable markers and growth on media lacking TRP, LEU, ADE and HIS (QDO media). Another two markers MEL1 and LacZ are also activated and can be used for blue-white screening and semi-quantitative β-galactosidase assays respectively.

I successfully cloned full length LSH into the GAL4BD vector and confirmed its expression in the AH109 yeast strain by Western blotting with α-LSH and α-GAL4BD antibodies (Figure 4.4). I used this fusion as a bait to screen a HeLa cDNA library of ~$1x10^6$ unique GAL4AD clones. Surprisingly the Y2H screen did not return any colonies that I was able to verify by DNA sequencing.

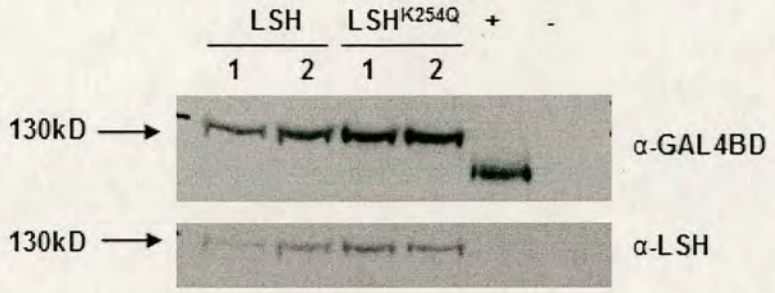


Figure 4.4. GAL4BD-LSH overexpression in yeast.
Yeast extract from strains expressing LSH or LSH^{K254Q} were separated by SDS-PAGE. GAL4BD-LSH was detected by Western blot with α-GAL4BD (top panel) or α-LSH (bottom panel) antibodies. GAL4BD-MBD1 and non transfected yeast were included as positive control and negative controls respectively.

4.3.2 GAL4BD-LSH represses the yeast 2-hybrid reporters

The lack of positive colonies obtained by the Y2H screen was surprising as most Y2H screens yield hundreds of colonies, including many false positives. This indicated that either the Y2H screen had failed, that LSH does not interact with many proteins, is inhibitory to yeast cell survival or represses activation of the Y2H reporter genes.

4.3.2.1 The Y2H screen successfully screened over one million clones

I first assessed the efficiency of yeast mating and how many individual clones had been screened. These experiments aimed to determine whether the screen was successful or not. The yeast strains used for the screen are genetically deficient for genes required to make tryptophan, leucine, adenine and histidine. The bait and activator plasmids used for the screen provide the missing genes required for tryptophan and leucine synthesis respectively. It is therefore possible to calculate mating efficiency by comparing the respective viability of the mated yeast on plates lacking tryptophan or leucine or both amino acids. The viability of the mated yeast on the single dropout plates is calculated and the strain with the lower viability is the 'limiting partner'. The Y187 library strain was limiting in this screen and had a viability of ~4 x 10^6 colony forming units / ml (cfu/ml). The mated strain had a

viability of ~1 x 10^5 cfu/ml, thus the efficiency of mating for this screen was ~2.5%. The number of individual clones screened can be calculated from the mated strain viability multiplied by the volume of yeast plated. I plated 11.5ml in this screen so managed to screen ~1.15 x 10^6 GAL4AD clones for interaction with GAL4BD-LSH. The relatively good mating efficiency combined with the high number of clones screened indicated that the Y2H screen had been successful.

I next tested whether GAL4BD-LSH was toxic to the yeast strains used in the Y2H screen. GAL4BD-LSH is present on the pGBKT7 plasmid that contains a tryptophan selectable marker. This plasmid is transformed into the yeast strain AH109 for the screen. Strains transformed with the empty vector or GAL4BD-LSH containing vector grew at the same rate under selection indicating that LSH is not toxic to this yeast strain. Also, these strains mated to the Y187 strain containing the activator plasmid pGADT7 grew at the same rate. Together these data suggest that GAL4BD-LSH is not toxic to either the AH109 or AH109/Y187 mated strains.

4.3.2.2 GAL4BD-LSH represses the Y2H reporter genes

I finally tested the possibility that GAL4BD-LSH repressed the reporter genes that assay for protein-protein interactions. I felt this was particularly likely as false positives are common in Y2H screens but none occurred during this one. To test this, I used the KIAA0737 protein that self activates the selectable Y2H reporter genes and causes yeast growth on QDO dropout media in the absence of an interacting protein (Table 4.4). Interestingly, like the repressive control protein GAL4BD-MeCP2, GAL4BD-LSH prevented yeast growth when mated to the KIAA0737 strain indicating it is able to repress transcription of the reporter genes (Table 4.2). I was able to map the repressive domain of LSH in this system using deletion constructs to the coiled-coils at the N-terminus of the protein. This indicates that a yeast protein interacting with this region of GAL4BD-LSH was responsible for the low number of positive clones obtained by the screen (Table 4.2).

	Growth on QDO media and β-galactosidase assay	
Bait Plasmid	**pGADT7**	**pACT2-KIAA0737**
pGBKT7	-	++
pGBKT7-LSH	-	-
pGBKT7-LSH(1-568)	-	-
pGBKT7-LSH(227-838)	-	++
pGBKT7-LSH(569-838)	-	++
pGBKT7-MeCP2(TRD)	-	-

Table 4.2. GAL4BD-LSH prevents activation of Y2H reporter genes
The AH109 yeast strain carrying the bait plasmid indicated was mated with the Y187 strain carrying either empty pGADT7 or pACT2-KIAA0737 plasmids. Yeast survival on QDO media was scored by growth (-/+) and β-galactosidase assays (++).

4.3.2.3 Y2H screen with GAL4BD-LSH(227-838) did not uncover any other interacting proteins

As the domain of LSH that repressed transcription mapped to the coiled-coil domains I repeated the Y2H screen with a deletion construct not containing this part of the protein, GAL4BD-LSH(227-838). This screen yielded only five clones that were sequenced to determine their identity. All of these were found to be false positives that did not encode recognised proteins. As previously, I confirmed the success of the screen by calculating the mating efficiency and the number of clones screened. I calculated the mating efficiency at ~2.3% and that ~$1.7x10^6$ clones had been screened. As the GAL4BD-LSH(227-838) fusion is not inhibitory to yeast survival and does not inhibit reporter expression I would conclude that this region of LSH does not have many interacting proteins (Table 4.2). The lack of interactions between the SNF2_N and Helicase_C domains of LSH and other proteins is not entirely surprising as these domains are believed to be primarily involved in interactions with DNA and not with other proteins (Durr et al., 2005).

Y2H screens are a powerful method of discovering protein-protein interactions that are not stable enough to be detected by conventional chromatography. In an attempt to determine the role of LSH in DNA methylation I have attempted to determine the

proteins it interacts with using this method. Using full length LSH to screen a HeLa cDNA library I did not uncover any proteins that interact with LSH. Interestingly, I have found that a possible reason for the low number of positive clones is the potent repression by GAL4BD-LSH of the Y2H reporter genes. I have mapped this repressive ability to the N-terminal coiled-coil domain of LSH, which would be expected to mediate protein-protein interactions (Table 4.2). Thus, this type of Y2H screen is not appropriate for detecting proteins that potentially interact with LSH via this domain. In the future it would be possible to screen for potential LSH interaction partners by other means such as fluorescence resonance energy transfer (FRET) or bimolecular fluorescence complementation (BiFC) (Ding et al., 2006; You et al., 2006). However, this result did open up other interesting avenues for research, which were subsequently undertaken and will now be discussed.

4.4 GAL4BD-LSH cooperates with DNMTs to repress transcription

4.4.1 GAL4BD-LSH is a HDAC dependant transcriptional repressor

A key role of SWI2/SNF2 enzymes is in regulating transcription (Abrams et al., 1986). Remodelling enzymes can have either activatory of repressive effects on transcription, with the specific effects in part depending on interacting proteins (Cheng et al., 1999; Tong et al., 1998; Zhang et al., 1998). LSH has crucial a role in determining proper DNA methylation in mammals. As DNA methylation is a transcriptionally repressive mark and as GAL4BD-LSH repressed transcription in yeast I investigated whether LSH acts as a transcriptional repressor in mammalian cells.

To achieve this I cotransfected HCT116 cells with a plasmid expressing full-length LSH fused to GAL4BD and two reporter plasmids. The first reporter plasmid carried five GAL4 binding sites upstream of a thymidine kinase (TK) promoter driving the expression of the firefly luciferase gene. The other, control plasmid lacked GAL4 binding sites and expressed β-galactosidase from an actin promoter (Figure 4.5A). The effect of either full length LSH or the coiled coil domain of LSH on transcription from the targeted and non targeted reporter was measured as a ratio of luciferase to β-galactosidase expression (Figure 4.5B). A GAL4BD-tagged transcriptional repression domain (TRD) of methyl-CpG binding protein MeCP2, which is known to strongly repress transcription in such assays, served as a control (Nan et al., 1997). Western blot experiments confirmed expression of all the GAL4BD-LSH fusions in HCT116 cells (Figure 3.5C) In these experiments the full-length GAL4BD-LSH as well as GAL4BD-MeCP2(TRD) consistently reduced the expression of the luciferase reporter to about 20% and 10%, respectively, of the levels observed in cells transfected with an empty vector (Figure 4.5D). This suggests that like MeCP2, LSH can function as an efficient transcriptional repressor when targeted to a promoter of a reporter gene in human cells.

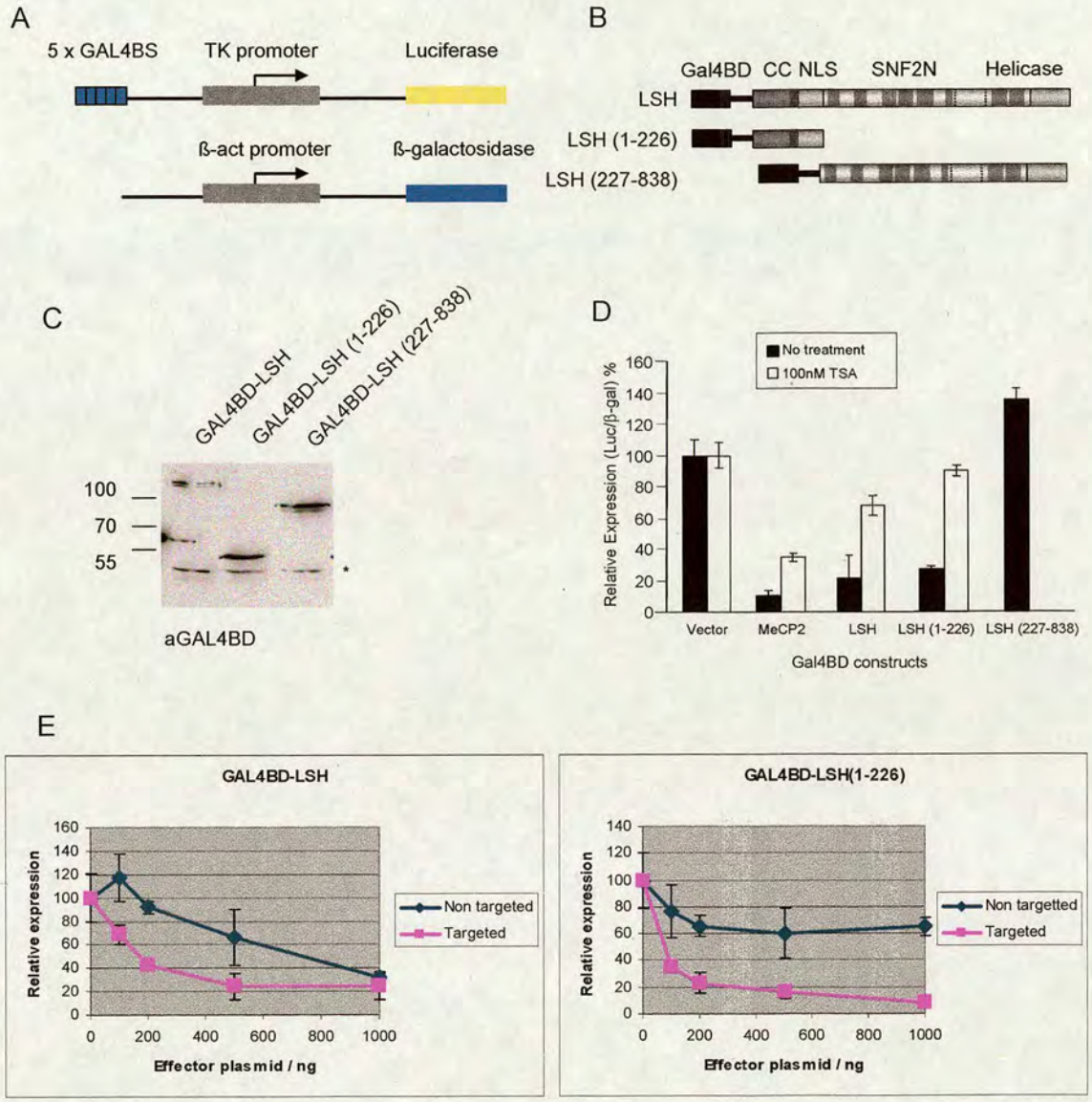


Figure 4.5. LSH efficiently represses transcription when targeted to the promoter of a reporter gene.

A. Schematic drawing of the reporter constructs used in this study. Fusion proteins are targeted to the 5xGAL4 binding sites upstream of the TK promoter. Expression of luciferase determines transcriptional effect. A control plasmid not containing GAL4 binding sites acts as a control for transfection and GAL4BD independent effects.

B. Schematic drawing of the full-length LSH and truncated LSH proteins fused to a GAL4BD. The functional domains of LSH, such as the coiled-coil domain (CC), the nuclear localization signal (NLS), and the eight conserved SNF2 motifs in the SNF2N and helicase domains, are indicated.

C. GAL4BD fusions of LSH were transfected into HCT116 cells, the cells lysed with 2 x SDS-PAGE buffer and proteins separated by SDS-PAGE. Fusion proteins were detected by Western blot with α-GAL4BD antibodies. Size markers are indicated to the left. Non-specific band that controls for equal loading indicated (*)

D. GAL4BD fusions of LSH were cotransfected into HCT116 cells with the two reporter constructs. The relative expression of the reporters represents the ratio of luciferase to β-

galactosidase products. MeCP2 was used as a control. The white bars represent experiments carried out in the presence of 100 nM TSA, which partially alleviates LSH and MeCP2-mediated repression of the luciferase reporter. The error bars represent standard deviations of the means.
E. Increasing amounts of GAL4BD-LSH and GAL4BD-LSH(1-226) were cotransfected into HCT116 cells with the two reporter constructs. Expression of the individual reporters was measured to determine GAL4BD independent transcriptional effects on β- galactosidase expression.

To investigate whether a specific domain of LSH was responsible for transcriptional repression, I tested a number of LSH deletion constructs in the luciferase reporter assay (Figure 4.5B). The expression of full length GAL4B-LSH and two deletion constructs was monitored by Western blot and found to be similar (Figure 4.5C). Interestingly, the N-terminal portion of LSH (amino acids 1 to 226), containing the predicted coiled-coil motif, was sufficient to repress the reporter gene to levels comparable to those observed with full-length LSH. A polypeptide corresponding to the SNF2 and helicase domains of LSH (amino acids 227 to 838), did not significantly affect the expression of the luciferase reporter (Figure 4.5C and D). These experiments indicate that the 226-amino-acid coiled-coil region of LSH functions as a TRD and is sufficient for silencing of the luciferase reporter. Importantly, this result uncouples the supposed catalytic role of LSH from its role in transcriptional repression. To control for GAL4BD independent repression of the non targeted β-gal reporter and to confirm that repression was dose dependent I repeated the reporter assays using a titration of effector plasmid (Figure 4.5E). As can be seen from the titration experiment both LSH and LSH(1-226) repress transcription in a dose dependent manner with LSH(1-226) being a slightly more efficient transcriptional repressor. Both proteins also showed a significant level of non specific repression of the β-gal control plasmid in particular LSH at 1000ng (Figure 4.5E). This non specific repression may be due to GAL4BD independent binding of the over-expressed protein to the control plasmid. If this is the case it would explain why high levels of full length LSH, containing the SNF2 DNA binding domain showed greater repression of the control plasmid than LSH(1-226). Due to this GAL4BD independent repression I determined the optimal amounts of each plasmid

to use in further experiments to maximise targeted, whilst minimising non-targeted repression to be 500ng LSH and 200ng LSH(1-226).

Coiled-coil regions of proteins often mediate protein-protein interactions. In an attempt to dissect the mechanism of LSH mediated transcriptional repression I next asked whether LSH functionally cooperates with co-repressor proteins such as HDACs as is the case for MeCP2 (Nan et al., 1997). Consistent with this hypothesis, transcriptional repression by GAL4BD-MeCP2(TRD), GAL4BD-LSH and GAL4BD-LSH(1-226) was partially alleviated when I performed the reporter assays in the presence of 100nM of the HDAC inhibitor trichostatin A (TSA) (Figure 4.5D). This indicates that transcriptional repression by LSH may be mediated via an interaction between the coiled-coil domain of LSH and HDACs. To provide further evidence for this hypothesis I asked whether endogenous HDACs and LSH can co-immunoprecipitate from HCT116 NE. Antibodies against LSH, but not control α-HA antibodies (mock) immunoprecipitated LSH and co-immunoprecipitated HDAC1 and HDAC2 (Figure 4.6A). I also detected LSH in reciprocal immunoprecipitates with α-HDAC1 and α-HDAC2 antibodies (Figure 4.6A). In addition to this I performed chromatin immunoprecipitation experiments in cells carrying stably integrated copies of the reporter construct transfected with GAL4BD-LSH (Ishizuka and Lazar, 2003). These experiments showed a three-fold decrease of acetylated lysine 9 of histone H3 and a two-fold decrease of acetylated lysine 12 of H4 at the targeted TK promoter compared to a non-targeted control (Figure 4.6B). I further characterised GAL4-LSH mediated repression of the reporter constructs by performing RNA interference knockdown experiments on HDAC1, HDAC2 and HDAC3 in HCT116 cells. The RNAi experiments were successful as each RNAi specifically depleted expression of its target (Figure 4.6C). These cells were subsequently used for reporter assays and revealed that successful, simultaneous depletion of HDAC1 and HDAC2, but not HDAC3, could similar to TSA treatment, alleviate the repression of luciferase reporter (Figure 4.6D). Together, these experiments provide strong evidence that transcriptional repression by GAL4-LSH in the reporter assay system requires HDAC1 and HDAC2 (Figure 4.6).

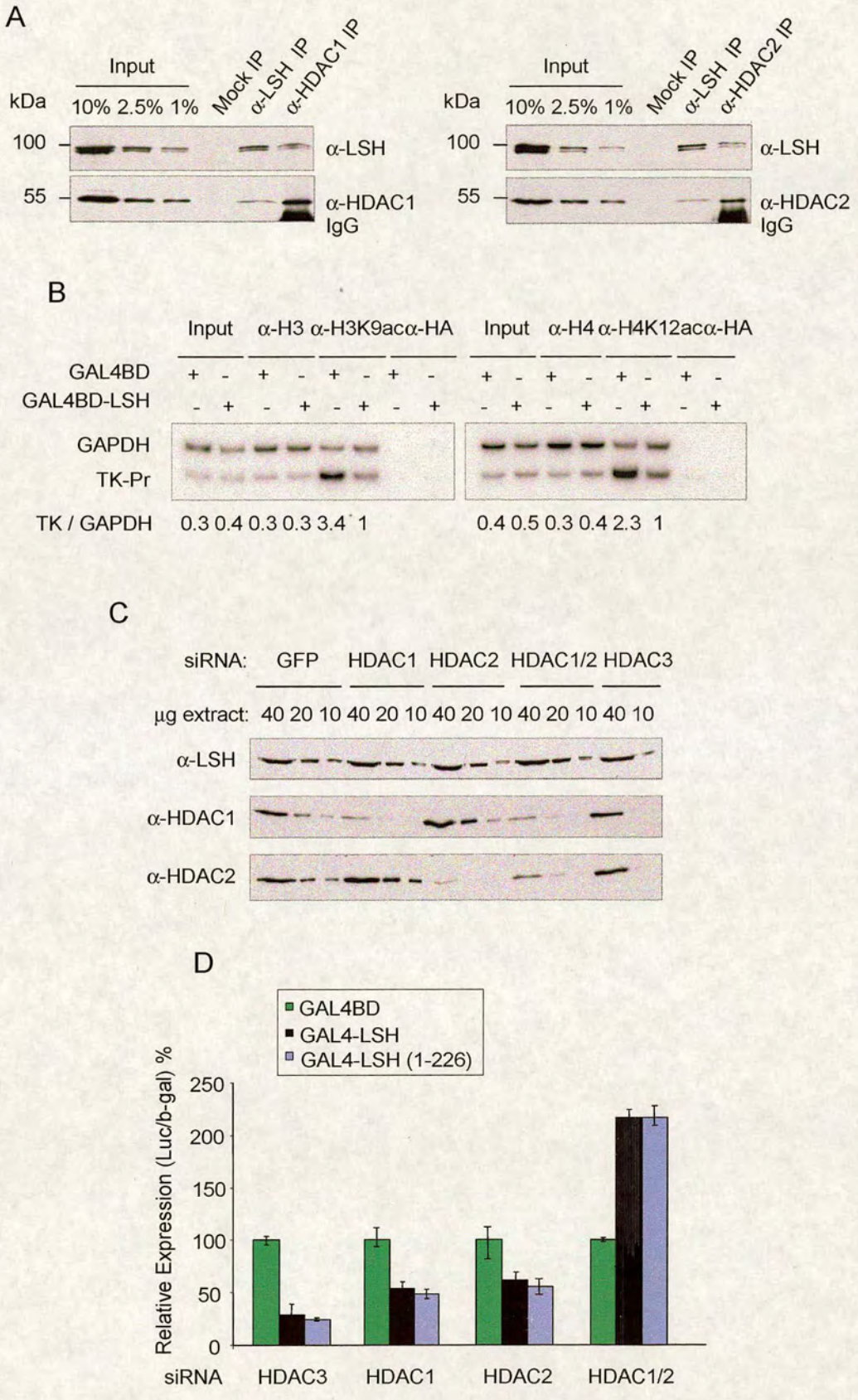

A
Input
Mock IP
α-LSH IP
α-HDAC1 IP
kDa
10% 2.5% 1%
100
55
α-LSH
α-HDAC1
IgG
Input
Mock IP
α-LSH IP
α-HDAC2 IP
kDa
10% 2.5% 1%
100
55
α-LSH
α-HDAC2
IgG
B
Input
α-H3
α-H3K9ac
α-HA
Input
α-H4
α-H4K12ac
α-HA
GAL4BD
GAL4BD-LSH
GAPDH
TK-Pr
TK / GAPDH
0.3 0.4 0.3 0.3 3.4 1
0.4 0.5 0.3 0.4 2.3 1
C
siRNA:
GFP
HDAC1
HDAC2
HDAC1/2
HDAC3
μg extract:
40 20 10 40 20 10 40 20 10 40 20 10 40 10
α-LSH
α-HDAC1
α-HDAC2
D
GAL4BD
GAL4-LSH
GAL4-LSH (1-226)
Relative Expression (Luc/b-gal) %
250
200
150
100
50
0
siRNA
HDAC3
HDAC1
HDAC2
HDAC1/2

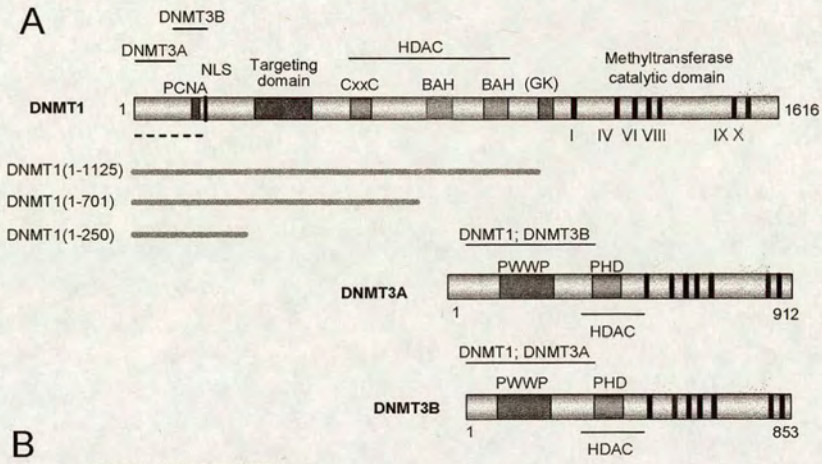
Figure 4.6. LSH functions as an HDAC-dependent transcriptional repressor in vivo.
A. Coimmunoprecipitation experiments with anti-LSH, anti-HDAC1, and anti-HDAC2 antibodies from HCT116 nuclear extracts. Anti-HA antibodies were used in mock immunoprecipitations (IP) to control for non-specific protein binding.
B. Chromatin immunoprecipitation with antibodies against acetylated H3-K9 and H4-K12. The TK promoter in cells transfected with GAL4BD-LSH is hypoacetylated compared to the GAPDH promoter used as an internal control. Anti-HA is a non-specific control antibody
C. SMART pool siRNAs efficiently reduce the levels of HDAC1 and HDAC2 in HCT116 cells compared with control GFP and HDAC3 siRNAs.
D. Reporter expression assay indicate that neither full length LSH (GAL4-LSH) nor the coiled coil domain of LSH repress the luciferase reporter in HCT116 cells treated simultaneously with HDAC1 and HDAC2 siRNA compared to controls.

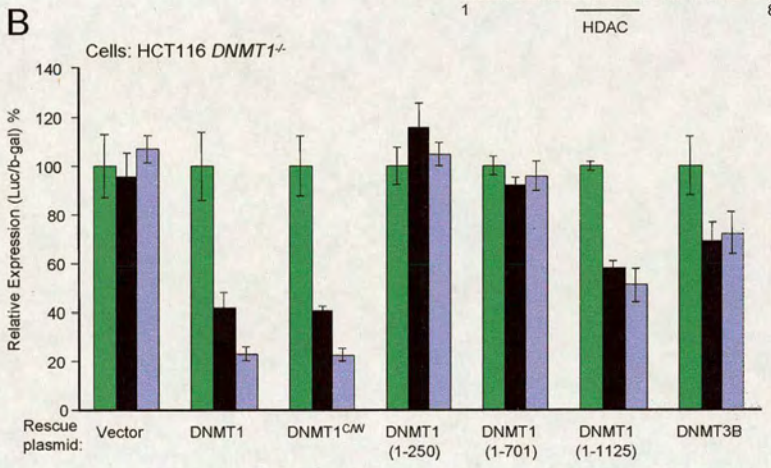
4.4.2 Transcriptional repression by GAL4BD-LSH requires DNMT1 and DNMT3B

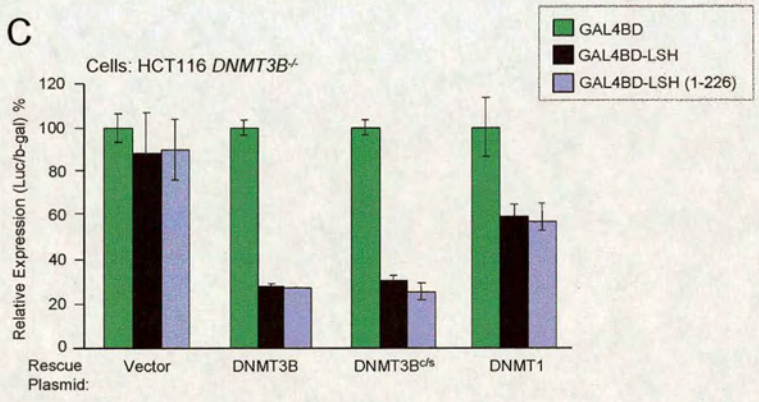
4.4.2.1 GALBD-LSH cannot repress transcription in cells deficient for DNMT1 or DNMT3B

The major phenotype of the $LSH^{-/-}$ mouse is a genome wide loss of DNA methylation (Dennis et al., 2001). A recent study has shown that Lsh interacts with Dnmt3a and Dnmt3b but not Dnmt1 in MEFs (Zhu et al., 2006). Dnmt1, Dnmt3a and Dnmt3b have also been shown to interact with each other and with histone deacetylases (Fuks et al., 2000; Fuks et al., 2001; Kim et al., 2002; Rountree et al., 2000) (Figure 4.7A). Given these complex interactions and the role of LSH in DNA methylation, we reasoned that DNMTs may also play a role in GAL4BD-LSH mediated transcriptional repression.

HCT116 cell lines genetically deficient for DNMT1 (*DNMT1*$^{-/-}$), DNMT3B (*DNMT3B*$^{-/-}$), DNMT1 and DNMT3B (*DKO*) and $DNMT1^{+/-}$, DNMT3A and DNMT3B (*TKO*) have been generated by homologous recombination (Jair et al., 2006; Rhee et al., 2002; Rhee et al., 2000). It has recently been shown, however, that the DNMT1 KO cell lines, including the DKO, express low levels of a truncated DNMT1 protein missing 150 amino acids of the N terminus. This region includes the sites of interaction with DNMT3A, DNMT3B, and PCNA (Egger et al., 2006; Spada et al., 2007) (Figure 4.7A). To assay for the involvement of DNMTs in LSH mediated repression I repeated the reporter assays in the *DNMT1*$^{-/-}$, *DNMT3B*$^{-/-}$ and *TKO* cell lines. Surprisingly, I found that GAL4BD-LSH and GAL4BD-LSH(1-226) mediated repression was completely absent in the KO cell lines tested (Figure 4.7)


A
DNMT3B
DNMT3A
NLS
PCNA
Targeting domain
HDAC
CxxC
BAH
BAH
(GK)
Methyltransferase catalytic domain
DNMT1
1
1616
I
IV
VI
VIII
IX
X
DNMT1(1-1125)
DNMT1(1-701)
DNMT1(1-250)
DNMT1; DNMT3B
PWWP
PHD
DNMT3A
1
912
HDAC
DNMT1; DNMT3A
PWWP
PHD
DNMT3B
1
853
HDAC


B
Cells: HCT116 DNMT1-/-
Relative Expression (Luc/b-gal) %
140
120
100
80
60
40
20
0
Rescue plasmid:
Vector
DNMT1
DNMT1C/W
DNMT1 (1-250)
DNMT1 (1-701)
DNMT1 (1-1125)
DNMT3B


C
GAL4BD
GAL4BD-LSH
GAL4BD-LSH (1-226)
Cells: HCT116 DNMT3B-/-
Relative Expression (Luc/b-gal) %
120
100
80
60
40
20
0
Rescue Plasmid:
Vector
DNMT3B
DNMT3Bc/s
DNMT1

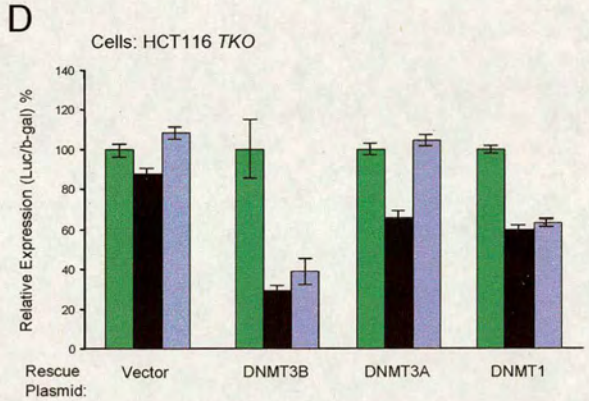

D
Cells: HCT116 TKO
Relative Expression (Luc/b-gal) %
140
120
100
80
60
40
20
0
Rescue Plasmid:
Vector
DNMT3B
DNMT3A
DNMT1

Figure 4.7. Transcriptional repression by LSH requires DNMT3B and the N-terminal portion of DNMT1.
A. Schematic representation of DNMT1 and DNMT3B proteins with their functional domains. The cysteine-rich putative DNA binding CxxC domain, bromo-adjacent homeobox motifs (BAH), GK-rich repeats, the domain involved in targeting to replication foci, and the catalytic part of DNMT1 are indicated. Mapped interactions with DNMT3A, DNMT3B, PCNA, and HDAC1 and -2 are shown above the diagram. The dashed line indicates the portion of DNMT1 that has been spliced out in DNMT1 KO HCT116 cells with targeted disruption of the DNMT1 gene (Egger et al., 2006; Rhee et al., 2002). The grey lines below the schematic indicate the fragments of the protein used for rescue experiment. The DNMT3A and DNMT3B proteins contain a DNA binding PWWP motif, a PHD, domain and a conserved catalytic DNMT domain. The portions of DNMT3A interacting with DNMT1 and DNMT3B and the portion of DNMT3B interacting with DNMT1 and DNMT3A are indicated above the schematic. The HDAC interacting region is indicated for both proteins below the drawing.
B. Neither full-length GAL4BD-LSH nor the TRD of LSH, GAL4-LSH(1–226), could silence the luciferase reporter in *DNMT1*$^{-/-}$ cells. Cotransfection of GAL4BD-LSH proteins together with wild-type GFP-tagged DNMT1, catalytically inactive DNMT1$^{C/W}$, and the N-terminal portion of DNMT1(1–1125) can rescue the repression of luciferase reporter in *DNMT1*$^{-/-}$ cells. Shorter N-terminal DNMT1 proteins [DNMT1(1–250) and DNMT1(1–701)] did not rescue the repression of the luciferase reporter gene. DNMT3B was used as an additional control. Error bars indicate standard deviations.
C. GAL4BD-LSH and GAL4BD-LSH(1–226) did not repress the luciferase reporter in *DNMT3B*$^{-/-}$ cells. Cotransfection of LSH proteins with either GFP-DNMT3B or a catalytically inactive GFP-DNMT3B$^{C/S}$ restored the repression of the reporter to levels observed in wild-type HCT116 cells. Cotransfection of GAL4BD-LSH with GFP-DNMT1 also reduced the repression of luciferase in DNMT3B KO cells, although not as efficiently as DNMT3B
D. GAL4BD-LSH and GAL4BD-LSH(1–226) did not repress the luciferase reporter in *TKO* cells. Cotransfection of LSH proteins with GFP-DNMT3B restored the repression of the reporter to levels observed in wild-type HCT116 cells. Cotransfection of GAL4BD-LSH with GFP-DNMT1 also reduced the repression of luciferase in DNMT3B KO cells, although not as efficiently as DNMT3B. Cotransfection of GAL4BD-LSH with GFP-DNMT3A did not rescue repression as efficiently as DNMT3B either.

To test that the observed effect was due to lack of DNMT protein and not a phenotypic effect in these cells, I measured GAL4BD-LSH mediated repression following co-transfection with the corresponding DNMT plasmid. In the *DNMT1*$^{-/-}$ cells, co-transfection of Dnmt1-GFP efficiently rescued GAL4BD-LSH and GAL4BD-LSH(1-226) mediated transcriptional repression to near wild-type levels (Figure 4.7B), Interestingly, repression was also rescued in *DNMT1*$^{-/-}$ cells by a catalytically inactive mutant of Dnmt1, C1229W (Schermelleh et al., 2005). This indicates that the transcriptional repression by GAL4BD-LSH does not depend on DNA methylation by DNMT1 (Figure 4.7B). I next sought to map the region of DNMT1 required to restore GAL4-LSH mediated repression in *DNMT1*$^{-/-}$ cells using a number of deletion constructs (Figure 4.7A). DNMT1(1-1125)-GFP, which contains the known interaction sites with DNMT3A, DNMT3B, and HDACs but is

lacking the C-terminal catalytic domain, could partially restore the repression by GAL4BD-LSH and GAL4BD-LSH(1-226), while the shorter proteins DNMT1(1-701)-GFP and DNMT1(1-250)-GFP could not (Figure 4.7B). This is suggestive that a fairly large portion of the DNMT1 N-terminus and perhaps some of the catalytic C-terminus is involved in protein-protein interactions that are critical for GAL4BD-LSH mediated repression. Cotransfection of DNMT3B-GFP with either LSH construct into the *DNMT1*$^{-/-}$ cells also led to a modest (~25 - 30%) decrease in transcription from the reporter gene (Figure 4.7B). This indicates that overexpression of DNMT3B, although having a slight effect is not sufficient to restore LSH-mediated repression in this cell line. I next attempted to rescue GAL4BD-LSH mediated repression in *DNMT3B*$^{-/-}$ and *TKO* cells. As in the case of DNMT1, when I cotransfected *DNMT3B*$^{-/-}$ cells with GAL4BD-LSH or GAL4BD-LSH(1-226) and either wild-type GFP-DNMT3B or catalytically inactive GFP-DNMT3B^{C640S}, the repression of the luciferase gene was largely restored (Figure 4.7C). Interestingly, cotransfection of Dnmt1-GFP and GAL4BD-LSH into the *DNMT3B*$^{-/-}$ cells also led to partial restoration of repression. This indicates that overexpression of Dnmt1 can partially compensate for loss of DNMT3B in GAL4BD-LSH mediated repression (Figure 4.7C). To investigate the role of DNMT3A in GAL4-LSH mediated repression I cotransfected the reporter constructs along with either DNMT3A-GFP or DNMT3B-GFP into *TKO* cells (Figure 4.7D). Interestingly, overexpression of DNMT3B-GFP alone was sufficient to rescue repression by both GAL4-LSH and GAL4BD-LSH(1-226) in these cells. Conversely, DNMT3A-GFP overexpression could only partially rescue repression by GAL4BD-LSH and not at all for GAL4BD-LSH(1-226) indicating that in this system DNMT3A is not absolutely required for LSH mediated repression. Again, DNMT1-GFP showed an ability to partially compensate for the lack of DNMT3A and DNMT3B (Figure 4.7D).

Together, these data indicate that DNMT1 and DNMT3B are both required for efficient GAL4-LSH mediated repression of the reporter gene but DNMT3A is not. The role of DNMT3A in LSH mediated repression in other cell types, particularly ES cells where Dnmt3A is known to be strongly expressed and to physically interact with Lsh should not be completely discounted. Interestingly, in the experiments

described above the catalytic activity of the DNMT proteins was not required for repression. To further investigate whether CpG methylation occurred upon targeting of GAL4BD-LSH to the reporter gene construct, I performed bisulphite sequencing on the promoter region of the luciferase reporter construct following 4 days targeting by GAL4BD-LSH (Figure 4.8). As can be seen from the sequenced clones no signification DNA methylation occurs at the TK promoter, the GAL4BS or the 5' of the luciferase gene after GAL4BD-LSH targeting compared to the GAL4BD targeted control (Figure 4.8). This is despite extended targeting of the region by GAL4BD-LSH leading to transcriptional repression and deacetylation of histones (Figures 4.5 and 4.6). Thus, the role of LSH in transcriptional repression is distinct from its presumed enzymatic activity and is independent of DNA methylation.

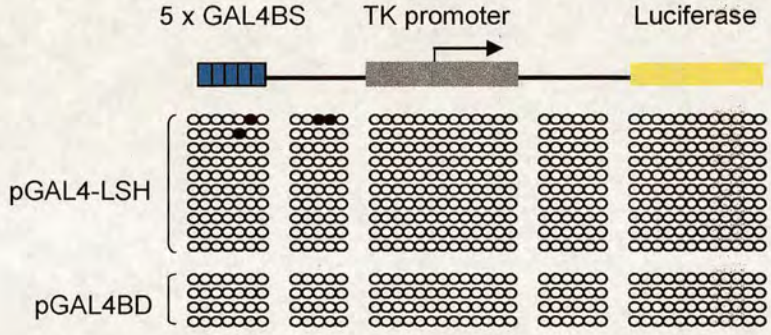


Figure 4.8. Recruitment of LSH to reporter plasmid does not immediately result in methylation of TK promoter sequences.
DNA was purified from HCT116 cells co-transfected either with luciferase reporter and either pGAL4BD or pGAL4BD-LSH 4 days post-transfection. DNA methylation was analyzed by bisulfite sequencing. The circles represent individual CpGs in the indicated plasmid regions. Each line of circles corresponds to an individual clone that was sequenced. Empty circles are unmethylated CpGs, black circles are methylated CpGs.

4.4.2.2 The interaction between LSH and HDACs is lost in cells deficient for DNMT1 or DNMT3B

Given that the repression by GAL4BD-LSH was sensitive to TSA and that HDAC1 and HDAC2 coimmunoprecipitated with LSH from nuclear extracts of wild-type HCT116 cells, I next examined whether these interactions remain intact in *DNMT1*$^{-/-}$

cells. Interestingly, when I immunoprecipitated endogenous LSH from either *DNMT1*$^{-/-}$ or *DNMT3B*$^{-/-}$ nuclear extract, I did not detect HDAC1 or HDAC2 in the α-LSH immunoprecipitation (Figure 4.9A-D, top panels; compare with Figure 4.6A). Consistent with this result the reciprocal α-HDAC1 and α-HDAC2 immunoprecipitations failed to co-immunoprecipitate LSH (Figure 4.9A-D, bottom panels; compare with Figure 4.6A). These experiments indicate that DNMT1 and/or DNMT3B could either directly or indirectly recruit HDAC1 and HDAC2 to LSH. Notably, the presence of both DNMTs is required to promote the association of HDACs with LSH.

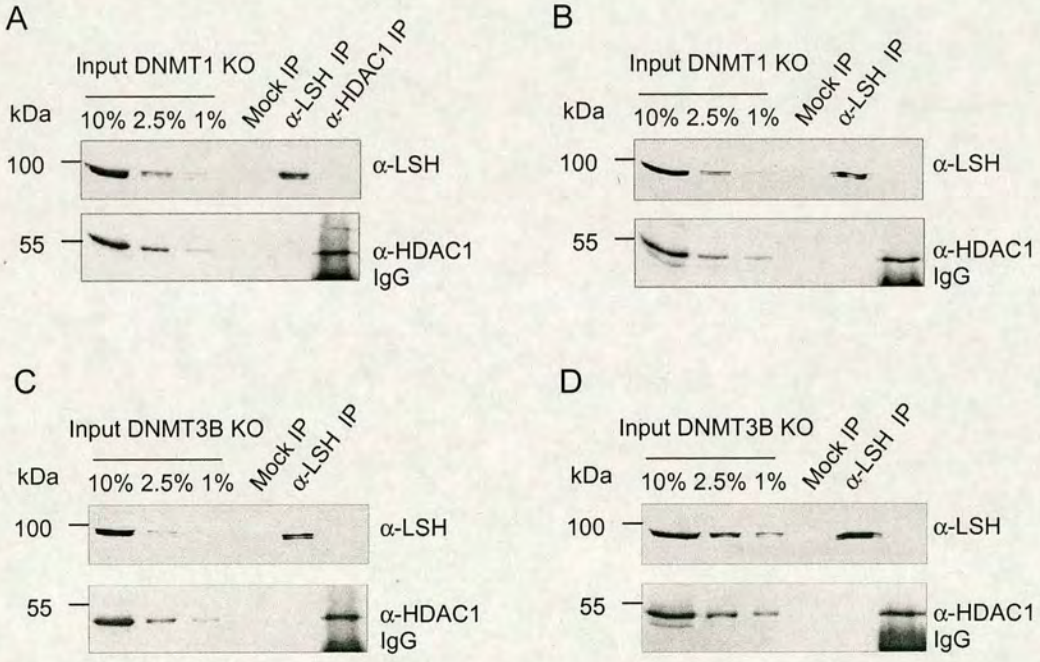


Figure 4.9. The interactions of LSH with HDACs are lost in DNMT KO cells.
A and B. Endogenous LSH, HDAC1, and HDAC2 could be efficiently immunoprecipitated (IP) from extracts of DNMT1 KO cells. However, LSH could not be detected in HDAC immunoprecipitations, nor was HDAC1 or HDAC2 detected in LSH immunoprecipitations. IgG, immunoglobulin G.
C and D. LSH, HDAC1 and HDAC2 do not coimmunoprecipitate from extracts of DNMT3B KO cells. Anti-HA antibodies were used as a control for non-specific interactions.

4.4.2.3 The TRD of LSH interacts with DNMT1 and DNMT3B

Thus far I have provided evidence supporting a functional role of DNMTs in GAL4BD-LSH mediated repression. To investigate whether LSH and DNMT1 physically interact I looked for an interaction between GAL4BD-LSH and Dnmt1-GFP by cotransfecting GAL4BD-LSH and Dnmt1-GFP in *DNMT1*$^{-/-}$ cells and used α-GAL4BD to immunoprecipitate LSH. DNMT1-GFP was detected by α-GFP antibodies in α-GAL4BD immunoprecipitates but not in control α-HA IPs, suggesting that GAL4BD-LSH and DNMT1-GFP interact with each other (Figure 4.10A, top panel). In a similar experiment I found that GAL4DB-LSH and DNMT3B-GFP cotransfected into *DNMT3B*$^{-/-}$ cells also co-immunoprecipitate (Fig. 4.10A bottom panel). Consistent with the reporter assays, when I cotransfected the cells with the coiled-coil TRD domain of LSH and either DNMT1-GFP or DNMT3B-GFP, I could detect GAL4BD-LSH(1-333) coimmunoprecipitating with each of the two DNMTs (Figure 4.10B). These experiments demonstrate that LSH and DNMTs physically interact *in vivo* and that the TRD of LSH is sufficient for these interactions.

4.4.2.4 LSH interacts indirectly with DNMT1 and HDAC1/2 via a direct interaction with DNMT3B

As the reporter assays and co-immunoprecipitations described above relied on overexpression of tagged proteins, we next examined whether endogenous LSH interacts with DNMTs in wild-type, *DNMT1*$^{-/-}$ and *DNMT3B*$^{-/-}$ HCT116 cells. We first sought to determine the precise nature of the KO cells lines. Consistent with other studies we found that an antibody against the C terminus of DNMT1 detected a truncated DNMT1 protein in nuclear extracts of *DNMT1*$^{-/-}$ and *DNMT1*$^{-/-}$/*DNMT3B*$^{-/-}$ *DKO1* cells (Egger et al., 2006; Spada et al., 2007) (Figure 4.11A, top panel). We found that this truncated form of DNMT1 was more abundant in the *DKO1* cell line than the *DNMT1*$^{-/-}$ cell line and that the cell line *DKO8* contained significant levels of full length DNMT1 (Figure 4.11A, top panel). We did not detect DNMT3B protein in extracts from any of the KO cell lines indicating that these cells are null for DNMT3B (Figure 4.11A, bottom panel). For further experiments we used extracts from four of the cell lines, HCT116 , *DNMT1*$^{-/-}$, *DNMT3B*$^{-/-}$ and *DNMT1*

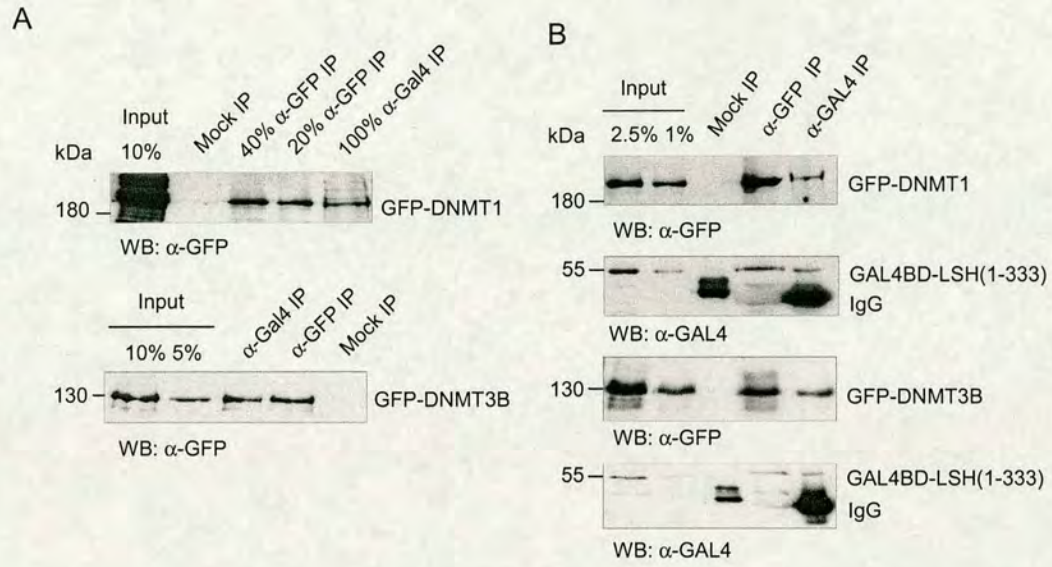


Figure 4.10. The coiled-coil TRD domain of LSH interacts with DNMTs.
A. GAL4BD-LSH coimmunoprecipitated with GFP-DNMT1 when both proteins were co expressed in DNMT1 KO cells. We could detect only about 20% of GFP-DNMT1 in the immunoprecipitations (IP) with anti-GAL4BD antibodies. GAL4BD-LSH coimmunoprecipitated more efficiently with GFP-DNMT3B when both proteins were co expressed in DNMT3B KO cells. WB, Western blot.
B. GAL4BD-LSH(1–333) protein, containing the coiled-coil TRD domain of LSH, coimmunoprecipitates with GFP-DNMT1 and GFP-DNMT3B from extracts of DNMT1 and DNMT3B KO cells, respectively. Anti-HA antibodies (mock immunoprecipitation) were used as a control.

$^{-/-}$/*DNMT3B*$^{-/-}$ *DKO1* for endogenous co -immunoprecipitations. We immunoprecipitated LSH from extracts of the different cell lines and then asked whether DNMT1 and DNMT3B could be detected in the immunoprecipitates (Figure 4.11B, top panel). DNMT3B was detected in α-LSH immunoprecipitates from HCT116 cells as expected, and was also detectable in α-LSH immunoprecipitates from *DNMT1*$^{-/-}$ cells (Figure 4.11B, middle panel). However, DNMT1 co-immunoprecipitated with LSH only from extracts of wild-type HCT116 cells and not from the extracts of *DNMT3B*$^{-/-}$ or *DNMT1*$^{-/-}$/*DNMT3B*$^{-/-}$ cells (Figure 4.11B, bottom panel). These results indicate that DNMT1 does not efficiently interact with LSH in the absence of DNMT3B. Conversely the presence of DNMT1 may not be required for the interaction of DNMT3B with LSH. This appears to be the case as

approximately equal amounts of DNMT3B co-immunoprecipitates with LSH from HCT116 and *DNMT1*$^{-/-}$ cells expressing a truncated DNMT1 that does not contain the DNMT3B interaction domain (Egger et al., 2006; Kim et al., 2002; Spada et al., 2007) (Figure 4.7A and 4.11A). Taken together with the luciferase reporter assays, these immunoprecipitation experiment suggest that LSH may exist in a complex with DNMT3B with or without DNMT1. However, an LSH complex containing DNMT1, HDAC1 and HDAC2 must also include DNMT3B.

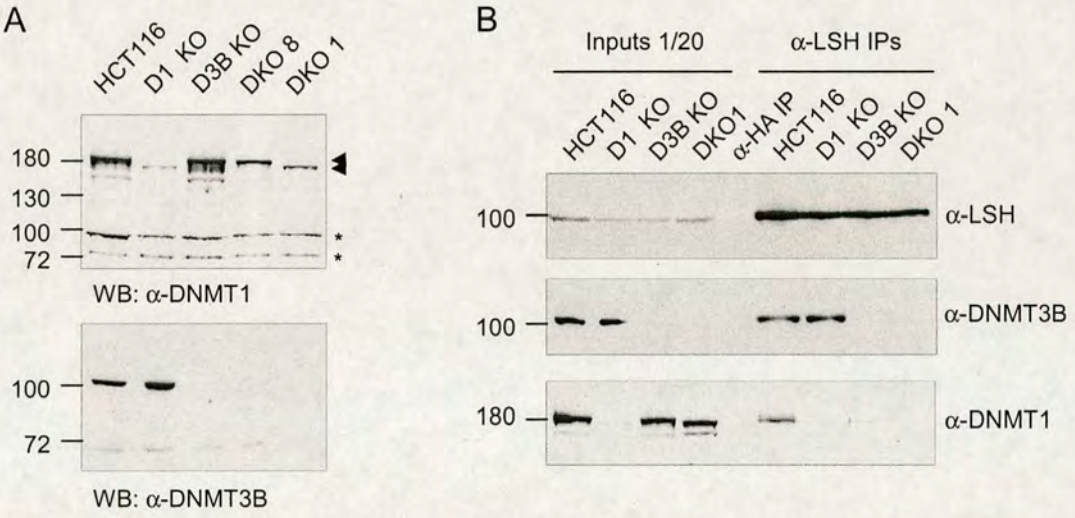


Figure 4.11. The interaction of LSH with DNMT1 in vivo requires DNMT3B.
These experiments were carried out by Dr Irina Stancheva.
A Western blots (WB) with antibodies against the C terminus of DNMT1 and the N terminus of DNMT3B on nuclear extracts of HCT116 and KO cell lines. Note that a truncated form of DNMT1 is detectable in *DNMT1*$^{-/-}$ (D1 KO) cells as well as in *DNMT1*$^{-/-}$/*DNMT3B*$^{-/-}$ DKO (DKO1) cells, while a second DKO cell line (DKO8) expresses full-length DNMT1. DNMT3B is detectable only in HCT116 and D1 KO cells. In the top panel, the full-length and the truncated DNMT1s are indicated with arrowheads. The asterisks indicate non-specific bands that serve as loading controls.
B Anti-LSH antibodies efficiently immunoprecipitate LSH from nuclear extracts of HCT116 and KO cells. In identical anti-LSH immunoprecipitations (IP), DNMT3B immunoprecipitates with LSH from HCT116 and D1 KO extracts, indicating that the N terminus of DNMT1 is not required for the interaction of DNMT3B with LSH. In contrast, DNMT1 coimmunoprecipitates with LSH only from HCT116 cells, suggesting that the presence of DNMT3B mediates the recruitment of DNMT1 to LSH.

To explore whether any of the proteins that co-immunoprecipitate with LSH bind it directly we expressed and purified from *E. coli* two glutathione S-transferase (GST) tagged recombinant mouse Lsh polypeptides. These were designated Lsh-N and Lsh-C (Figure 4.12A). Lsh-N contains amino acids 1 to 503 of Lsh including the N-terminal coiled-coil and the SNF2_N domain. A shorter fragment consisting of just the coiled-coils domain was insoluble in *E. coli* so this larger fragment had to be used. Lsh-C contains amino acids 248-883 and includes the SNF2_N and Helicase_C domains. We used these two polypeptides bound to glutathione-Sepharose beads or a control GST-GFP protein to pull down *in vitro* translated HA-tagged DNMT1(1-1125) and full length DNMT3B (Figure 4.12B). Interestingly, neither Lsh-N or Lsh-C could bind DNMT1(1-1125) in these assays (Figure 4.12B). Lsh-N but not Lsh-C or the GFP control could efficiently bind DNMT3B indicating that LSH and DNMT3B can interact directly (Figure 4.12B).

DNMT1 and DNMT3B are known to interact with each other via the extreme N-terminus of DNMT1 and the N-terminus of DNMT3B (Kim et al., 2002). As LSH did not co-immunoprecipitate with DNMT1 from extracts of *DNMT3B*$^{-/-}$ cells but did co-immunoprecipitate in the extracts of wild-type cells we next asked whether DNMT3B is required for the interaction between LSH and DNMT1. To investigate this we added increasing amounts of recombinant DNMT3B purified from insect cells to *in vitro* translated DNMT1(1-1125) and asked whether LSH-N could now pull down DNMT1(1-1125) (Figure 4.12C). We could detect increasing amounts of DNMT1(1-1125) being pulled down by GST-LSH-N only when the purified DNMT3B was present (Figure 4.12C). Consistent with our co-immunoprecipitation assays, these experiments provide evidence for a direct interaction between DNMT3B and the N-terminus of LSH, mediating an indirect interaction between LSH and DNMT1 (Figure 4.11B and 4.12B and C)

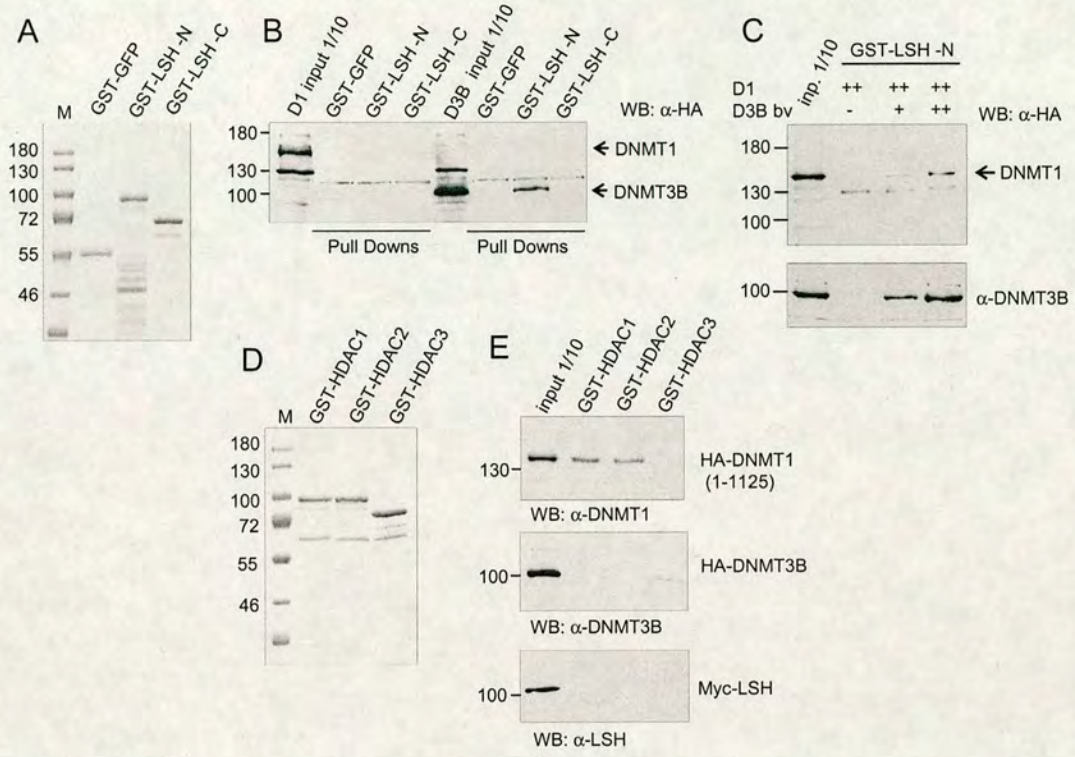


Figure 4.12. LSH directly interacts with DNMT3B but not with DNMT1 or HDACs
These experiments were carried out by Dr Irina Stancheva.
A. A Coomassie blue-stained gel shows GST-tagged purified GFP, LSH-N, and LSH-C proteins used in the pull-down assays (B and C).
B. The N terminus of LSH binds in vitro-translated DNMT3B. Neither the N terminus nor the C terminus of LSH pulls down in vitro-translated DNMT1 (amino acids 1 to 1125). WB, Western blot.
C. GST-LSH-N can pull down DNMT1 in the presence of recombinant DNMT3B (0.5 and 1 μg).
D. A Coomassie blue-stained gel shows purified GST-HDAC1, HDAC2, and HCAC3.
E. GST-HDAC1 and GST-HDAC2 pull down in vitro-translated DNMT1 but not DNMT3B or LSH.

We next sought to examine whether HDAC1 and HDAC2 could directly bind to LSH, DNMT1 or DNMT3B. For this experiment we expressed GST-tagged full-length HDAC1, HDAC2 and as a control, HDAC3 in *E. coli* and bound them to glutathione-Sepharose beads (Figure 4.12D). We used the immobilised HDAC proteins to pull down *in vitro* translated HA-tagged DNMT1(1-1125), HA-DNMT3B and Myc-LSH (Figure 4.11E). In agreement with previous reports, we could detect

an interaction between DNMT1(1-1125) and HDAC1 and HDAC2 but not with HDAC3 (Figure 4.12E; top panel) (Fuks et al., 2000; Robertson et al., 2000; Rountree et al., 2000). However, Myc-LSH could not bind any of the GST-HDACs in this *in vitro* system and HA-DNMT3B could only very weakly interact with HDAC3 (Figure 4.12E; middle and bottom panels). These *in vitro* experiments are consistent with our reporter assays and co-immunoprecipitation results and allow us to construct a model of GAL4BD-LSH mediated transcriptional repression in our reporter assay system (Figure 4.13).

The data presented from this study suggest that GAL4BD-LSH recruited to the reporter promoter by the GAL4BD recruits DNMT3B via a direct interaction. DNMT3B in turn interacts with the DNMT1-HDAC1-HDAC2 complex and recruits it to LSH. The HDAC containing LSH complex is then associated with the promoter region of the reporter gene leading to deacetylation of histone H3 and H4 and transcriptional repression (Figure 4.13). An alternative, thus far unexplained mechanism of recruiting the DNMT1-HDAC complex to LSH may also exist (Figure 4.13; arrow) as it which would explain the partial rescue of repression by GFP-DNMT1 in *DNMT3B*$^{-/-}$ cells (Figure 4.7C and D). Our model does not indicate how this repressive complex is recruited to chromatin *in vivo*. This may occur via one or a combination of the members binding directly to chromatin or via an as yet unidentified recruitment factor.

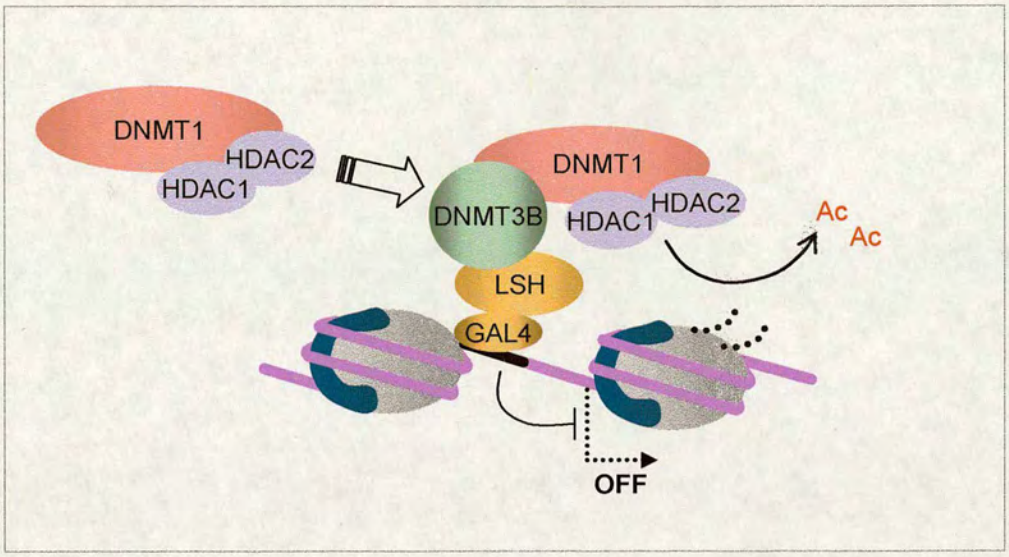


Figure 4.13. A model of how the LSH-associated protein complex acts to repress transcription.
Our experiments are consistent with a model where DNA-bound LSH recruits a complex that includes DNMT3B, DNMT1, HDAC1, and HDAC2. LSH-associated HDACs remove acetyl groups (Ac) from histone tails, generating deacetylated chromatin incompatible with transcriptional activation. LSH does not directly interact with DNMT1 and HDACs but requires DNMT3B for the assembly of the repressive complex. On the other hand, HDAC1 and HDAC2 require both DNMT1 and DNMT3B for association with the LSH complex. The order of these interactions explains why in cells expressing N-terminally truncated DNMT1, which does not bind DNMT3B, or in cells lacking DNMT3B, LSH-mediated repression is disrupted.

4.5 Summary

The plant SNF2 protein DDM1 and its mammalian homologue LSH were initially discovered as proteins essential for the establishment of proper DNA methylation patterns *in vivo* (Dennis et al., 2001; Jeddeloh et al., 1999; Vongs et al., 1993). DDM1 deficient plants and mice lacking functional Lsh develop with dramatically reduced levels of DNA methylation and are defective in silencing transposable elements and a number of genes (Fan et al., 2005b; Huang et al., 2004; Miura et al., 2001). Despite the requirement of Lsh for genomic methylation it is not required for embryonic development but leads to death a few hours after birth and runted growth (Geiman et al., 2001). This phenotype is dissimilar to the Dnmt mutants indicating that the role of Lsh in DNA methylation is unlikely to act as a stimulatory factor for a single Dnmt in a manner similar to Dnmt3L (Kaneda et al., 2004). Cytological studies and experiments with ES cells have shown that although Lsh co localises with Dnmt1 at replication foci during late S phase, Lsh is not required for maintenance methylation of satellite DNA or on replicating episomal plasmids. Lsh has also been shown to be required for *de novo* methylation and to co-immunoprecipitate with Dnmt3a and Dnmt3b (Zhu et al., 2006). Taken together these studies indicate that the primary role of Lsh is in *de novo* methylation and it is dispensable for maintaining methylation during DNA replication.

Despite the crucial role of Lsh in assisting efficient DNA methylation in mammalian cells, very little is known about how it interacts with the DNA methylation machinery. As mentioned above the interaction partners of SNF2 enzymes are crucial to their function and identifying them provides much important information to dissecting their function. To provide insight into the molecular role of LSH I have attempted to identify proteins that associate with LSH in a stable complex. By combining the biophysical properties of LSH obtained using gel filtration chromatography and sucrose gradients I have shown that LSH is present mostly as a free monomeric peptide in HCT116 and other human and mouse cells. I have furthermore been unable to detect an LSH complex at any cell cycle stage or with

any protein extraction method used, which indicates that unlike most SNF2 proteins LSH does not stably interact with other proteins. Therefore, using biochemical fractionation to purify a low abundance, transient or unstable LSH complex would not have been feasible. The lack of detectable, stably bound interaction partners is unusual for SNF2 enzymes and indicates that LSH most likely interacts transiently with the DNA methylation machinery.

In addition to using protein chromatography to identify a stable LSH complex I have also utilised a Y2H screen to identify transiently interacting proteins. This screen did not identify any interacting proteins which I believe was not due to failure of the Y2H screen or a lack of other LSH interacting proteins. On the contrary, I was able to demonstrate that LSH prevents activation of the Y2H reporter genes through its N-terminal coiled-coil domain and this was the most likely explanation the Y2H screen did not identify any targets. Unfortunately, this domain of LSH is the one most likely to interact with other proteins, as shown in further experiments, so this type of screen is not an effective tool to screen for proteins that interact with LSH. Although this result complicated identification of LSH interacting partners, it allied with many studies on the transcriptional effects of SNF2 enzymes, led me to ask whether LSH can repress transcription in mammalian cells.

I used luciferase reporter assays to show that LSH can act as a transcriptional repressor in mammalian cells and used these assays as a tool to identify proteins that functionally interact with LSH. By recruiting various deletion fragments of LSH to the promoter of the luciferase reporter via GAL4 binding sites I was able to show that the N-terminal coiled-coil domain of LSH spanning amino acids 1 to 226 was necessary and sufficient for transcriptional silencing of the reporter. I designated this fragment of LSH the TRD and reasoned it likely that this region interacts with other co-repressor proteins that modify chromatin into a transcriptionally nonpermissive state.

I was able to use the reporter assays to investigate the mechanism of LSH mediated transcriptional repression. I first showed that repression is sensitive to TSA treatment

indicating a role of HDAC enzymes in this process. This led me to perform co-immunoprecipitation experiments between LSH and HDAC1 and HDAC2. I found that LSH successfully co-immunoprecipitated both of these proteins indicating a physical interaction between these proteins. Furthermore, I was able to demonstrate that targeting LSH to the promoter of the reporter plasmid induced deacetylation of histone H3K9 and H4K12. Together, these data indicate that LSH can function as an HDAC dependent transcriptional repressor.

Regions of transcriptionally silenced chromatin share a number of signature chromatin signals. These include histone hypoacetylation, H3K9me3 and H3K27me3 and DNA methylation (Trojer and Reinberg, 2007). Given that these chromatin marks are found at silenced regions, that Lsh has a key role in DNA methylation and that a previous study has shown an interaction between Lsh and Dnmt3a and Dnmt3b, I also examined whether LSH mediated repression requires DNMT enzymes. In human colorectal carcinoma HCT116 cells with targeted disruptions of the DNMT3B (*DNMT3B*$^{-/-}$) gene, the DNMT3A and DNMT3B (*TKO*) genes as well as cells expressing low levels of a truncated DNMT1 proteins (*DNMT1*$^{-/-}$), I observed a significant reduction in LSH mediated repression. I was able to rescue repression to wild type levels by over expressing DNMT3B in the *DNMT3B*$^{-/-}$ or *TKO* cell lines and DNMT1 in the *DNMT1*$^{-/-}$ cell line indicating that DNMT3B and DNMT1, but not DNMT3A is involved in recruiting HDAC1 and HDAC2 to LSH in this system. Interestingly, the DNMT function of DNMT3B or DNMT1 was not required for this rescuing ability and targeting of LSH to the reporter construct did not lead to CpG methylation. This result is in agreement with previous studies that have shown the repressive ability of DNMTs to be independent of their role in DNA methylation (Fuks et al., 2001; Robertson et al., 2000; Rountree et al., 2000). In addition, studies on the Oct4 locus, which becomes transcriptionally silenced during development indicate that histone deacetylation is an early event that precedes DNA methylation *in vivo* (Deb-Rinker et al., 2005; Feldman et al., 2006; Jeong-Heon Lee, 2004). Further experiments confirmed that LSH is able to co-immunoprecipitate with both tagged and endogenous DNMT3B and DNMT1 indicating that these proteins functionally interact. Interestingly, I was unable to co-immunoprecipitate LSH with

HDACs from extracts of *DNMT3B*$^{-/-}$ and *DNMT1*$^{-/-}$ cells. This indicates that the primary role of DNMT1 and DNMT3B in LSH mediated repression in this system is to provide a scaffold for the interaction between LSH and HDAC1 and HDAC2. This was further verified as LSH did not show a direct interaction with either HDAC1 or HDAC2 in *in vitro* pull down assays.

Previous studies have reported that both DNMT1 and DNMT3B interact with HDAC1 and HDAC2 *in vitro* and *in vivo* (Fuks et al., 2000; Geiman et al., 2004a; Robertson et al., 2000). It was therefore possible that either one of the DNMTs, or a combination of both, could recruit HDACs to LSH-targeted chromatin. Given that LSH mediated repression was deficient in both *DNMT3B*$^{-/-}$ and *DNMT1*$^{-/-}$ cells the association of the different members of this complex was investigated. Interestingly, we found that LSH was able to co-immunoprecipitate with DNMT3B in all the cell lines tested but DNMT1 could not interact with LSH in cells lacking DNMT3B. This indicates that the interaction between LSH and DNMT1 requires DNMT3B *in vivo*. This result was confirmed using pulldown assays, which detected a direct interaction between LSH and DNMT3B but no direct interaction between LSH and DNMT1. In addition to this, only DNMT1 was shown to interact with HDAC1 and HDAC2 in our pull down assays. Taken together, the data suggest that DNMT1-bound HDAC1 and HDAC2 are recruited to LSH indirectly via DNMT3B (Figure 4.13). Interestingly, we observed that when DNMT1 was overexpressed in *DNMT3B*$^{-/-}$ cells, it could partially rescue the transcriptional repression effect of GAL4BD-LSH. As we did not detect a direct interaction between LSH and DNMT1, it is possible that recruitment of DNMT1 to LSH may be an artefact of overexpression or may occur by some other, inefficient mechanism. The most likely candidate for this would be DNMT3A due to its previously detected interaction with Lsh and its homology and functional similarity to DNMT3B (Zhu et al., 2006). This does not appear likely as DNMT3A expression was not detected in these cells and overexpression of DNMT3A in *TKO* cells did not completely rescue LSH mediated repression. However, I do not wish to rule out the potential role of DNMT3A in this process as it may have an effect in cells where it is normally highly expressed. Given that overexpression of DNMT3B in *DNMT1*$^{-/-}$ cells transfected with GAL4DB-LSH

also reduced the expression of the luciferase reporter by about 25-30%, it is possible that DNMT3B can either function in transcriptional repression independently of the truncated DNMT1 or, to some extent but not very efficiently, interact with the C terminus of DNMT1. The second interpretation seems plausible, since the N-terminal portion of DNMT1(1-1125) did not fully rescue LSH-mediated repression in DNMT1 KO cells compared to the full-length DNMT1, indicating a potential role for the C-terminus of DNMT1 in this process. Again, I would caution against drawing strong conclusions from this result as it is based on experiments involving protein overexpression.

In summary, I have used a variety of techniques to investigate whether and how LSH interacts with the DNA methylation machinery in human cells. I have identified an LSH associated complex containing at least four proteins: DNMT3B, DNMT1, HDAC1 and HDAC2. This complex is not detectable using conventional chromatography techniques indicating that it is either very low abundance, transient or not stable to the methods used to extract proteins. The ability to detect protein-protein interactions by co-immunoprecipitation and *in vitro* pull down but not by chromatography is not particularly unusual and has been previously shown for other SNF2 and chromatin related enzymes such as the interaction between SNF2h and DNMT3B and the interaction between MeCP2 and Sin3A (Geiman et al., 2004a; Klose and Bird, 2004). I was able to utilise the ability of GAL4BD-LSH to repress transcription of a targeted reporter to dissect proteins that functionally interact with LSH for this process. I subsequently confirmed these interactions using physical methods such as co-immunoprecipitate and *in vitro* pull down. I showed that repression is dependent on recruitment of HDACs via DNMT3B and DNMT1 and but does not required the DNMT activity of the enzymes or lead to DNA methylation. Thus although LSH protein serves as a scaffold for assembly of the DNMT/HDAC complex the primary function of the LSH complex in this system may not be to methylate DNA but to establish deacetylated inactive chromatin. Transcriptional repression caused by deacetylation of histone tails by LSH-interacting HDAC1/2 can be viewed as an initial and, perhaps reversible step in LSH-mediated gene silencing. A longer-term association of LSH with specific loci and a

persistently high local concentration of DNMT1 and DNMT3B may result in methylation of CpGs at these loci. As the experiments described utilise a targeted reporter system, it is important to confirm them on endogenous targets for two reasons. Firstly, it is important to establish how general these findings are and to test whether endogenous LSH targets lose histone acetylation and transcriptional repression upon loss of LSH. A number of studies have provided evidence that in $Lsh^{-/-}$ MEFs histone hyperacetylation and transcriptional activation occurs at pericentromeric major satellite repeats and at a subset of Hox genes believed to be Lsh targets (Huang et al., 2004; Xi et al., 2007). Thus my data provides a mechanistic explanation of these findings by indicating that gain of histone acetylation is due to loss of HDAC1 and HDAC2 from these loci. Further experiments in human cells using siRNA against LSH and looking at a number of specific LSH target genes would also be useful to confirm these results. Secondly, it will be important to determine how this complex is recruited to chromatin *in vivo* and how it is targeted to specific genomic loci. As the reporter assays used the GAL4BD to target the LSH complex to chromatin it is impossible to determine if LSH recruits the complex *in vivo*. Thus, a ChIP-chip approach or similar should be utilised to identify genomic targets of LSH. Upon identification of targets different members of the complex can be knockdown using siRNA to determine which member of the complex is responsible for recruiting it to chromatin. In addition to this a biochemical approach can be undertaken to determine which member of the complex binds DNA and chromatin most strongly *in vitro*. It is also unclear how LSH and/or the LSH-associated DNMT complex could be recruited to specific genomic loci. One could envisage that as a potential DNA translocase LSH continuously scans chromatin for regions that challenge the processivity of DNMT enzymes or alternatively, sequence specific recruitment to particular genomic loci may occur.

Experiments with unmethylated episomal plasmids capable of replicating in mammalian cells indicate that Lsh-facilitated DNA methylation was observed weeks rather than days after the plasmids were introduced into these cells, indicating the long timescale of this process (Zhu et al., 2006). Together with my results, this data suggests that LSH cooperates with DNMTs and HDAC1/2 to act as a general

transcriptional repressor in mammalian cells which may lead to later DNA methylation events.

The previous chapter described experiments investigating the catalytic function of LSH. These experiments showed that LSH is a DNA stimulated ATPase with an activity much lower than that recorded for other SNF2 enzymes. I reasoned that LSH activity may depend on additional protein co-factors so sought to identify proteins that interact with LSH *in vivo*. This chapter has outlined experiments that have identified a transient, unstable or low abundance complex of LSH, DNMT3B, DNMT1, HDAC1 and HDAC2. If LSH activity is modulated by protein co-factors then these proteins would be interesting to test.

Chapter five - Analysis of LSH-DNMT complex *in vitro*

5.1 Outline – LSH is required for high levels of DNA methylation in mammals

The assembly of euchromatic DNA into chromatin provides an inhibitory structure to enzymes that require access to it. The demonstration that LSH is required for high levels of DNA methylation in mammals outlined the potential importance of SNF2 chromatin remodelling enzymes to this process. SNF2 enzymes use the energy derived from ATP hydrolysis to transiently disrupt histone:DNA contacts in a manner that increases DNA accessibility (Narlikar et al., 2002). Thus, an attractive hypothesis for the role LSH plays in DNA methylation is that it disrupts nucleosomal DNA in a manner that facilitates its accessibility to DNMT enzymes. If this is the role of LSH, then it could be envisaged to occur either by repositioning nucleosomes or by liberating loops of DNA from their surface. Another characteristic activity displayed by SNF2 enzymes is the ability to translocate DNA (Lia et al., 2006; Saha et al., 2002, 2005; Whitehouse et al., 2003). Thus, if LSH also possess this activity it may be involved in scanning the genome for DNMT target sites. Alternatively, LSH may have a more passive role that does not involve its putative ATPase activity, for example targeting or tethering DNMTs to chromatin.

5.1.1 LSH does not exhibit nucleosome remodelling activity *in vitro*

In chapter 3 I discussed experiments designed to test whether LSH is an active ATPase that can remodel nucleosomes *in vitro*. These experiments would be indicative of the *in vivo* function of LSH. Interestingly, although exhibiting modest DNA stimulated ATPase activity, LSH did not display detectable levels of nucleosome remodelling activity using two distinct assays. These assays were a classical sliding assay to assess the ability of LSH to slide nucleosomes on DNA, and an SssI methyltransferase protection assay to investigate if LSH can liberate DNA

loops from the surface of the nucleosome. However, these assays have limitations as chromatin remodelling may not necessarily lead to nucleosome repositioning and the SssI assay is dependant on DNA loops being formed at CpG sites. Therefore, I am not able to completely rule out LSH having chromatin remodelling ability. It is also possible that LSH requires a co factor, or specific stimulus to achieve full activity and efficiently remodel nucleosomes. For example, interactions with partner proteins may lead to conformational changes in LSH which could stimulate ATPase activity.

5.1.2 LSH interacts with DNMT1 and DNMT3B *in vivo*

The ATPase activity of purified recombinant SNF2 enzymes is usually highly similar to that of purified native protein complexes. However, in the case of ISWI, it is clear that its activity is modulated by its cofactor Acf1 (Langst et al., 1999). Thus, it is possible that LSH requires additional proteins to remodel chromatin to its full capacity. Chapter 4 described experiments that aimed to identify proteins that interact with LSH *in vivo*. Although, no proteins were identified that stably bound LSH, I was able to identify a transient, or low abundance repressive complex of LSH, DNMT3B, DNMT1, HDAC1 and HDAC2. Within this complex, it appears that LSH directly interacts with DNMT3B and indirectly with the other members via this interaction. Thus, if LSH activity is dependant on a protein co-factor it would be useful to test DNMT3B. A further interesting question raised by these experiments is how this complex is targeted to chromatin *in vivo*.

This chapter will discuss experiments devised to ask if a single member of the repressive LSH complex could have a prominent role in tethering it to DNA. These experiments provide the basis for future work to identify how this complex is targeted *in vivo*. I will also discuss experiments to test if LSH activity is stimulated by DNMT3B and reciprocally, whether the ability of DNMTs to methylate nucleosomes is stimulated by LSH.

5.2 DNA binding properties of LSH and DNMT3B

5.2.1 Comparative binding of LSH, DNMT1 and DNMT3B to DNA

The experiments described in chapter 4 gave rise to a model of LSH mediated transcriptional repression requiring DNMT3B and DNMT1. The model did not allude to a mechanism by which this complex is targeted to chromatin or which member/s of the complex is/are required for binding DNA (Figure 4.13). To address these questions I expressed recombinant 6xHIS tagged DNMT1 and DNMT3B in insect cells using the baculovirus system and purified these proteins using Cobalt affinity chromatography (Figure 5.1A). I used the recombinant proteins to investigate whether a single member of this complex has a dominant role in binding DNA. To assay this I optimised electrophoretic mobility shift assays (EMSAs) using equal concentrations of the different recombinant protein and a variety of competitor DNA (Figure 5.1B). As can be seen from this experiment LSH and DNMT3B both shifted the DNA probe with LSH forming a single DNA:protein complex and DNMT3B giving rise to multiple products. DNMT1 showed little or no ability to bind DNA in this assay. DNMT3B also formed non-specific aggregates and addition of competitor DNA, in particular dAdT gave rise to clearer DNA:protein complexes. The additional bands formed with DNMT3B were not removed upon addition of unlabelled competitor and their origin is not known. As the EMSA with dAdT gave rise to the clearest DNA:protein complexes of the conditions tested it was added in subsequent experiments. I repeated the EMSA over a range of protein concentrations for LSH and DNMT3B with two distinct DNA probes. These probes were 36mers either containing a central CG site and runs of Gs or no central CG site and no runs of Gs. In these experiments it was clear that LSH bound the probe with ~10 fold greater affinity than DNMT3B (Figure 5.1C; left and right panels). This result perhaps suggests that LSH is the member of the complex that is responsible for tethering it to DNA. This data provides tentative evidence for this model and requires further verification. In particular, future work could focus on confirmation of this result *in vivo* using ChIP and siRNA knockdown against LSH, DNMT1 and DNMT3B.

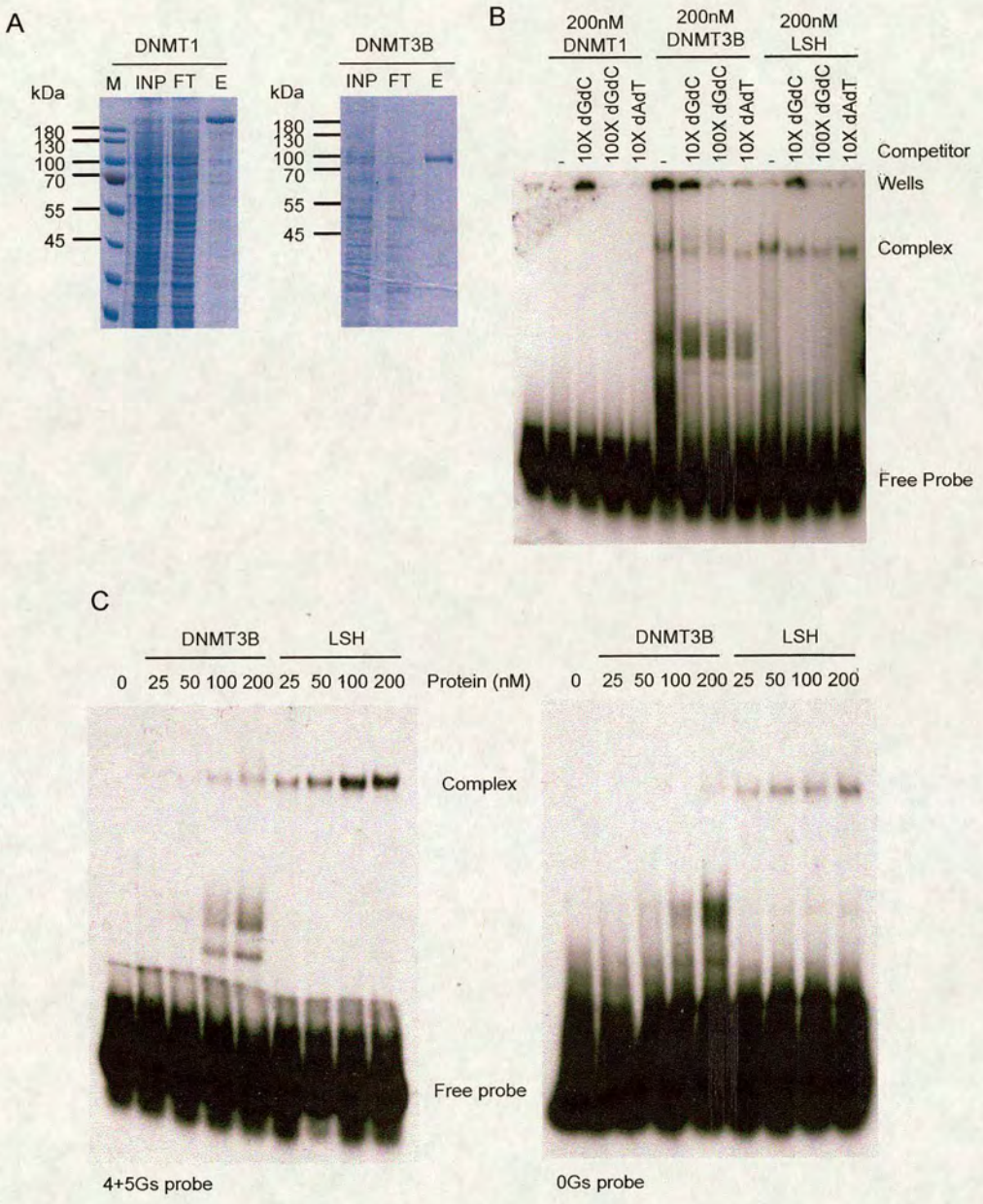


Figure 5.1. LSH binds DNA with greater affinity than DNMTs.
A. Coomassie stained gel outlining the purification of HIS-DNMT1 and HIS-DNMT3B via Cobalt affinity chromatography. 5µl of cell lysate input (INP), 5µl of flow-through (FT) and 1µl of elution (E) from the Cobalt column loaded. Size markers (M) in kDa listed to the left.
B. Electrophoretic mobility shift assay (EMSA) with 200nM recombinant DNMT1, DNMT3B and LSH (indicated above the gel) and radio labelled 4+5Gs DNA probe. 10nM probe with or without various non specific competitor DNA (indicated above the gel) was incubated with protein for 30 minutes at 37°C. Free probe separated from specific complexes by native polyacrylamide gel electrophoresis. Following electrophoresis, the gel was dried and

exposed to photographic film. Free probe, specific complexes and wells indicated to the right.

C. Electrophoretic mobility shift assay (EMSA) with increasing concentrations of DNMT3B or LSH (indicated above the gel) and 10nM 4+5Gs (left panel) or 0Gs (right panel) radio labelled probe. Radio labelled probe was incubated with protein for 30 minutes at 37°C. Free probe separated from specific complexes by native polyacrylamide gel electrophoresis. Following electrophoresis, the gel was dried and exposed to photographic film. Free probe and specific complexes indicated between the two panels.

5.3 Analysis of the effect of LSH on DNMT activity *in vitro*

Biochemical characterisation of LSH *in vitro* has shown that it is a DNA binding protein and a weak DNA stimulated ATPase. Both of these activities are relatively inefficient compared to other SNF2 enzymes, in particular the ability to hydrolyse ATP. I have not been able to detect the ability of LSH to remodel nucleosomes indicating that either this is not an activity LSH possesses or that the assays I have used are not sensitive enough. It is possible that LSH requires additional co factors in order to efficiently hydrolyse ATP and remodel chromatin. I have identified a LSH complex that contains DNMT3B, DNMT1, HDAC1 and HDAC2. Thus, if LSH requires a protein co-factor to exhibit full ATPase and chromatin remodelling activity it may be one of these proteins. As DNMT3B interacts with LSH directly, I thought this protein would be most likely to influence these potential abilities. The role of LSH in DNA methylation is well established genetically but there is no data on its molecular function in this process. I attempted to elucidate this function by asking what effect LSH has on the ability of DNMTs to methylate a variety of templates *in vitro.*

5.3.1 Recombinant DNMT3B and DNMT1 do not stimulate the activity of LSH *in vitro*

In order to test if recombinant DNMT3B and DNMT1 can affect the rate of ATP hydrolysis by LSH, reconstitution of this complex was attempted. Equimolar ratios of these three proteins were mixed on ice for one hour and formation of a complex monitored by Superose 12 gel filtration. Formation of a stable protein complex should lead to a shift in elution of the three proteins to a higher molecular weight. No shift in the elution fraction was observed following incubation indicating that the majority of these proteins stayed monomeric (Figure 5.2A). As I did not detect a stable complex of these three proteins *in vivo* I reasoned it possible that this complex may be functional if only transiently associated. Thus, I assayed its ATPase activity upon addition of a range of stimuli. Similar to the previously observed results with recombinant LSH, the pre-mixed proteins displayed a low level of ATPase activity

that was stimulated ~2 fold upon addition of DNA or nucleosomes (Figure 5.2B). Therefore, the presence of DNMT1 and DNMT3B do not significantly alter the ATPase activity of LSH *in vitro*.

5.3.2 Recombinant DNMT1 and DNMT3B have methyltransferase activity towards DNA

In order to assess if the presence of LSH affects the activity of DNMT1 and DNMT3B I tested the activity of the recombinant DNMTs I had purified. I incubated 40nM of each enzyme with poly dGdC template in the presence of ^{3}H-SAM for 1h at 37C. The poly dGdC substrate contains many CpG sites that are the substrate of DNMTs and is therefore a good initial starting point for determining DNMT activity. After incubation I spotted the reactions on DE81 filter paper and washed extensively with ammonium carbonate, water and ethanol. Incorporation of ^{3}H to the washed DNA was determined by scintillation counting. Both of the recombinant proteins incorporated ^{3}H into the DNA compared to no protein controls indicating that they are active DNMTs (Figure 5.3A). 0.05u of the bacterial DNMT SssI served as a positive control. Interestingly, the *de novo* activity of DNMT3B to this substrate was very low, and much lower than DNMT1, which is in agreement with previously published data (Li et al., 2007a; Okano et al., 1998a) (Figure 5.3A). Addition of equimolar amounts of LSH along with ATP did not have a stimulatory effect on the activity of the methyltransferases towards this substrate (Figure 5.3A).

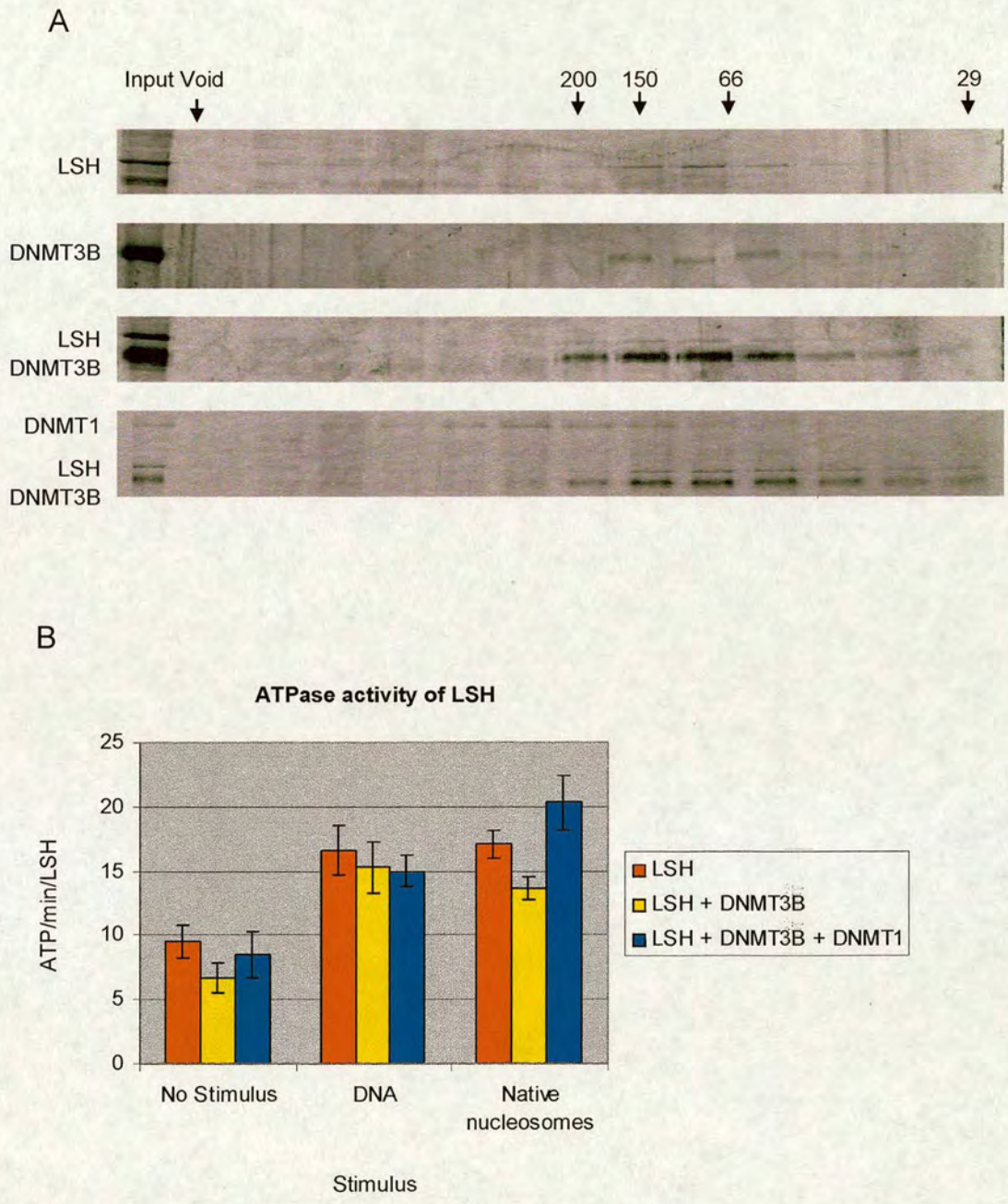
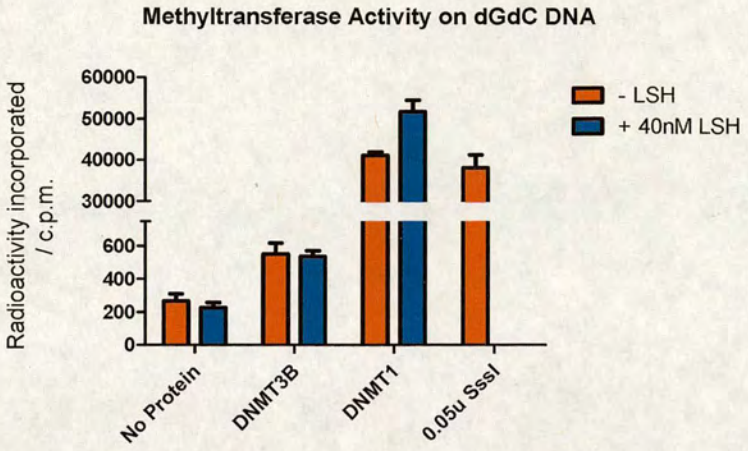


Figure 5.2 DNMTs do not stimulate LSH activity

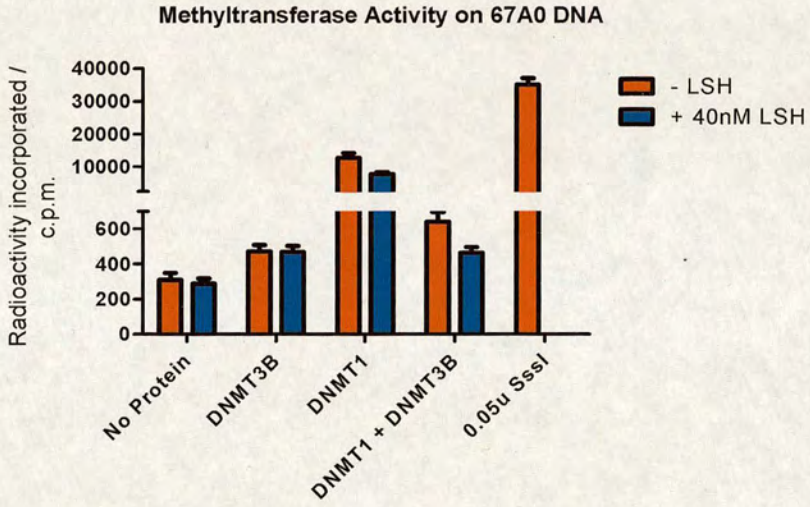
A. Recombinant LSH, DNMT1 and DNMT3B purified from insect cells were mixed on ice for one hour and fractionated on a Superose 12 size exclusion column. The fractions were run on sodium dodecyl sulfate-polyacrylamide gels and silver stained. The molecular masses of marker proteins used to calibrate the column are indicated at the top.

B. Recombinant LSH, DNMT3B and DNMT1 (100nM of each) were premixed and used in ATPase assays with 50μM ATP in the presence of 200nM DNA or nucleosomes (indicated on the X axis). The number of ATP molecules hydrolysed per minute per LSH calculated by analysis of thin layer chromatography plates (Y axis). ATP hydrolysis rate observed for LSH (red bars), LSH+DNMT3B (yellow bars) or LSH+DNMT3B+DNMT1 (blue bars) for each stimulus is displayed. Error bars denote standard deviation from the mean of three independent experiments.

A


Methyltransferase Activity on dGdC DNA
Radioactivity incorporated / c.p.m.
60000
50000
40000
30000
600
400
200
0
- LSH
+ 40nM LSH
No Protein
DNMT3B
DNMT1
0.05u SssI

B


Methyltransferase Activity on 67A0 DNA
Radioactivity incorporated / c.p.m.
40000
30000
20000
10000
600
400
200
0
- LSH
+ 40nM LSH
No Protein
DNMT3B
DNMT1
DNMT1 + DNMT3B
0.05u SssI

C

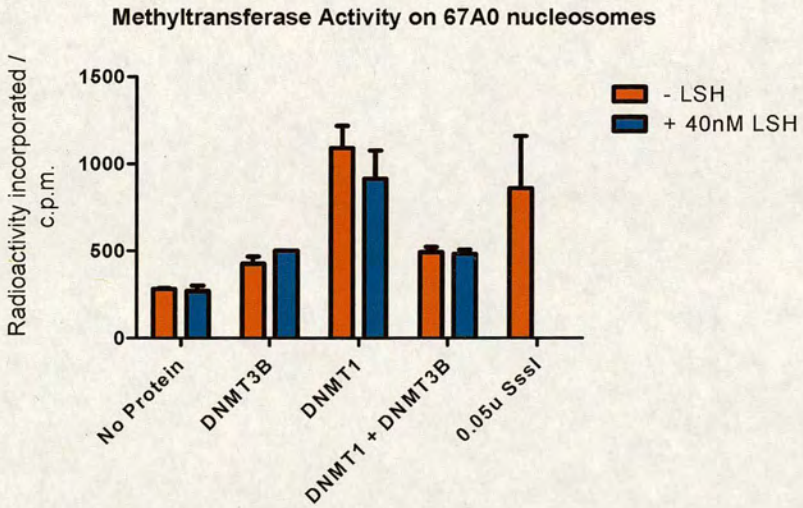

Methyltransferase Activity on 67A0 nucleosomes
Radioactivity incorporated / c.p.m.
1500
1000
500
0
- LSH
+ 40nM LSH
No Protein
DNMT3B
DNMT1
DNMT1 + DNMT3B
0.05u SssI

Figure 5.3 LSH does not stimulate DNMTs activity
A. 40nM recombinant DNMT1, DNMT3B or 0.05 units SssI was used in DNA methyltransferase assays on 200ng dGdC template in the absence (red bars) or presence (blue bars) of 40nM LSH. Reactions including ^{3}H-SAM were incubated at 37°C for 1 hour, spotted on DE81 paper, washed extensively and radioactivity incorporation measured by scintillation counting. Error bars denote standard deviation from the mean of three independent experiments.
B. 40nM recombinant DNMT1, DNMT3B, DNMT1 + DNMT3B or 0.05 units SssI was used in DNA methyltransferase assays on 100nM 67A0 DNA in the absence (red bars) or presence (blue bars) of 40nM LSH. Reactions including ^{3}H-SAM were incubated at 37°C for 1 hour, spotted on DE81 paper, washed extensively and radioactivity incorporation measured by scintillation counting. Error bars denote standard deviation from the mean of three independent experiments.
C. 40nM recombinant DNMT1, DNMT3B, DNMT1 + DNMT3B or 0.05 units SssI was used in DNA methyltransferase assays on 100nM 67A0 nucleosomes in the absence (red bars) or presence (blue bars) of 40nM LSH. Reactions including ^{3}H-SAM were incubated at 37°C for 1 hour, spotted on DE81 paper, washed extensively and radioactivity incorporation measured by scintillation counting. Error bars denote standard deviation from the mean of three independent experiments.

5.3.3 DNMT1 and DNMT3B show DNMT activity towards 67A0 DNA which is not stimulated by LSH

In addition to showing activity towards the CpG rich poly dGdC I tested the activity of purified DNMTs towards the 67A0 nucleosome positioning DNA sequence (Flaus and Richmond, 1998). This sequence contains 10 CpGs in total so approximately 1 per 20 bp of DNA. DNA methyltransferase assays showed that the purified DNMTs could methylate this template, but with lower affinity compared to the poly dGdC template presumably due to the reduction of total number of CpGs in the substrate. This was particularly noticeable for DNMT1 which showed a ~4 fold reduction in activity towards this substrate (Figure 5.3B). Surprisingly, when mixed together, DNMT3B appeared to have an inhibitory effect on DNMT1 when added to the same reaction. This could be explained if the less active DNMT3B is saturating the substrate and preventing binding by the more active DNMT1. Addition of LSH and ATP to this assay did not increase the methyltransferase activity of DNMT1, DNMT3B or DNMT1 mixed with DNMT3B towards this substrate (Figure 5.3B). I next assayed the activity of DNMT1 and DNMT3B on nucleosomal DNA using the 67A0 template assembled into nucleosomes. I found that the ability of DNMT1 and SssI to methylate the 67A0 nucleosomal arrays was strongly inhibited (>10 fold),

with the low level of DNMT3B activity not affected (Figure 5.3C). Addition of LSH and ATP did not significantly increase the activity of the DNMTs towards this substrate (Figure 5.3C). In conclusion, it appears that the presence of LSH in these assays does not enhance the ability of DNMT1 or DNMT3B to methylate DNA or nucleosomes.

5.3.4 Co-purification of LSH-DNMT3B complex from insect cells

For LSH, DNMT3B and DNMT1 to stimulate each others activity they may have to be bound together in an active complex. The previous experiments used individual purified recombinant proteins, mixed in equimolar ratios. Mixing purified protein does not guarantee that an active complex will be formed. In fact, when run on Superose 6 or Superose 12 sizing columns I did not see LSH and DNMT3B co-eluting with DNMT1 indicating a complex between these proteins had not been formed (Figure 5.2A). Additionally, I did not detect a shift in elution of LSH when mixed with DNMT3B as would be expected if these proteins were bound to each other (Figure 5.2A). In order to ensure these proteins were present together in a complex I attempted to co-purify them following expression in insect cells. I used PCR mutagenesis to remove the 6xHIS tags from the pFAST-BAC vectors and used them to make recombinant baculovirus. I infected insect cells at a multiplicity of infection of ~10 with each recombinant virus to ensure co-infection and expression. Initial attempts to purify the complex via HIS-LSH failed as no DNMT3B or DNMT1 were detected (Figure 5.4A). However, I was able to successfully co-purify LSH and DNMT3B using HIS-DNMT3B indicating that these proteins form a purifiable complex when overexpressed in insect cells (Figure 5.4A). Surprisingly, DNMT1 was only detected at very low levels in this purification despite its previously reported interaction with DNMT3B (Figure 5.4A) (Kim et al., 2002). I used SDS-PAGE to monitor the purity of the purified complex compared to DNMT3B purified the same way without LSH (Figure 5.4B). The complex was impure with only a small quantity of LSH and several other contaminating proteins present in both samples (Figure 5.4B). The most likely reason for the relative impurity of the protein complex compared to the individually purified proteins was

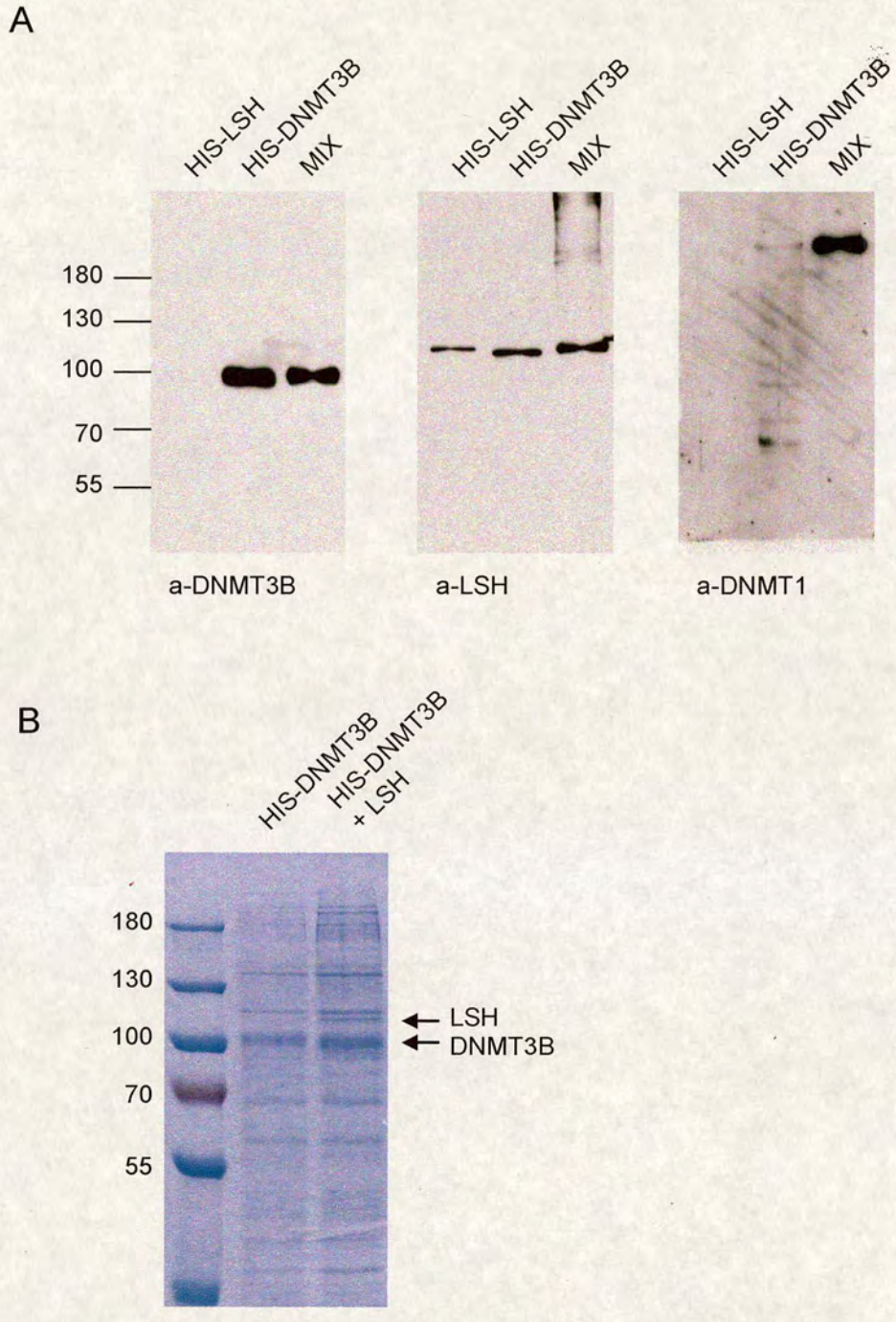


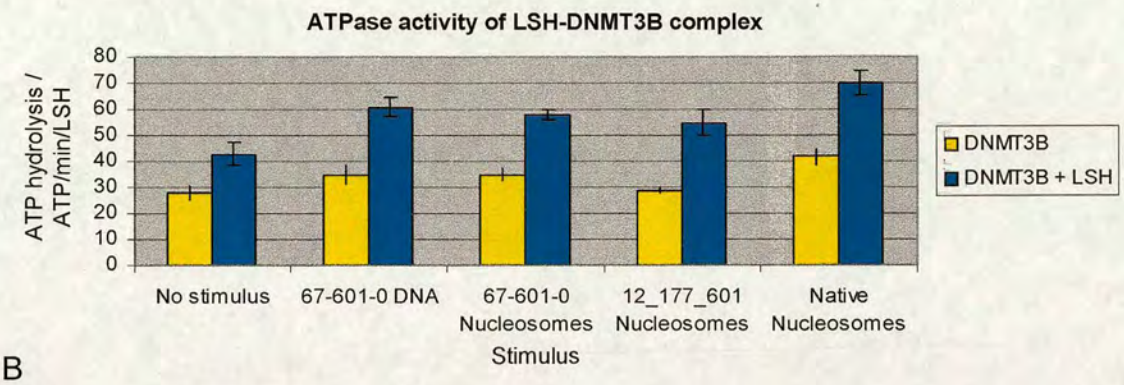
Figure 5.4 Co-purification of LSH and DNMT3B from insect cells
A. Sf9 insect cells were infected with baculovirus encoding recombinant HIS-LSH, DNMT3B and DNMT1 (labelled HIS-LSH) or HIS-DNMT3B, LSH and DNMT1 (labelled HIS-DNMT3B). The proteins were purified via Cobalt affinity chromatography, separated by SDS-PAGE and detected by the appropriate antibodies (indicated below panels). As a positive control, 100ng of each individually purified recombinant proteins were mixed and run alongside the purified complexes (labelled MIX).
B. Coomassie stained gel showing 4µl of co-purified HIS-DNMT3B + LSH and HIS-DNMT3B alone. Size markers in kDa listed to the left.

the low stringency purification protocol utilised in purification. I had previously used 1M NaCl for solubilisation and purification of DNMT3B. This high level of salt was not efficient for co-purification of LSH so lower salt (200mM) was used. I think it likely that the relatively low level of salt used for the co-purification and subsequent washing resulted in the high level of background proteins.

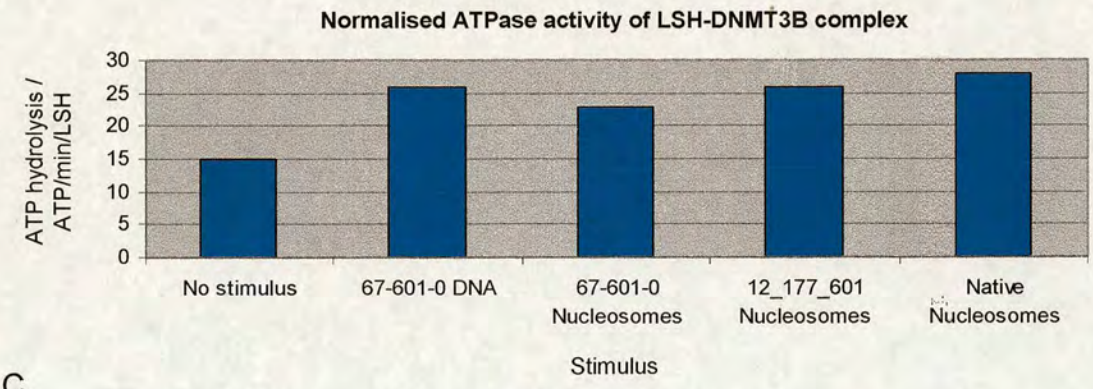
I performed ATPase assays with the partially purified LSH-DNMT3B complex using the partially purified DNMT3B as a control. The LSH-DNMT3B complex exhibited a low level of ATPase activity that was significantly higher than the background level exhibited by the DNMT3B control. As before this activity was weakly stimulated by DNA, nucleosomes with linker DNA and native nucleosomes (Figure 5.5A). I subtracted the background level of ATPase activity exhibited by DNMT3B to determine if LSH activity was higher when bound to DNMT3B. This revealed that neither the activity nor the level of stimulation were significantly higher than that of recombinant LSH alone (Figure 5.5B). Together, this indicates that DNMT3B does not stimulate LSH activity *in vitro*.

I also used the partially purified complex to assess the impact of the LSH interaction on DNMT3B activity. I found that, similar to recombinant DNMT3B the complex had very low activity *in vitro* (Figure 5.5C). Again, I found LSH had no stimulatory effect on DNMT3B activity on native nucleosomal templates (Figure 5.5C). Together, these results indicate that recombinant LSH is unable to affect the activity of DNMT3B *in vitro*.

A



B



C

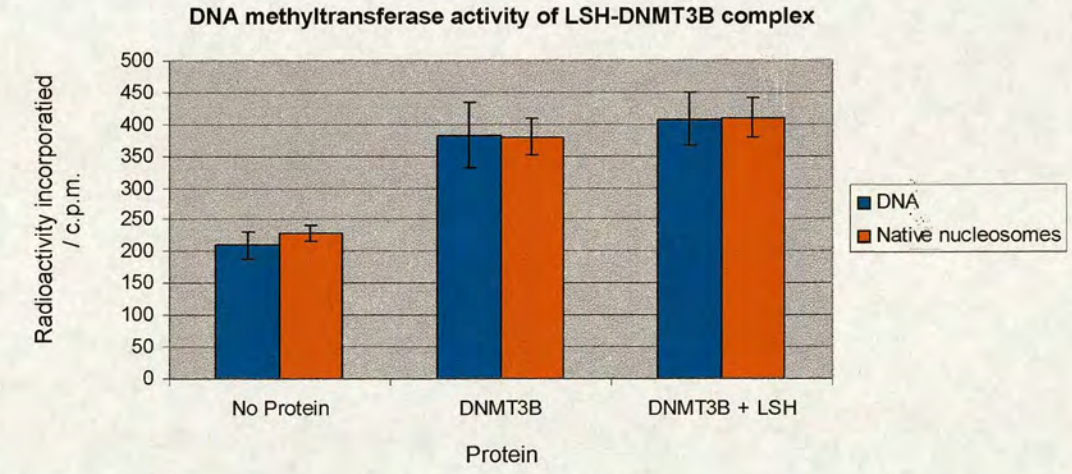


Figure 5.5 Analysis of the activity of LSH-DNMT3B complex

A. Co-purified LSH and DNMT3B or DNMT3B (25nM of each) were premixed and used in ATPase assays with 50μM ATP and 200nM of a range of stimuli (indicated on the X axis). The number of ATP molecules hydrolysed per minute per LSH was calculated by analysis of thin layer chromatography plates (Y axis). ATP hydrolysis rate observed for DNMT3B+LSH (blue bars) or DNMT3B (yellow bars) for each stimulus is displayed. Error bars denote standard deviation from the mean of three independent experiments.

B. Rate of ATP hydrolysis by the LSH-DNMT3B complex subtracted from that of DNMT3B background levels.
C. 50nM recombinant DNMT3B or LSH+DNMT3B complex was used in DNA methyltransferase assays on 200ng dGdC DNA or 100nM nucleosomes. Reactions including ^{3}H-SAM were incubated at 37°C for 1 hour, spotted on DE81 paper, washed extensively and radioactivity incorporation measured by scintillation counting. Error bars denote standard deviation from the mean of three independent experiments.

5.4 Summary

The genetic function of LSH as a crucial modulator of mammalian CpG methylation has been well established in studies of mice lacking in this protein (Dennis et al., 2001; Sun et al., 2004). The finding that cells missing LSH are deficient in *de novo* but not maintenance methylation of replicating episomal plasmids and interacts with Dnmt3a and Dnmt3b indicates LSH has a direct role in this process (Zhu et al., 2006). How the enzymatic activity of LSH relates to its function *in vivo* are currently unknown. As a member of the SNF2 family of chromatin remodelling enzymes it would be expected that its enzymatic activity is related to its *in vivo* function. SNF2 enzymes use ATP hydrolysis to disrupt chromatin structure and can alter the accessibility of nucleosomal DNA (Narlikar et al., 2002). Thus, it could be envisaged that LSH could increase the accessibility of DNA to DNMT enzymes. Experiments described in chapter 3 sought to determine the catalytic function of LSH. I determined that LSH is an active ATPase whose activity is stimulated by DNA. However, both the activity and the level of stimulation were relatively low compared to other SNF2 enzymes. Additionally, LSH did not exhibit the ability to slide nucleosomes using two distinct assays. These results led me to question whether the recombinant LSH I had used was fully active or if LSH perhaps required additional co factors for maximal activity. I therefore used a number of approaches to identify proteins that interact with LSH and could potentially modulate its activity. I used a reporter assay system to identify proteins required for LSH-mediated repression. This identified members of the DNA methylation and histone modification pathways as crucial for transcriptional silencing by LSH. Further experiments confirmed that LSH directly interacts with DNMT3B and indirectly with DNMT1, HDAC1 and HDAC2. This raised two important questions that I have attempted to answer in this chapter. How the repressive LSH complex is targeted *in vivo* and does this complex have increased ATPase and DNMTase activity compared to its individual components?

By comparing the relative DNA binding affinities of the DNA associated components of the LSH complex I determined that LSH is the most efficient binder.

It binds DNA with ~10 fold higher affinity than DNMT3B, whereas DNMT1 shows no detectable binding to DNA. Thus it would appear that LSH is the predominant DNA binding protein in this complex. However, the cellular situation is far more complex than this *in vitro* binding experiment and it is entirely possible that other members of this complex would bind DNA in the context of chromatin with greater efficiency. It will be interesting to test this binding *in vivo* as this will take into account other factors that may influence the tethering of this complex to genomic loci. . The tethering of DNMTs to chromatin *in vivo* is likely to be due to cumulative binding affinities of several proteins to both DNA and histones. A recent example of this is the proposed role of Dnmt3L in recruiting Dnmt3a to genomic loci lacking H3K4me (Ooi et al., 2007). In addition to this Dnmt3a has been shown to preferentially methylate CpG sites spaced 10bp apart *in vitro* (Jia et al., 2007). Thus it can be envisaged that targeting Dnmt3a activity *in vivo* may occur via a combination of Dnmt3L tethering it to unmodified H3K4 and its intrinsic substrate preference. Thus determining if and how LSH is involved in targeting and tethering DNMTs to chromatin would provide useful insight into its role in DNA methylation. Future work could focus on identifying *in vivo* targets of LSH using a ChIP-chip approach. As *in vivo* LSH targets are largely unknown it would be preferable to use an unbiased approach across a chromosome. Once identified, the presence of the rest of the LSH complex at the target sites could be attempted. If confirmed this would allow assaying of *in vivo* targeting of the complex by knocking down individual components of the complex and using ChIP to determine the crucial members required for targeting. These experiments were outside the scope of the current project but would provide interesting insights into the *in vivo* targeting of LSH.

In order to test the *in vitro* activity of the LSH complex I expressed and purified recombinant DNMT1 and DNMT3B from insect cells using a baculovirus system. I confirmed that both enzymes were catalytically active and attempted to reconstitute the LSH complex *in vitro*. Reconstitution of the complex was monitored by size exclusion chromatography but appeared to be unsuccessful. Also, the mixed proteins did not enhance each others activity. Reasoning it possible that the proteins may require being in a complex to influence each others activity, I co-expressed and co-

purified them from insect cells. Using relatively low salt extraction and washes I successfully co-purified LSH, but not DNMT1, via tagged DNMT3B. Although not particularly pure, this complex was used to test how DNMT3B affects LSH activity and vice versa. Disappointingly, I did not detect a significant change in the ATPase activity of LSH or the DNMTase activity of DNMT3B in this complex compared to monomeric protein. These results indicate that if LSH requires a co-factor for maximal activity it is not DNMT3B. I also failed to find a stimulatory role for LSH on DNMT3B activity. It is entirely possible that LSH, if fully active, would augment the activity of DNMT3B on DNA or nucleosomal templates. As LSH was bound to DNMT3B in these experiments but did not influence its activity, I think it is unlikely that it stimulates DNMT3B in a manner analogous to how Dnmt3L enhances the activity of Dnmt3a. In this example stimulation is believed to occur via stabilisation of the active site loop of Dnmt3a by Dnmt3L (Jia et al., 2007). Thus, Dnmt3L appears to be an inactive, protein co factor for Dnmt3a. Due to the complete lack of effect on DNMT3B activity by the relatively inactive LSH *in vitro* I think it unlikely that LSH provides a similar role. It would be possible to address this issue *in vivo* using a genetic approach. $Lsh^{-/-}$ MEFs have been obtained from the Muegge lab and will be used to generate cell lines stably expressing LSH and LSH^{K254Q} (Dennis et al., 2001). Investigating the dynamics of recovery of DNA methylation in these cell lines will establish whether the catalytic activity of LSH is required for DNA methylation *in vivo*. This will provide insight into whether LSH is simply providing a passive anchor to DNA for DNMTs or whether it is actively altering chromatin to increase its accessibility. The result of this experiment would indicate whether it is worthwhile to continue searching for potential modulators of LSH activity.

Chapter six - Discussion

Methylation of the carbon-5 position of cytosine is the major covalent modification of DNA found in eukaryotes. In mammals, this modification plays a number of diverse roles in development and cellular differentiation. These include transcriptional gene silencing, chromosomal integrity, repression of transposable elements, parental imprinting and X chromosome inactivation (Bird, 2002). Patterns of DNA methylation are established and maintained by proteins of the DNA methyltransferase (DNMT) family (Bestor, 2000a; Bird, 2002). Interestingly, genetic studies in *A. thaliana* and mouse have identified a protein from a functionally distinct family that is also required for high levels of DNA methylation. This protein, termed DDM1 in plants (deficient in DNA methylation 1) or Lsh in mouse (lymphoid specific helicase), belongs to the SNF2 chromatin remodelling family of ATPases (Dennis et al., 2001; Jeddeloh et al., 1999; Vongs et al., 1993). SNF2 enzymes have a variety of chromatin related functions that are based upon the ability to disrupt DNA:histone contacts in an ATP dependant manner (Narlikar et al., 2002). Despite the genetic characterisation of these proteins, little is known about their molecular function. A number of studies using Lsh deficient cells have indicated that it is primarily involved in *de novo* methylation of DNA (Zhu et al., 2006). Furthermore, a physical interaction between Lsh and Dnmt3a and Dnmt3b has been demonstrated, indicative of a direct role for Lsh in establishing genomic DNA methylation (Zhu et al., 2006). As a member of the SNF2 family of ATPases an attractive hypothesis is that Lsh may disrupt chromatin in a manner that makes it more accessible to DNMT enzymes. Alternatively, LSH may act as a recruitment factor for DNMTs or increase their catalytic activity in some other way. The aim of this study was to characterise LSH biochemically and attempt to relate this to its role in DNA methylation *in vivo*.

6.1 Determining the enzymatic function of LSH

To investigate the enzymatic function of LSH I purified recombinant protein from insect cells. I used this protein to demonstrate that LSH is a DNA binding protein with low levels of DNA stimulated ATPase activity. This data correlates well with that shown for DDM1 which is also a DNA stimulated ATPase (Brzeski and Jerzmanowski, 2003). However, the maximal velocity of ATP hydrolysis displayed by recombinant LSH is much lower than that recorded for other SNF2 enzymes. I also tested whether LSH is able to remodel chromatin using a nucleosome sliding assay and an SssI accessibility assay. I did not detect any remodelling activity by LSH in either assay indicating that recombinant LSH is not able to remodel chromatin. The inability of LSH to remodel chromatin does not correlate with the nucleosome repositioning activity demonstrated for DDM1 (Brzeski and Jerzmanowski, 2003) and confirmed by myself in this study. There are three possible explanations for this. The first is that LSH does not possess such an activity and is therefore functionally different from DDM1. The second is that LSH displays a type remodelling activity out with the detection limitations of these assays. Such an activity could involve transient disruptions of DNA:histone contacts that do not lead to nucleosome repositioning. This has been previously demonstrated for ATRX protein (Xue et al., 2003). As the SssI based assay was not particularly successful and is limited to detecting exposed sites containing CpGs it may not have detected remodeling events like this. The third possible explanation is that due to the inefficiency of ATP hydrolysis, very low levels of chromatin remodelling by LSH is not detectable in these assays. Obtaining fully active protein or identifying a co-factor required for activation may lead to LSH displaying chromatin remodelling activity.

The low level of ATPase activity exhibited by LSH and its inability to remodel chromatin suggest one of four possibilities – (1) LSH is an inefficient SNF2 ATPase. (2) The conditions are not optimised for optimal activity. (3) The recombinant protein is not as active as the native one. (4) A co-factor is missing. As the first two

issues are difficult to address I focused on dissecting if native LSH is more active than recombinant and whether LSH requires a protein co-factor. I purified Flag-LSH expressed in HeLa cells and assessed its ability to hydrolyse ATP. Surprisingly, this protein was completely inactive indicating that this method was not efficient at generating active LSH for biochemical studies. I was only able to purify very small quantities of protein by this method. Therefore, it is possible that the ATPase assay I used is not sensitive enough to detect the very low levels of ATPase activity exhibited by LSH. The purification of Flag-LSH involved stringent washes with a buffer containing 500mM NaCl. It is possible that the inability of this protein to hydrolyse ATP was due to a required protein co-factor being washed from it during purification. It is worth noting that most SNF2 enzymes reside in multi-subunit complexes and it has been reported that these subunits can modulate enzymatic activity (Langst et al., 1999). Thus, I next sought to determine whether LSH is associated with other proteins *in vivo* and if so, determine their functional significance.

6.2 LSH cooperates with DNMTs to repress transcription

My initial attempts to identify proteins that interact with LSH focused on biophysical analysis of the protein using gel filtration chromatography and sucrose gradients. These data strongly suggest that LSH is present predominantly as a free monomeric peptide in mammalian cells. The lack of detectable, stably bound interaction partners is unusual for SNF2 enzymes, and indicates that LSH most likely interacts transiently with the DNA methylation machinery. In addition to using protein chromatography to identify a stable LSH complex, I also performed a Y2H screen in search of transiently interacting proteins. This screen did not identify any interacting proteins, which I believe was due to LSH preventing activation of the Y2H reporter genes. Although this result complicated identification of LSH interacting partners it, allied with many studies on the transcriptional effects of SNF2 enzymes, led me to ask whether LSH can repress transcription in mammalian cells.

I used luciferase reporter assays to show that LSH can act as a transcriptional repressor in mammalian cells and used these assays as a tool to identify proteins that functionally interact with LSH. Using this system, I successfully identified a key role for histone deacetylase activity, DNMT1 and DNMT3B in LSH mediated repression. I used co-immunoprecipitation experiments to identify an interaction between LSH and HDAC1 and HDAC2 indicating that LSH is an HDAC dependant transcriptional repressor. Furthermore, in collaboration with Dr Irina Stancheva I was able to show using co-immunoprecipitation and *in vitro* pull down experiments that LSH physically interacts with DNMT3B and DNMT1 and both of these proteins are required for the interaction with HDACs. These experiments led us to propose a model for LSH mediated repression (Figure 4.13). In this model LSH recruits DNMT1, HDAC1 and HDAC2 indirectly, via a direct interaction with DNMT3B. Thus LSH directly binds at least one member of the DNA methylation machinery, strongly indicative of a direct role in the DNA methylation process. Interestingly, although repression by LSH required DNMT enzymes, it was not dependant on their catalytic function, and did not lead to DNA methylation of the targeted promoter in

the time scale measured. This is in agreement with previous studies which have shown that the repressive ability of DNMTs is independent of their role in DNA methylation (Fuks et al., 2001; Robertson et al., 2000; Rountree et al., 2000). Additionally, several studies have indicated that histone deacetylation and transcriptional silencing precede DNA methylation. For example, at the Oct4 locus, which becomes transcriptionally silenced during development, histone deacetylation is an early event indicate that that precedes DNA methylation *in vivo* (Deb-Rinker et al., 2005; Feldman et al., 2006; Jeong-Heon Lee, 2004). Also, during silencing of integrated transgenes and the tumour suppressor RASSF1A, histone deacetylation is an early event that comes before DNA methylation (Mutskov and Felsenfeld, 2004; Strunnikova et al., 2005). Thus, the HDAC dependent repressive function of LSH appears to be independent of its role in DNA methylation. Transcriptional repression caused by deacetylation of histone tails by LSH-interacting HDAC1/2 can be viewed as an initial and, perhaps reversible step in LSH-mediated gene silencing. A longer-term association of LSH with specific loci and a persistently high local concentration of DNMT1 and DNMT3B may result in methylation of CpGs at these loci. Whether long-term LSH association with specific regions does lead to DNA methylation could be tested by artificially targeting LSH to the promoter of a stably integrated reporter gene in HEK293 cells (Ishizuka and Lazar, 2003). This experiment could also help resolve whether the catalytic activity of LSH is required for its role in DNA methylation by comparing the dynamics of DNA methylation with the LSH^{K254Q} mutant. Importantly, accumulation of DNA methylation could be designated as an LSH, not transcriptional repression, mediated effect by using another transcriptional repressor, such as GAL4BD-MBD1 as a control. MBD1 has not been reported to interact with DNMTs so would determine whether DNA methylation, if it occurs, is specific to recruitment by LSH.

The experiments that identified LSH as an HDAC dependant transcriptional repressor utilised a targeted reporter system. Therefore, it is important to confirm these observations on endogenous targets to ensure that these findings are generally applicable. A number of endogenous targets of Lsh have been identified by chromatin immunoprecipitations. These include major and minor satellite repeats,

LINE and SINE elements, the differentially methylated region of the *cdkn1c* promoter and the promoters of several Hox genes (Fan et al., 2005b; Huang et al., 2004; Xi et al., 2007). Interestingly, these studies also provided evidence that these regions are hyperacetylated in $Lsh^{-/-}$ MEFs (Huang et al., 2004; Xi et al., 2007). Thus my data provides a mechanistic explanation of these findings by indicating that lack of histone acetylation is due to loss of HDAC1 and HDAC2 from these loci. An interesting future project could involve confirming that these loci are bound by LSH and identifying new binding sites using an unbiased chromatin immunoprecipitation approach. Crucially, $Lsh^{-/-}$ MEFs have been obtained and are available to use as a negative control in this experiment (Dennis et al., 2001). As mentioned above an unbiased approach should be used to identify Lsh targets, such as a ChIP-chip tiling array across a single chromosome or Solexa sequencing. This is because little is known about Lsh target sites *in vivo* so using promoter arrays or such like may miss crucial targets of the protein. Additionally, microarray analysis of RNA obtained from $Lsh^{-/-}$ MEFs did not detect deregulation of many genes indicating that promoter arrays would not be efficient at finding Lsh binding sites (Huang et al., 2004). Once identified, Lsh targets could then be used to test for recruitment of Dnmt1, Dnmt3b, Hdac1 and Hdac2. DNA methylation and histone acetylation levels at the identified loci could also be analysed. Following the identification of Lsh binding sites it would also be interesting to ask how this complex is targeted to chromatin. In the reporter assays the GAL4BD fused to LSH targeted it but the situation *in vivo* may be completely different. By knocking down the separate members of the LSH complex individually and performing chromatin immunoprecipitations against the other members this issue could be addressed.

6.3 Analysing the activity of the LSH-DNMT3B complex *in vitro*

The identification of proteins that interact with LSH *in vivo* encouraged me to test if they modulated the activity of LSH *in vitro*. As DNMT3B and LSH directly interact this was a particularly promising candidate for activating LSH. Mixing individually purified recombinant DNMT3B and DNMT1 with LSH did not lead to generation of a stable complex as analysed by gel filtration. Additionally, these proteins, when mixed together, did not affect the ATPase activity of LSH. As the individually purified proteins did not appear to form a complex when added together I was unable to rule out the possibility that they did not affect the ability of LSH to hydrolyse ATP as they did not reconstitute a stable complex. I attempted to overcome this problem by co-expressing and co-purifying the three proteins via HIS-tagged DNMT3B. This method was successful in co-purifying LSH but not DNMT1 although the purity and yield of the protein complex was low. Despite this, I analyzed whether the DNMT3B-LSH complex was more efficient in hydrolysing ATP than the monomeric LSH. Unfortunately, the complex did not display levels of ATPase activity that were significantly higher than the monomeric protein. These experiments led me to conclude that if recombinant LSH requires a co-factor to reach maximal rates of ATP hydrolysis it is not DNMT3B. The biochemical characterisation of LSH focused on the use of recombinant protein purified from insect cells. In the future it may be pertinent to attempt purification of native LSH from human cells. This could be attempted by conventional chromatography techniques as accomplished for human NuRD complex (Zhang et al., 1998). However, there are two potential problems that could be envisaged with such a project. The first is that I tested the activity of recombinant LSH produced and purified from human cells and found it to be indistinguishable from the controls. Thus, it is not immediately apparent that the production method of recombinant LSH is the reason for it being an inefficient ATPase. The second potential problem that can be foreseen is that LSH may require transient interactions with other protein co-factors for full activity. Biophysical analysis of LSH determined that the majority of the native protein is not part of a stable complex. Thus attempts to biochemically purify LSH with transiently

associated proteins would likely be inefficient. In this respect it would perhaps be more advisable to partially purify Lsh from mouse embryonic fibroblasts by immunoprecipitation using the $Lsh^{-/-}$ cells as a negative control. Thus, comparing the purified samples could distinguish between background and Lsh specific activity. Additionally, this technique could potentially be used to identify additional Lsh interacting partners and potential co-factors.

One possible explanation for the low level of ATPase activity exhibited by LSH is if it is a naturally inefficient enzyme. It is possible that it acts merely as a recruitment or stabilisation factor for DNMTs and is not required to alter chromatin structure in any way. This hypothesis was tested *in vitro* by assessing the ability of DNMT1 and DNMT3B to methylate DNA or nucleosomes with or without LSH. I reasoned that if the presumed chromatin remodelling activity of LSH is not required to enhance DNMT function the relatively inactive LSH should still have a positive effect on DNMT activity. Using mixed proteins that had been individually purified and the co-purified DNMT3B-LSH complex I did not detect any stimulatory effect of LSH on the ability of DNMTs to methylate DNA or nucleosomes. Of course, this result may be completely dependant on the ability of LSH to remodel chromatin and a highly active protein may significantly affect DNMT function. A key experiment in determining this will be to use LSH and LSH^{K254Q} to attempt to rescue DNA methylation patterns in $Lsh^{-/-}$ MEFs. By comparing the dynamics of DNA methylation recovery between these samples it will become clear whether the catalytic activity of LSH is required for its function *in vivo*. Initially, the recovery of DNA methylation at major and minor satellites will be analysed by Southern blot (Dennis et al., 2001). This could be complemented by looking at specific promoter regions that lose methylation in these cells. Such promoter could potentially be identified by methylated CpG affinity chromatography (Illingworth et al., 2008). These experiments are currently being attempted using retroviral constructs to stably infect $Lsh^{-/-}$ MEFs. Unfortunately, stable cell lines have not yet been generated and the timescale for this experiment is unknown. As *de novo* DNA methylation in somatic cells is relatively inefficient, extended incubation periods may be required to observe restoration of DNA methylation patterns. Additionally, it is not clear

whether a rescue experiment like this would be successful as the time window for establishing DNA methylation patterns, which occur in early embryogenesis, may have been passed (Bird, 2002). Despite these reservations, these experiments, if successful, would give excellent insight into whether the proposed catalytic activity of LSH is required for its role in DNA methylation.

A number of preliminary experiments investigating the relative binding efficiencies of different members of the LSH, DNMT3B and DNMT1 complex were described in chapter five. These showed that LSH binds DNA more efficiently than DNMT3B or DNMT1. The physiological relevance of these experiments is unclear at present. However, they may provide useful biochemical verification of experiments performed *in vivo*, particularly if the chromatin immunoprecipitation and knockdown experiments discussed above revealed that LSH tethers this complex to chromatin.

6.4 Concluding Remarks

The results obtained during my studies have provided an improved understanding of the molecular function of LSH. They have shown that recombinant LSH is a weak, DNA stimulated ATPase that does not remodel chromatin *in vitro*. Additionally, I have used a variety of techniques to identify an LSH complex consisting of DNMT3B, DNMT1, HDAC1 and HDAC2. This has established a mechanism by which LSH can repress transcription and perhaps give rise to a silenced chromatin state that is susceptible to stable silencing via DNA methylation. Disappointingly, I have not been able to detect a role for LSH in stimulating the activity of the DNA methylation machinery *in vitro*. It is possible that this is due to the low ATPase activity demonstrated by the recombinant protein. Future studies should attempt to resolve two key issues arising from this study. Firstly, is the LSH complex I identified found at endogenous LSH binding sites and if so how is it targeted? And secondly, is the proposed catalytic activity of LSH required for its function *in vivo*?

Appendix I

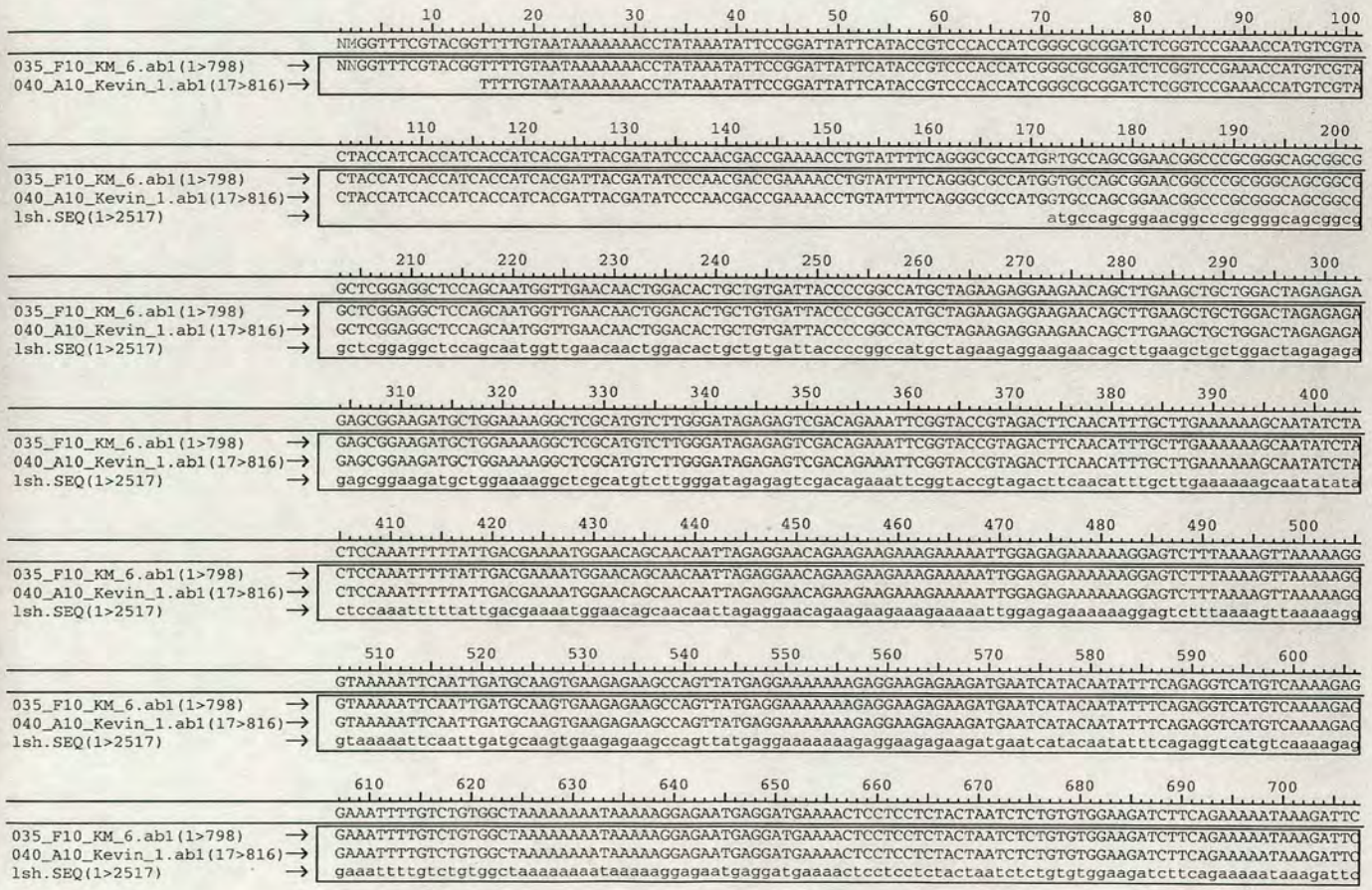
MMTV promoter DNA sequence (NucA)

CAAAAACTTATGGCATGAGTTATTATGAATAGCCTTTATTGGCCCAACCTTGCGGTT
CCCAGGGCTTAAGTAAGTTTTTGGTTACAAACTGTTCTTAAAACGAGGATGTGAGAC
AAGTGGTTTCCTGACTTGGTTTGGTATCAAAGGTTCTGATCTGAGCTCTGAGTGTTC
TATTTTCCTATGTTCTTTTGGAATTTATCCAAATCTTATGTAAATGCTTATGTAAAC
CAAGATATAAAAGAGTGCTGATTTTTTGAGTAAACTTGCAACAGTCCTAACATTCAC
CTCTTGTGTGTTTGTGTCTGTTCGCCATCCCGTCTCCGCTCGTCACTTA**T**CCTTCAC
TTTCCAGAGGGTCCCCCCGCAGACCCCGGCGACCCTCAGGTCGGCCGACTGCGGCAC
AGTTTTTTGCTCCTTTTTCTAGATGTAATTTTTAAAGCTTATTTTTTAACTTTCACA
TGTGCTACACTCACATGTGCAATGAGTGA

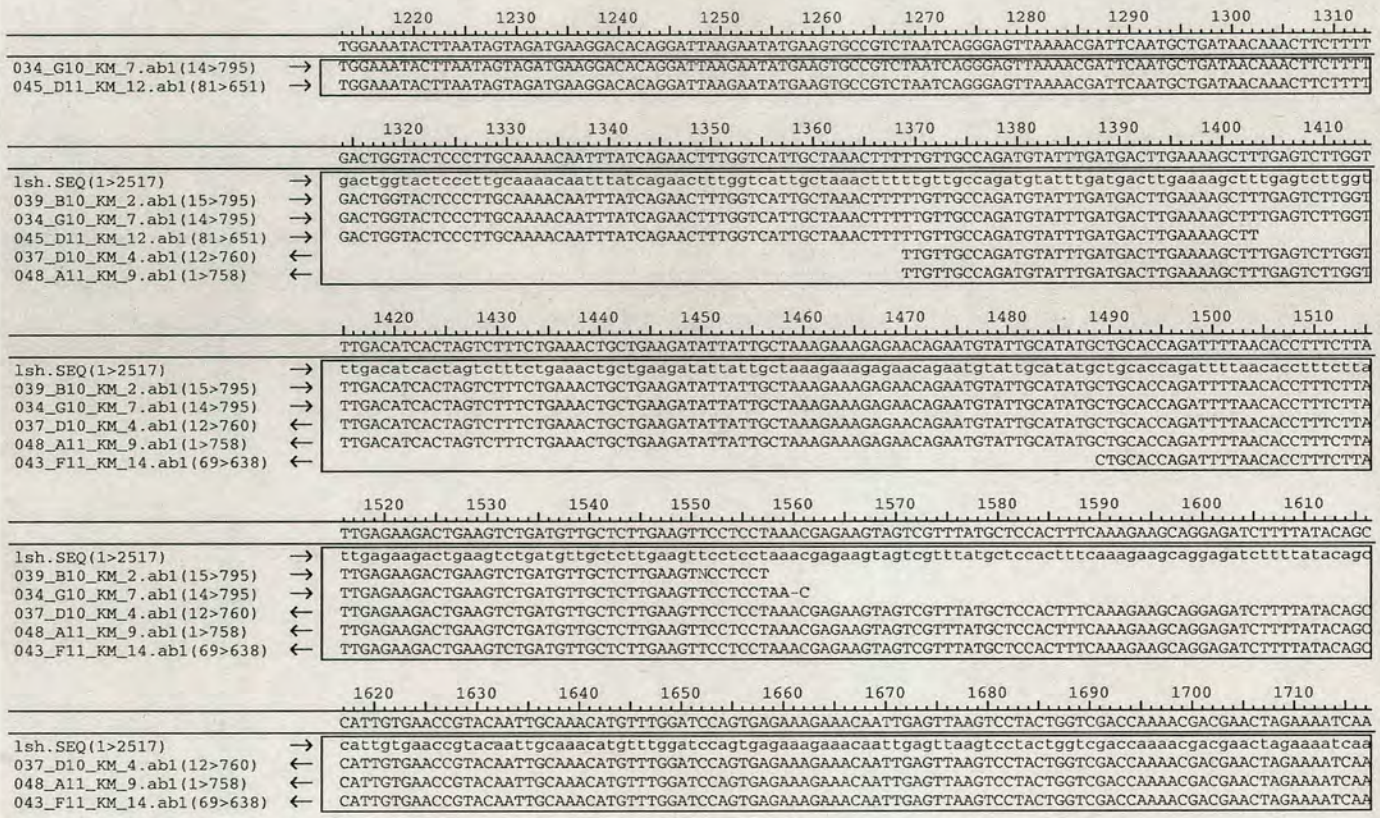
Underlined: NucA sequence
Bold: Central base in nucleosome

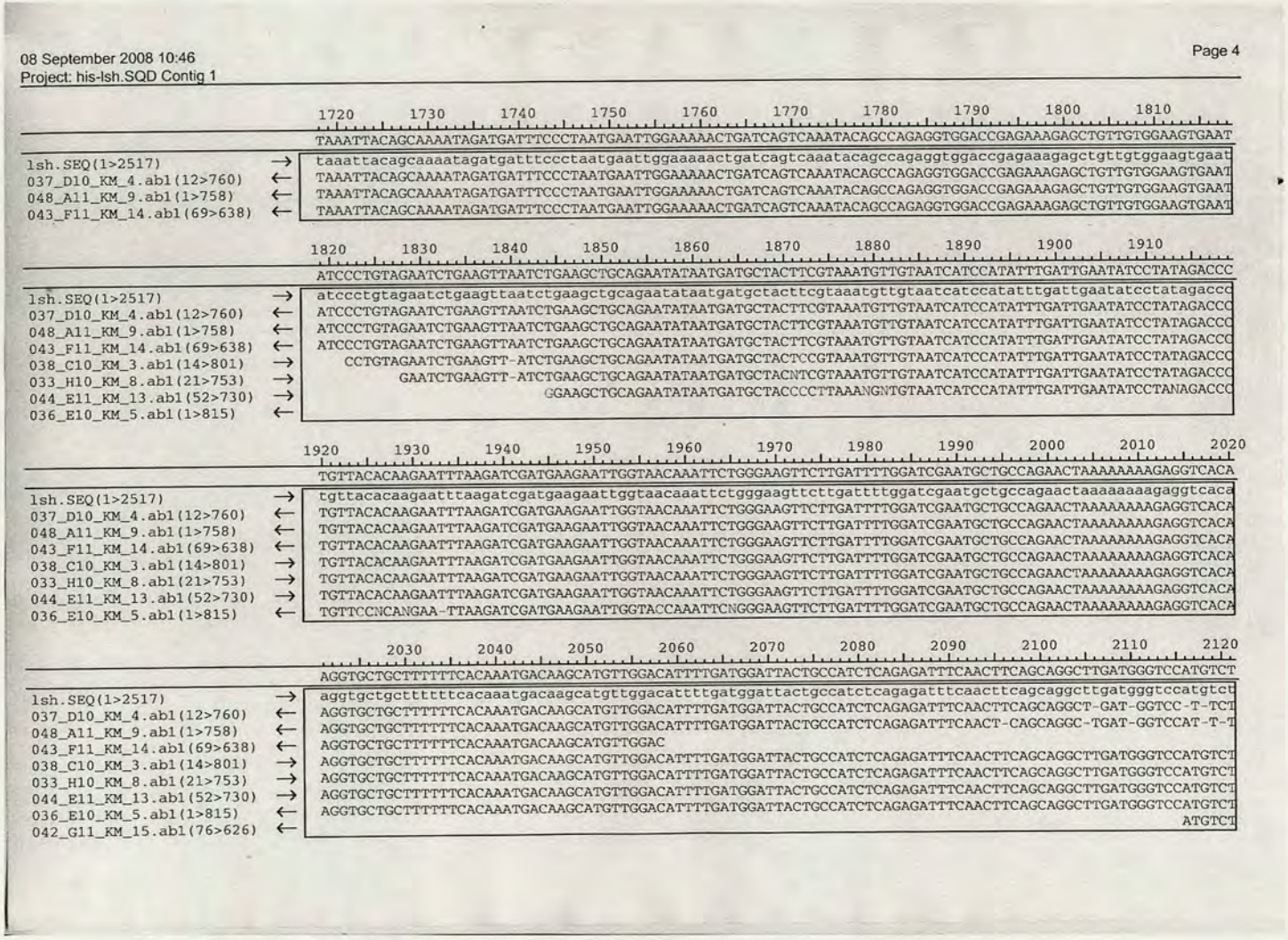
Appendix II

LSH cDNA sequenced following cloning into pFAST-BAC vector

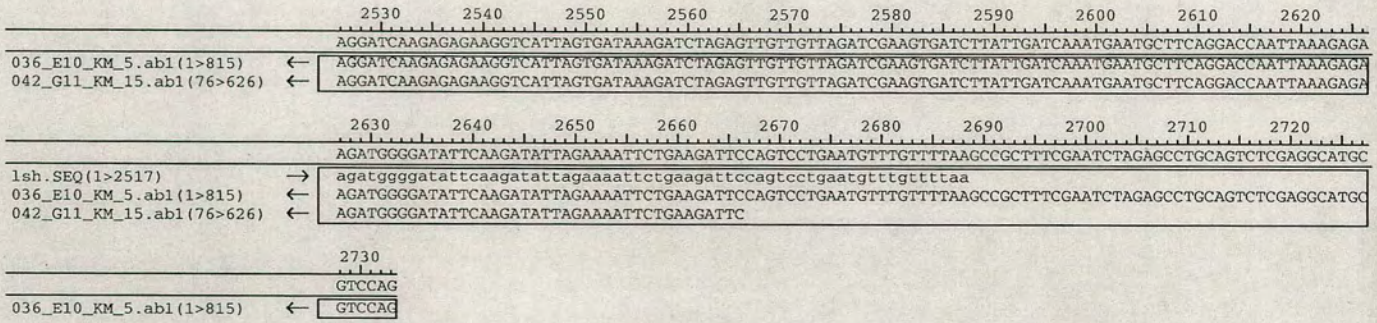

10 20 30 40 50 60 70 80 90 100
NNGGTTTCGTACGGTTTTGTAATAAAAAAACCTATAAATATTCCGGATTATTCATACCGTCCCACCATCGGGCGCGGATCTCGGTCCGAAACCATGTCGTA
035_F10_KM_6.ab1(1>798) →
040_A10_Kevin_1.ab1(17>816) →
110 120 130 140 150 160 170 180 190 200
CTACCATCACCATCACCATCACGATTACGATATCCCAACGACCGAAAACCTGTATTTTCAGGGCGCCATGRTGCCAGCGGAACGGCCCGCGGGCAGCGGCG
035_F10_KM_6.ab1(1>798) →
040_A10_Kevin_1.ab1(17>816) →
lsh.SEQ(1>2517) →
atgccagcggaacggcccgcgggcagcggcg
210 220 230 240 250 260 270 280 290 300
GCTCGGAGGCTCCAGCAATGGTTGAACAACTGGACACTGCTGTGATTACCCCGGCCATGCTAGAAGAGGAAGAACAGCTTGAAGCTGCTGGACTAGAGAGA
035_F10_KM_6.ab1(1>798) →
040_A10_Kevin_1.ab1(17>816) →
lsh.SEQ(1>2517) →
gctcggaggctccagcaatggttgaacaactggacactgctgtgattaccccggccatgctagaagaggaagaacagcttgaagctgctggactagagaga
310 320 330 340 350 360 370 380 390 400
GAGCGGAAGATGCTGGAAAAGGCTCGCATGTCTTGGGATAGAGAGTCGACAGAAATTCGGTACCGTAGACTTCAACATTTGCTTGAAAAAAGCAATATCTA
035_F10_KM_6.ab1(1>798) →
040_A10_Kevin_1.ab1(17>816) →
lsh.SEQ(1>2517) →
gagcggaagatgctggaaaaggctcgcatgtcttgggatagagagtcgacagaaattcggtaccgtagacttcaacatttgcttgaaaaaagcaatatata
410 420 430 440 450 460 470 480 490 500
CTCCAAATTTTTATTGACGAAAATGGAACAGCAACAATTAGAGGAACAGAAGAAGAAAGAAAAATTGGAGAGAAAAAAGGAGTCTTTAAAAGTTAAAAAGG
035_F10_KM_6.ab1(1>798) →
040_A10_Kevin_1.ab1(17>816) →
lsh.SEQ(1>2517) →
ctccaaatttttattgacgaaaatggaacagcaacaattagaggaacagaagaagaaagaaaaattggagagaaaaaaggagtctttaaaagttaaaaagg
510 520 530 540 550 560 570 580 590 600
GTAAAAATTCAATTGATGCAAGTGAAGAGAAGCCAGTTATGAGGAAAAAAAGAGGAAGAGAAGATGAATCATACAATATTTCAGAGGTCATGTCAAAAGAG
035_F10_KM_6.ab1(1>798) →
040_A10_Kevin_1.ab1(17>816) →
lsh.SEQ(1>2517) →
gtaaaaattcaattgatgcaagtgaagagaagccagttatgaggaaaaaaagaggaagagaagatgaatcatacaatatttcagaggtcatgtcaaaagag
610 620 630 640 650 660 670 680 690 700
GAAATTTTGTCTGTGGCTAAAAAAAATAAAAAGGAGAATGAGGATGAAAACTCCTCCTCTACTAATCTCTGTGTGGAAGATCTTCAGAAAAATAAAGATTC
035_F10_KM_6.ab1(1>798) →
040_A10_Kevin_1.ab1(17>816) →
lsh.SEQ(1>2517) →
gaaattttgtctgtggctaaaaaaaataaaaaggagaatgaggatgaaaactcctcctctactaatctctgtgtggaagatcttcagaaaaataaagattc


710 720 730 740 750 760 770 780 790 800
GAATAGTATAATTAAaGATAGATTGTCTGAAACGGTTAGGCAGAATACTAAATTCTTTTTTGACCCAGTCCGGAaGTGTAATGGTCAGCCAGTACCTTTTC
035_F10_KM_6.ab1(1>798)
040_A10_Kevin_1.ab1(17>816)
lsh.SEQ(1>2517)
039_B10_KM_2.ab1(15>795)
034_G10_KM_7.ab1(14>795)
GAATAGTATAATTAA-GATAGATTGTCTGAAACGGTTAGGCAGAATACTAAATTCTTTTTTGACCCAGTCCGGAAGTGTAATGGTCAGCCAG
GAATAGTATAATTAAAGATAGATTGTCTGAAACGGTTAGGCAGAATACTAAATTCTTTTTTGACCCAGTCCGGAAGTGTAATGGTCAGCCAGTACCTTTTC
gaatagtataattaaagatagattgtctgaaacggttaggcagaatactaaattcttttttgacccagtccggaagtgtaatggtcagccagtacctttt
GTCCGGA-GTGTAATGGTCAGCCAGTACCTTTTC
N-CGGA-GTGTAATGGTCAGCCAGTACCTTTTC
810 820 830 840 850 860 870 880 890 900
AACAACCAAAGCACTTCACTGGAGGAGTGATGCGATGGTACCAAGTAGAAGGCATGGAATGGCTTAGGATGCTTTGGGAAAATGGAATTAATGGCATTTTA
040_A10_Kevin_1.ab1(17>816)
lsh.SEQ(1>2517)
039_B10_KM_2.ab1(15>795)
034_G10_KM_7.ab1(14>795)
045_D11_KM_12.ab1(81>651)
-ACAACN
aacaaccaaagcacttcactggaggagtgatgcgatggtaccaagtagaaggcatggaatggcttaggatgctttgggaaaatggaattaatggcatttta
AACAACCAAAGCACTTCACTGGAGGAGTGATGCGATGGTACCAAGTAGAAGGCATGGAATGGCTTAGGATGCTTTGGGAAAATGGAATTAATGGCATTTTA
AACAACCAAAGCACTTCACTGGAGGAGTGATGCGATGGTACCAAGTAGAAGGCATGGAATGGCTTAGGATGCTTTGGGAAAATGGAATTAATGGCATTTTA
GAGTGATGCGATGGTACCAAGTAGAAGGCATGGAATGGCTTAGGATGCTTTGGGAAAATGGAATTAATGGCATTTTA
910 920 930 940 950 960 970 980 990 1000 1010
GCAGATGAAATGGGATTGGGTAAGACAGTTCAGTGCATTGCTACTATTGCATTGATGATTCAGAGAGGAGTACCAGGACCTTTTCTTGTCTGTGGCCCTTT
lsh.SEQ(1>2517)
039_B10_KM_2.ab1(15>795)
034_G10_KM_7.ab1(14>795)
045_D11_KM_12.ab1(81>651)
gcagatgaaatgggattgggtaagacagttcagtgcattgctactattgcattgatgattcagagaggagtaccaggacctttttcttgtctgtggccctt
GCAGATGAAATGGGATTGGGTAAGACAGTTCAGTGCATTGCTACTATTGCATTGATGATTCAGAGAGGAGTACCAGGACCTTTTCTTGTCTGTGGCCCTTT
GCAGATGAAATGGGATTGGGTAAGACAGTTCAGTGCATTGCTACTATTGCATTGATGATTCAGAGAGGAGTACCAGGACCTTTTCTTGTCTGTGGCCCTTT
GCAGATGAAATGGGATTGGGTAAGACAGTTCAGTGCATTGCTACTATTGCATTGATGATTCAGAGAGGAGTACCAGGACCTTTTCTTGTCTGTGGCCCTTT
1020 1030 1040 1050 1060 1070 1080 1090 1100 1110
GTCTACACTTCCTAACTGGATGGCTGAATTCAAAAGATTTACACCAGATATCCCTACAATGTTATATCATGGAACCCAGGAGGAACGTCAAAAATTGGTAA
lsh.SEQ(1>2517)
039_B10_KM_2.ab1(15>795)
034_G10_KM_7.ab1(14>795)
045_D11_KM_12.ab1(81>651)
gtctacacttcctaactggatggctgaattcaaaagatttacaccagatatccctacaatgttatatcatggaacccaggaggaacgtcaaaaattggtaa
GTCTACACTTCCTAACTGGATGGCTGAATTCAAAAGATTTACACCAGATATCCCTACAATGTTATATCATGGAACCCAGGAGGAACGTCAAAAATTGGTAA
GTCTACACTTCCTAACTGGATGGCTGAATTCAAAAGATTTACACCAGATATCCCTACAATGTTATATCATGGAACCCAGGAGGAACGTCAAAAATTGGTAA
GTCTACACTTCCTAACTGGATGGCTGAATTCAAAAGATTTACACCAGATATCCCTACAATGTTATATCATGGAACCCAGGAGGAACGTCAAAAATTGGTAA
1120 1130 1140 1150 1160 1170 1180 1190 1200 1210
GAAATATTTACAAACGGAAAGGGACTTTGCAGATTCATCCTGTGGTAATCACGTCATTTGAAATAGCCATGAGAGACCGAAATGCGTTACAGCATTGCTAT
lsh.SEQ(1>2517)
039_B10_KM_2.ab1(15>795)
034_G10_KM_7.ab1(14>795)
045_D11_KM_12.ab1(81>651)
gaaatatttacaaacggaaagggactttgcagattcatcctgtggtaatcacgtcatttgaaatagccatgagagaccgaaatgcgttacagcattgctat
GAAATATTTACAAACGGAAAGGGACTTTGCAGATTCATCCTGTGGTAATCACGTCATTTGAAATAGCCATGAGAGACCGAAATGCGTTACAGCATTGCTAT
GAAATATTTACAAACGGAAAGGGACTTTGCAGATTCATCCTGTGGTAATCACGTCATTTGAAATAGCCATGAGAGACCGAAATGCGTTACAGCATTGCTAT
GAAATATTTACAAACGGAAAGGGACTTTGCAGATTCATCCTGTGGTAATCACGTCATTTGAAATAGCCATGAGAGACCGAAATGCGTTACAGCATTGCTAT
1220 1230 1240 1250 1260 1270 1280 1290 1300 1310
TGGAAATACTTAATAGTAGATGAAGGACACAGGATTAAGAATATGAAGTGCCGTCTAATCAGGGAGTTAAAACGATTCAATGCTGATAACAAACTTCTTTT
lsh.SEQ(1>2517)
039_B10_KM_2.ab1(15>795)
tggaaatacttaatagtagatgaaggacacaggattaagaatatgaagtgccgtctaatcagggagttaaaacgattcaatgctgataacaaacttcttttt
TGGAAATACTTAATAGTAGATGAAGGACACAGGATTAAGAATATGAAGTGCCGTCTAATCAGGGAGTTAAAACGATTCAATGCTGATAACAAACTTCTTTT


1220 1230 1240 1250 1260 1270 1280 1290 1300 1310
TGGAAATACTTAATAGTAGATGAAGGACACAGGATTAAGAATATGAAGTGCCGTCTAATCAGGGAGTTAAAACGATTCAATGCTGATAACAAACTTCTTTT
034_G10_KM_7.ab1(14>795) →
045_D11_KM_12.ab1(81>651) →
1320 1330 1340 1350 1360 1370 1380 1390 1400 1410
GACTGGTACTCCCTTGCAAAACAATTTATCAGAACTTTGGTCATTGCTAAACTTTTTGTTGCCAGATGTATTTGATGACTTGAAAAGCTTTGAGTCTTGGT
lsh.SEQ(1>2517) →
039_B10_KM_2.ab1(15>795) →
034_G10_KM_7.ab1(14>795) →
045_D11_KM_12.ab1(81>651) →
037_D10_KM_4.ab1(12>760) ←
048_A11_KM_9.ab1(1>758) ←
1420 1430 1440 1450 1460 1470 1480 1490 1500 1510
TTGACATCACTAGTCTTTCTGAAACTGCTGAAGATATTATTGCTAAAGAAAGAGAACAGAATGTATTGCATATGCTGCACCAGATTTTAACACCTTTCTTA
lsh.SEQ(1>2517) →
039_B10_KM_2.ab1(15>795) →
034_G10_KM_7.ab1(14>795) →
037_D10_KM_4.ab1(12>760) ←
048_A11_KM_9.ab1(1>758) ←
043_F11_KM_14.ab1(69>638) ←
1520 1530 1540 1550 1560 1570 1580 1590 1600 1610
TTGAGAAGACTGAAGTCTGATGTTGCTCTTGAAGTTCCTCCTAAACGAGAAGTAGTCGTTTATGCTCCACTTTCAAAGAAGCAGGAGATCTTTTATACAGC
lsh.SEQ(1>2517) →
039_B10_KM_2.ab1(15>795) →
034_G10_KM_7.ab1(14>795) →
037_D10_KM_4.ab1(12>760) ←
048_A11_KM_9.ab1(1>758) ←
043_F11_KM_14.ab1(69>638) ←
1620 1630 1640 1650 1660 1670 1680 1690 1700 1710
CATTGTGAACCGTACAATTGCAAACATGTTTGGATCCAGTGAGAAAGAAACAATTGAGTTAAGTCCTACTGGTCGACCAAAACGACGAACTAGAAAATCAA
lsh.SEQ(1>2517) →
037_D10_KM_4.ab1(12>760) ←
048_A11_KM_9.ab1(1>758) ←
043_F11_KM_14.ab1(69>638) ←


1720 1730 1740 1750 1760 1770 1780 1790 1800 1810
TAAATTACAGCAAAATAGATGATTTCCCTAATGAATTGGAAAAACTGATCAGTCAAATACAGCCAGAGGTGGACCGAGAAAGAGCTGTTGTGGAAGTGAAT
lsh.SEQ(1>2517) → taaattacagcaaaatagatgatttccctaatgaattggaaaaactgatcagtcaaatacagccagaggtggaccgagaaagagctgttgtggaagtgaat
037_D10_KM_4.ab1(12>760) ← TAAATTACAGCAAAATAGATGATTTCCCTAATGAATTGGAAAAACTGATCAGTCAAATACAGCCAGAGGTGGACCGAGAAAGAGCTGTTGTGGAAGTGAAT
048_A11_KM_9.ab1(1>758) ← TAAATTACAGCAAAATAGATGATTTCCCTAATGAATTGGAAAAACTGATCAGTCAAATACAGCCAGAGGTGGACCGAGAAAGAGCTGTTGTGGAAGTGAAT
043_F11_KM_14.ab1(69>638) ← TAAATTACAGCAAAATAGATGATTTCCCTAATGAATTGGAAAAACTGATCAGTCAAATACAGCCAGAGGTGGACCGAGAAAGAGCTGTTGTGGAAGTGAAT
1820 1830 1840 1850 1860 1870 1880 1890 1900 1910
ATCCCTGTAGAATCTGAAGTTAATCTGAAGCTGCAGAATATAATGATGCTACTTCGTAAATGTTGTAATCATCCATATTTGATTGAATATCCTATAGACCC
lsh.SEQ(1>2517) → atccctgtagaatctgaagttaatctgaagctgcagaatataatgatgctacttcgtaaatgttgtaatcatccatatttgattgaatatcctatagaccc
037_D10_KM_4.ab1(12>760) ← ATCCCTGTAGAATCTGAAGTTAATCTGAAGCTGCAGAATATAATGATGCTACTTCGTAAATGTTGTAATCATCCATATTTGATTGAATATCCTATAGACCC
048_A11_KM_9.ab1(1>758) ← ATCCCTGTAGAATCTGAAGTTAATCTGAAGCTGCAGAATATAATGATGCTACTTCGTAAATGTTGTAATCATCCATATTTGATTGAATATCCTATAGACCC
043_F11_KM_14.ab1(69>638) ← ATCCCTGTAGAATCTGAAGTTAATCTGAAGCTGCAGAATATAATGATGCTACTTCGTAAATGTTGTAATCATCCATATTTGATTGAATATCCTATAGACCC
038_C10_KM_3.ab1(14>801) → CCTGTAGAATCTGAAGTT-ATCTGAAGCTGCAGAATATAATGATGCTACTCCGTAAATGTTGTAATCATCCATATTTGATTGAATATCCTATAGACCC
033_H10_KM_8.ab1(21>753) → GAATCTGAAGTT-ATCTGAAGCTGCAGAATATAATGATGCTACNTCGTAAATGTTGTAATCATCCATATTTGATTGAATATCCTATAGACCC
044_E11_KM_13.ab1(52>730) → GGAAGCTGCAGAATATAATGATGCTACCCCTTAAANGNTGTAATCATCCATATTTGATTGAATATCCTANAGACCC
036_E10_KM_5.ab1(1>815) ←
1920 1930 1940 1950 1960 1970 1980 1990 2000 2010 2020
TGTTACACAAGAATTTAAGATCGATGAAGAATTGGTAACAAATTCTGGGAAGTTCTTGATTTTGGATCGAATGCTGCCAGAACTAAAAAAAAGAGGTCACA
lsh.SEQ(1>2517) → tgttacacaagaatttaagatcgatgaagaattggtaacaaattctgggaagttcttgattttggatcgaatgctgccagaactaaaaaaaagaggtcaca
037_D10_KM_4.ab1(12>760) ← TGTTACACAAGAATTTAAGATCGATGAAGAATTGGTAACAAATTCTGGGAAGTTCTTGATTTTGGATCGAATGCTGCCAGAACTAAAAAAAAGAGGTCACA
048_A11_KM_9.ab1(1>758) ← TGTTACACAAGAATTTAAGATCGATGAAGAATTGGTAACAAATTCTGGGAAGTTCTTGATTTTGGATCGAATGCTGCCAGAACTAAAAAAAAGAGGTCACA
043_F11_KM_14.ab1(69>638) ← TGTTACACAAGAATTTAAGATCGATGAAGAATTGGTAACAAATTCTGGGAAGTTCTTGATTTTGGATCGAATGCTGCCAGAACTAAAAAAAAGAGGTCACA
038_C10_KM_3.ab1(14>801) → TGTTACACAAGAATTTAAGATCGATGAAGAATTGGTAACAAATTCTGGGAAGTTCTTGATTTTGGATCGAATGCTGCCAGAACTAAAAAAAAGAGGTCACA
033_H10_KM_8.ab1(21>753) → TGTTACACAAGAATTTAAGATCGATGAAGAATTGGTAACAAATTCTGGGAAGTTCTTGATTTTGGATCGAATGCTGCCAGAACTAAAAAAAAGAGGTCACA
044_E11_KM_13.ab1(52>730) → TGTTACACAAGAATTTAAGATCGATGAAGAATTGGTAACAAATTCTGGGAAGTTCTTGATTTTGGATCGAATGCTGCCAGAACTAAAAAAAAGAGGTCACA
036_E10_KM_5.ab1(1>815) ← TGTTCCNCANGAA-TTAAGATCGATGAAGAATTGGTACCAAATTCNGGGAAGTTCTTGATTTTGGATCGAATGCTGCCAGAACTAAAAAAAAGAGGTCACA
2030 2040 2050 2060 2070 2080 2090 2100 2110 2120
AGGTGCTGCTTTTTTCACAAATGACAAGCATGTTGGACATTTTGATGGATTACTGCCATCTCAGAGATTTCAACTTCAGCAGGCTTGATGGGTCCATGTCT
lsh.SEQ(1>2517) → aggtgctgctttttttcacaaatgacaagcatgttggacattttgatggattactgccatctcagagatttcaacttcagcaggcttgatgggtccatgtct
037_D10_KM_4.ab1(12>760) ← AGGTGCTGCTTTTTTCACAAATGACAAGCATGTTGGACATTTTGATGGATTACTGCCATCTCAGAGATTTCAACTTCAGCAGGCT-GAT-GGTCC-T-TCT
048_A11_KM_9.ab1(1>758) ← AGGTGCTGCTTTTTTCACAAATGACAAGCATGTTGGACATTTTGATGGATTACTGCCATCTCAGAGATTTCAACT-CAGCAGGC-TGAT-GGTCCAT-T-T
043_F11_KM_14.ab1(69>638) ← AGGTGCTGCTTTTTTCACAAATGACAAGCATGTTGGAC
038_C10_KM_3.ab1(14>801) → AGGTGCTGCTTTTTTCACAAATGACAAGCATGTTGGACATTTTGATGGATTACTGCCATCTCAGAGATTTCAACTTCAGCAGGCTTGATGGGTCCATGTCT
033_H10_KM_8.ab1(21>753) → AGGTGCTGCTTTTTTCACAAATGACAAGCATGTTGGACATTTTGATGGATTACTGCCATCTCAGAGATTTCAACTTCAGCAGGCTTGATGGGTCCATGTCT
044_E11_KM_13.ab1(52>730) → AGGTGCTGCTTTTTTCACAAATGACAAGCATGTTGGACATTTTGATGGATTACTGCCATCTCAGAGATTTCAACTTCAGCAGGCTTGATGGGTCCATGTCT
036_E10_KM_5.ab1(1>815) ← AGGTGCTGCTTTTTTCACAAATGACAAGCATGTTGGACATTTTGATGGATTACTGCCATCTCAGAGATTTCAACTTCAGCAGGCTTGATGGGTCCATGTCT
042_G11_KM_15.ab1(76>626) ← ATGTCT


2130 2140 2150 2160 2170 2180 2190 2200 2210 2220
TACTCAGAGAGAGAAAAAAACATGCACAGCTTCAACACGGATCCAGAGGTGTTTATCTTCTTAGTGAGTACACGAGCTGGTGGCCTGGGCATTAATCTGAC
lsh.SEQ(1>2517) → tactcagagagagaaaaaaacatgcacagcttcaacacggatccagaggtgtttatcttcttagtgagtacacgagctggtggcctgggcattaatctgac
048_A11_KM_9.ab1(1>758) ← T--TCGG-GA-ANT
038_C10_KM_3.ab1(14>801) → TACTCAGAGAGAGAAAAAAACATGCACAGCTTCAACACGGATCCAGAGGTGTTTATCTTCTTAGTGAGTACACGAGCTGGTGGCCTGGGCATTAATCTGAC
033_H10_KM_8.ab1(21>753) → TACTCAGAGAGAGAAAAAAACATGCACAGCTTCAACACGGATCCAGAGGTGTTTATCTTCTTAGTGAGTACACGAGCTGGTGGCCTGGGCATTAATCTGAC
044_E11_KM_13.ab1(52>730) → TACTCAGAGAGAGAAAAAAACATGCACAGCTTCAACACGGATCCAGAGGTGTTTATCTTCTTAGTGAGTACACGAGCTGGTGGCCTGGGCATTAATCTGAC
036_E10_KM_5.ab1(1>815) ← TACTCAGAGAGAGAAAAAAACATGCACAGCTTCAACACGGATCCAGAGGTGTTTATCTTCTTAGTGAGTACACGAGCTGGTGGCCTGGGCATTAATCTGAC
042_G11_KM_15.ab1(76>626) ← TACTCAGAGAGAGAAAAAAACATGCACAGCTTCAACACGGATCCAGAGGTGTTTATCTTCTTAGTGAGTACACGAGCTGGTGGCCTGGGCATTAATCTGAC
2230 2240 2250 2260 2270 2280 2290 2300 2310 2320
TGCAGCAGATACAGTTATCATTTATGATAGTGATTGGAACCCCCAGTCGGATCTTCAGGCCCAGGATAGATGTCATAGAATTGGTCAGACAAAGCCAGTTG
lsh.SEQ(1>2517) → tgcagcagatacagttatcatttatgatagtgattggaacccccagtcggatcttcaggcccaggatagatgtcatagaattggtcagacaaagccagttg
038_C10_KM_3.ab1(14>801) → TGCAGCAGATACAGTTATCATTTATGATAGTGATTGGAACCCCCAGTCGGATCTTCAGGCCCAGGATAGATGTCATAGAATTGGTCAGACAAAGCCAGTTG
033_H10_KM_8.ab1(21>753) → TGCAGCAGATACAGTTATCATTTATGATAGTGATTGGAACCCCCAGTCGGATCTTCAGGCCCAGGATAGATGTCATAGAATTGGTCAGACAAAGCCAGTTG
044_E11_KM_13.ab1(52>730) → TGCAGCAGATACAGTTATCATTTATGATAGTGATTGGAACCCCCAGTCGGATCTTCAGGCCCAGGATAGATGTCATAGAATTGGTCAGACAAAGCCAGTTG
036_E10_KM_5.ab1(1>815) ← TGCAGCAGATACAGTTATCATTTATGATAGTGATTGGAACCCCCAGTCGGATCTTCAGGCCCAGGATAGATGTCATAGAATTGGTCAGACAAAGCCAGTTG
042_G11_KM_15.ab1(76>626) ← TGCAGCAGATACAGTTATCATTTATGATAGTGATTGGAACCCCCAGTCGGATCTTCAGGCCCAGGATAGATGTCATAGAATTGGTCAGACAAAGCCAGTTG
2330 2340 2350 2360 2370 2380 2390 2400 2410 2420
TTGTTTATCGCCTTGTTACAGCAAATACTATCGATCAGAAAATTGTGGAAAGAGCAGCTGCTAAAAGGAAACTGGAAAAGTTGATCATCCATAAAAATCAT
lsh.SEQ(1>2517) → ttgtttatcgccttgttacagcaaatactatcgatcagaaaattgtggaaagagcagctgctaaaaggaaactggaaaagttgatcatccataaaaatcat
038_C10_KM_3.ab1(14>801) → TTGTTTATCGCCTTGTTACAGCAAATACTATCGATCAGAAAATTGTGGAAAGAGCAGCTGCTAAAAGGAAACTGGAAAAGTTGATCATCCATAAAAATCAT
033_H10_KM_8.ab1(21>753) → TTGTTTATCGCCTTGTTACAGCAAATACTATCGATCAGAAAATTGTGGAAAGAGCAGCTGCTAAAAGGAAACTGGAAAAGTTGATCATCCATAAAAATCAT
044_E11_KM_13.ab1(52>730) → TTGTTTATCGCCTTGTTACAGCAAATACTATCGATCAGAAAATTGTGGAAAGAGCAGCTGCTAAAAGGAAACTGGAAAAGTTGATCATCCATAAAAATCAT
036_E10_KM_5.ab1(1>815) ← TTGTTTATCGCCTTGTTACAGCAAATACTATCGATCAGAAAATTGTGGAAAGAGCAGCTGCTAAAAGGAAACTGGAAAAGTTGATCATCCATAAAAATCAT
042_G11_KM_15.ab1(76>626) ← TTGTTTATCGCCTTGTTACAGCAAATACTATCGATCAGAAAATTGTGGAAAGAGCAGCTGCTAAAAGGAAACTGGAAAAGTTGATCATCCATAAAAATCAT
2430 2440 2450 2460 2470 2480 2490 2500 2510 2520
TTCAAAGGTGGTCAGTCTGGATTAAATCTGTCTAAGAATTTCTTAGATCCTAAGGAATTAATGGAATTATTAAAATCTAGAGATTATGAAAGGGAAATAAA
lsh.SEQ(1>2517) → ttcaaaggtggtcagtctggattaaatctgtctaagaatttcttagatcctaaggaattaatggaattattaaaatctagagattatgaaagggaaataaa
038_C10_KM_3.ab1(14>801) → TTCAAAGGTGGTCAGTCTGGATTAAATCTGTCTAAGAATTTCTTAGATCCTAAGGAATTAATGGAATTATTAAAATCTAGAGATTATGAAAGGGAAATAAA
033_H10_KM_8.ab1(21>753) → TTCAAAGGTGGTCAGTCTGGATTAAATCTGTCTAAGAATTTCTTAGATCCTAAGGAATTAATGGAATTATTAAAATCTAGAGATTATGAAAGGGAAATAAA
044_E11_KM_13.ab1(52>730) → TTCAAAGGTGGTCAGTCTGGATTAAATCTGTCTAAGAATTTCTTAGATCCTAAGGAATTAATGGAATTATTAAAATCTAGAGATTATGAAAGGGGAA
036_E10_KM_5.ab1(1>815) ← TTCAAAGGTGGTCAGTCTGGATTAAATCTGTCTAAGAATTTCTTAGATCCTAAGGAATTAATGGAATTATTAAAATCTAGAGATTATGAAAGGGAAATAAA
042_G11_KM_15.ab1(76>626) ← TTCAAAGGTGGTCAGTCTGGATTAAATCTGTCTAAGAATTTCTTAGATCCTAAGGAATTAATGGAATTATTAAAATCTAGAGATTATGAAAGGGAAATAAA
2530 2540 2550 2560 2570 2580 2590 2600 2610 2620
AGGATCAAGAGAGAAGGTCATTAGTGATAAAGATCTAGAGTTGTTGTTAGATCGAAGTGATCTTATTGATCAAATGAATGCTTCAGGACCAATTAAAGAGA
lsh.SEQ(1>2517) → aggatcaagagagaaggtcattagtgataaagatctagagttgttgttagatcgaagtgatcttattgatcaaatgaatgcttcaggaccaattaaagaga
038_C10_KM_3.ab1(14>801) → AGGATCAAGAGAGAAGGTCATTAGTGATAAAGATCTAGAGTTGTTGTTAGATCGAAGTGATCTTATTGATCAA-TGANTGCTTCN
033_H10_KM_8.ab1(21>753) → AGGATCAAGAGAGAAGGTCATTAGTGATAAAGATC


2530 2540 2550 2560 2570 2580 2590 2600 2610 2620
AGGATCAAGAGAGAAGGTCATTAGTGATAAAGATCTAGAGTTGTTGTTAGATCGAAGTGATCTTATTGATCAAATGAATGCTTCAGGACCAATTAAAGAGA
036_E10_KM_5.ab1(1>815) ← AGGATCAAGAGAGAAGGTCATTAGTGATAAAGATCTAGAGTTGTTGTTAGATCGAAGTGATCTTATTGATCAAATGAATGCTTCAGGACCAATTAAAGAGA
042_G11_KM_15.ab1(76>626) ← AGGATCAAGAGAGAAGGTCATTAGTGATAAAGATCTAGAGTTGTTGTTAGATCGAAGTGATCTTATTGATCAAATGAATGCTTCAGGACCAATTAAAGAGA
2630 2640 2650 2660 2670 2680 2690 2700 2710 2720
AGATGGGGATATTCAAGATATTAGAAAATTCTGAAGATTCCAGTCCTGAATGTTTGTTTTAAGCCGCTTTCGAATCTAGAGCCTGCAGTCTCGAGGCATGC
lsh.SEQ(1>2517) → agatggggatattcaagatattagaaaattctgaagattccagtcctgaatgtttgttttaa
036_E10_KM_5.ab1(1>815) ← AGATGGGGATATTCAAGATATTAGAAAATTCTGAAGATTCCAGTCCTGAATGTTTGTTTTAAGCCGCTTTCGAATCTAGAGCCTGCAGTCTCGAGGCATGC
042_G11_KM_15.ab1(76>626) ← AGATGGGGATATTCAAGATATTAGAAAATTCTGAAGATTC
2730
GTCCAG
036_E10_KM_5.ab1(1>815) ← GTCCAG

Appendix III

LSH Cooperates with DNA Methyltransferases To Repress Transcription

MOLECULAR AND CELLULAR BIOLOGY, Jan. 2008, p. 215–226 Vol. 28, No. 1
0270-7306/08/$08.00+0 doi:10.1128/MCB.01073-07

LSH Cooperates with DNA Methyltransferases To Repress Transcription▽†

Kevin Myant and Irina Stancheva*

Wellcome Trust Centre for Cell Biology, University of Edinburgh, Michael Swann Building, Mayfield Road, Edinburgh EH9 3JR, United Kingdom

Received 17 June 2007/Returned for modification 16 August 2007/Accepted 11 October 2007

LSH, a protein related to the SNF2 family of chromatin-remodeling ATPases, is required for efficient DNA methylation in mammals. How LSH functions to support DNA methylation and whether it associates with a large protein complex containing DNA methyltransferase (DNMT) enzymes is currently unclear. Here we show that, unlike many other chromatin-remodeling ATPases, native LSH is present mostly as a monomeric protein in nuclear extracts of mammalian cells and cannot be detected in a large multisubunit complex. However, when targeted to a promoter of a reporter gene, LSH acts as an efficient transcriptional repressor. Using this as an assay to identify proteins that are required for LSH-mediated repression we found that LSH cooperates with the DNMTs DNMT1 and DNMT3B and with the histone deacetylases (HDACs) HDAC1 and HDAC2 to silence transcription. We show that transcriptional repression by LSH and interactions with HDACs are lost in DNMT1 and DNMT3B knockout cells but that the enzymatic activities of DNMTs are not required for LSH-mediated silencing. Our data suggest that LSH serves as a recruiting factor for DNMTs and HDACs to establish transcriptionally repressive chromatin which is perhaps further stabilized by DNA methylation at targeted loci.

In vertebrate genomes, DNA methylation patterns are established during gametogenesis, embryo development, and cell differentiation by enzymes of the DNA cytosine methyltransferase family, which includes the maintenance DNA methyltransferase (DNMT) DNMT1 and the de novo methyltransferases DNMT3A and DNMT3B (1, 24, 45). DNMT1 binds to PCNA and functions primarily during S phase to restore fully methylated CpGs on hemimethylated daughter DNA strands generated during DNA replication (3, 4). DNMT3A and -3B are able to methylate unmethylated DNA and in mouse embryogenesis are required during gastrulation, when DNA methylation patterns are established in differentiating cell lineages of the embryo (24). Mice lacking DNMT proteins die early during embryogenesis and display aberrant expression of retrotransposons and various imprinted and nonimprinted genes (14, 18, 24). Genetic studies with plants and mammals have revealed that additional factors besides DNMTs are required for the establishment of DNA methylation patterns in vivo. Loss-of-function mutations in SNF2 family-related putative chromatin-remodeling ATPases such as the *Arabidopsis thaliana* DDM1 (*d*ecrease in *D*NA *m*ethylation 1) protein and its murine homolog Lsh (*l*ymphoid-*s*pecific *h*elicase) lead to dramatic hypomethylation of the genome in *Arabidopsis thaliana* and mice, respectively (5, 16). Unlike animals that are null for DNMTs, Lsh-deficient mice develop to term but die soon after birth with symptoms of renal failure (5). Interestingly, mice expressing a hypomorph allele of Lsh with targeted disruption of the SNF2 domain survive much longer and display modest hypomethylation of DNA and premature aging (39). Collectively, these studies indicate that the low levels of DNA methylation (~30 to 35% of the wild-type level) in Lsh-deficient mice are compatible with embryonic development.

It is largely unknown how DDM1 and Lsh function to assist DNA methylation. Given the homology of DDM1 and Lsh with chromatin-remodeling ATPases of the SNF2 family, it has been suggested that chromatin remodeling by DDM1 and Lsh may facilitate the processivity of DNMT enzymes on nucleosomal DNA templates (20, 28). In agreement with the requirement for SNF2-like proteins to support DNA methylation in vivo, several studies have demonstrated that in vitro DNMT enzymes methylate DNA assembled into chromatin 10- to 20-fold less efficiently than naked DNA (11, 25, 31). Whether DDM1 or Lsh is able to stimulate DNA methylation on nucleosomal templates in vitro has not been demonstrated. However, it was reported that DDM1 exhibits features common to the SNF2 family of chromatin-remodeling proteins, such as ATP hydrolysis upon stimulation by DNA or nucleosomes and active sliding of mononucleososomes on DNA templates in vitro (2).

The biochemical properties of mammalian Lsh protein have not been investigated in detail, nor has it been determined whether, similar to many other chromatin-remodeling ATPases, Lsh associates with a large protein complex (22). It is also unclear whether Lsh directly interacts with DNMTs and actively recruits them to chromatin. A recent study has found that Lsh coimmunoprecipitates with de novo DNMTs Dnmt3a and Dnmt3b and stimulates methylation of nonmethylated episomal plasmids introduced into mouse embryonic fibroblasts (45). No interaction between Lsh and maintenance methyltransferase Dnmt1 was detected in these assays, suggesting

* Corresponding author. Mailing address: Wellcome Trust Centre for Cell Biology, University of Edinburgh, Michael Swann Building, Mayfield Road, Edinburgh EH9 3JR, United Kingdom. Phone: 44 131 650 7029. Fax: 44 131 650 7063. E-mail: istancheva@ed.ac.uk.

† Supplemental material for this article may be found at http://mcb.asm.org/.

▽ Published ahead of print on 29 October 2007.

that Lsh is involved primarily in de novo DNA methylation of nonmethylated sequences (45).

To address some of these questions, we focused on the properties of human Lsh protein, also known as HELLS (for *he*licase *l*ymphoid-*s*pecific) or PASG (for *p*roliferation-*a*ssociated *S*NF2-like *g*ene). For convenience we refer to this protein as LSH throughout this paper. Using size exclusion chromatography and sedimentation in sucrose gradients, we demonstrate that native LSH can be detected mostly as a free monomeric protein in nuclear extracts of human cells. Thus, if LSH associates with other proteins, such complexes are either unstable or not very abundant. Nevertheless, when targeted to GAL4 binding sites upstream of a reporter gene promoter, LSH acts as an efficient histone deacetylase (HDAC)-dependent transcriptional repressor. The region of LSH required for transcriptional silencing maps to the N-terminal coiled-coil motif, suggesting that transcriptional silencing by LSH is independent of its putative chromatin-remodeling activity. We further demonstrate that LSH coimmunoprecipitates with the HDACs HDAC1 and HDAC2 and that the DNMTs DNMT1 and DNMT3B interact with LSH in vitro and are essential for recruitment of HDACs to LSH in vivo. In cells expressing N-terminally truncated DNMT1 or null for DNMT3B, HDACs do not efficiently coimmunoprecipitate with LSH, suggesting that stable association of HDACs with the rest of the LSH complex requires both DNMTs to be present simultaneously. Taken together, our data demonstrate that LSH forms a transient or not very abundant protein complex in vivo and directly recruits DNMT1, DNMT3B, and HDACs to establish transcriptionally inactive chromatin. Transcriptional silencing by the LSH complex does not immediately result in methylation of DNA, but LSH-mediated increase in the local concentration of DNMTs on chromatin may eventually lead to DNA methylation and further stabilize a silenced chromatin state.

MATERIALS AND METHODS

Size exclusion chromatography and sucrose gradients. A Superose 6 10/300GL gel filtration column was calibrated with gel filtration standards thyroglobulin (669 kDa; Stokes radius [R_s] = 8.5), apoferritin (443 kDa; R_s = 6.1), alcohol dehyrdogenase (150 kDa; R_s = 4.55), bovine serum albumin (66 kDa; R_s = 3.55), and carbonic anhydrase (29 kDa; R_s = 2). One milligram of HeLa nuclear extract or 1 μg of purified recombinant LSH was loaded onto a column preequilibrated with buffer GF150 (20 mM HEPES-KOH [pH 7.9], 3 mM $MgCl_2$, 10% glycerol, and 150 mM KCl). Fractions (0.5 ml) were collected, trichloroacetic acid precipitated, and separated on 7% or 10% polyacrylamide gels. Native and recombinant LSH and native BRG1 were detected by Western blotting using anti-LSH (sc-46665) and anti-BRG1 (sc-H88) antibodies (Santa Cruz). The R_s of LSH was derived from the plotted standards. For sucrose gradients, 50 μg of carbonic anhydrase (2.8S), 50 μg of bovine serum albumin (4.3S) 50 μg of alcohol dehydrogenase (7.4S), and 30 μg of β-amylase (9S) were loaded as standards through a 2-ml linear 5 to 20% sucrose gradient made in buffer NE2 without glycerol (20 mM HEPES-KOH [pH 7.0], 10 mM KCl, 1 mM MgCl2, 0.5 mM dithiothreitol, 0.1% Triton X-100, 420 mM NaCl, and protease inhibitor mix [Sigma P8340]). The protein standard gradient together with an identical gradient loaded with 500 μg HeLa nuclear extract or 1 μg recombinant LSH was spun for 4 h at 50,000 rpm in a Beckman TLS-55 rotor at 4C. Fractions (100 μl) were taken from the top of the gradient and run on 10% polyacrylamide gels. Protein standards were detected by Coomassie blue staining, and LSH was detected by Western blotting. The sedimentation coefficient of LSH was derived from the plotted standards. Calculations to determine the molecular weight of native LSH were applied as described previously (35) using the equation $M_r = 6\pi\eta_{20,\omega} \cdot s_{20,\omega} \cdot R_s \cdot N/(1 - \rho_{20,\omega}\upsilon)$, where R_s is the Stoke's radius (nm), $s_{20,w}$ is the sedimentation velocity ($S \times 10^{-13}$), $\eta_{20,\omega}$ is the viscosity of water at 20°C (0.01002 $g \cdot s^{-1}\ cm^{-1}$), N is Avogadro's number ($6.022 \times 10^{23} \cdot molecules^{-1}$), $\rho_{20,\omega}$ is the density of water at 20°C (0.9981 $g \cdot cm^3$), and υ is the partial specific volume of protein (used 0.725 cm^3/g).

Plasmids and reporter assays. Full-length human LSH (kindly provided by Robert Arceci, The John Hopkins University School of Medicine) or LSH fragments corresponding to amino acids 1 to 226 and 227 to 838 as well as the GAL4 DNA binding domain were PCR amplified and cloned into the pcDNA 3.1 vector (Invitrogen). DNMT3B and DNMT1 were cloned into the pEGFP vector (Clontech). The catalytically inactive $DNMT1^{c/w}$ point mutant was previously described (34) and was provided by Heinrich Leonhardt. $DNMT3B^{c/s}$ was generated by site-directed mutagenic PCR. In reporter assays, GAL4-TK-LUC (500 ng) and pact-βgeo (500 ng) plasmids were cotransfected with 500 ng of the indicated effector plasmids into 2.5×10^5 HCT116 cells or DNMT knockout (KO), cells kindly provided by Bert Vogelstein, using JetPEI reagent (Autogen Bioclear). To test whether LSH-mediated repression was sensitive to HDAC inhibitors, trichostatin A (TSA) was added to the tissue culture medium after transfection to a 100 nM final concentration 24 h prior to lysing the cells. For rescue experiments with HCT116 KO cell lines, 500 ng of DNMT1-green fluorescent protein (GFP) or DNMT3B-GFP plasmid was cotransfected with the reporter plasmids in the presence or absence of GAL4-LSH or GAL4-LSH(1–226). All transfection experiments were performed in triplicate. At 48 h posttransfection, cells were harvested using reporter lysis buffer (Promega), and detection of luciferase activity was carried out according to the manufacturer's guidelines. Luminescence was measured in a TD20/20 luminometer (Turner Designs). Detection of β-galactosidase activity in the same lysates (as described above) was carried out as described previously (44).

Coimmunoprecipitation experiments. Immunoprecipitations were performed from HCT116 nuclear extracts (400 μg of total nuclear protein) with 4 μg anti-LSH (sc-46665), anti-HDAC1 (sc-7872), anti-HDAC2 (sc-7899), polyclonal anti-GFP antibody (a kind gift from K. Sawin, Wellcome Trust Centre for Cell Biology, Edinburgh), anti-Myc (CRUK), or mock control antihemagglutinin (anti-HA) (CRUK) antibodies. Immunoprecipitated complexes bound to protein G beads were washed with NE150 buffer (20 mM HEPES-KOH [pH 7.0], 10 mM KCl, 1 mM MgCl2, 0.5 mM dithiothreitol, 0.1% Triton X-100, 150 mM NaCl, and protease inhibitors) and run on 7% or 10% sodium dodecyl sulfate-polyacrylamide gel electrophoresis. The gels were transferred to nitrocellulose membranes and proteins detected by appropriate primary antibodies (as described above), secondary horseradish peroxidase-conjugated anti-mouse or anti-rabbit antibodies (Sigma), and ECL reagent (Amersham). The antibodies used for the detection of DNMTs were anti-DNMT3B (PA1-884) and anti-DNMT1 (sc-20701).

In vitro pull-down assays. LSH fragments corresponding to amino acids 1 to 503 and 248 to 883 were cloned into the pGEX 4T1 plasmid in frame with glutathione *S*-transferase (GST) tag and designated GST-LSH-N and LSH-C, respectively. pGEX-HDAC1, -2, and -3 were provided by Ronald Evans (Salk Institute, La Jolla, CA) and GST-GFP by Ken Sawin. All GST fusion proteins were expressed in *Escherichia coli* BL21(DE3) and purified on glutathione-Sepharose by standard methods. Full-length DNMT3B and a fragment of DNMT1 corresponding to amino acids 1 to 1125 were cloned into pGAD-T7Rec (Clontech) in frame with an HA tag. Full-length LSH was cloned into pGBK-T7 (Clontech) in frame with a Myc tag. One microgram of pGADT7-DNMT3B, pGADT7-DNMT1, or pGBK-LSH was translated in rabbit reticulocyte lysates (TnT T7 kit; Promega). One hundred nanograms of each GST fusion protein immobilized on glutathione-Sepharose beads and 5 μl of the translation reaction mixtures were used in pull-down experiments performed in 100 μl of NE2 buffer for 1 h at 4C on a rotating wheel. The beads were washed four times with 700 μl NE2 buffer, and LSH-bound DNMT proteins were detected on Western blots with appropriate antibodies. To detect interactions of LSH with DNMT1 in the presence of DNMT3B, 0.5 or 1 μg of baculovirus-produced DNMT3B was added to the pull-down assay mixtures. LSH and DNMT3B were expressed in insect cells and purified as described previously (2, 38).

Chromatin immunoprecipitations. Chromatin immunoprecipitation assays were carried out as previously described (13). Briefly, HEK293 cells containing a stably integrated 5× GAL4-TK-LUC construct (13) were transiently transfected with 10 μg of GAL4-LSH 226 or 10 μg of pCDNA3.1 GALYBD control plasmid. At 48 h posttransfection, the cells were cross-linked with formaldehyde, sheared by sonication, and immunoprecipitated for 2 h at 4°C with the following antibodies: anti-HA control (CRUK), anti-H3 (Upstate; 05-499, lot 25198), anti-H3K9ac (Upstate; 07-352, lot 28741), anti-H4 (Upstate; 07-108, lot 25296), and anti-H4K12ac (Upstate; 07-595, lot 28885). Immunoprecipitated complexes were then captured on protein G Dynabeads (Invitrogen). Following extensive washing, DNA was recovered and 1 μl of chromatin-immunoprecipitated DNA was used in 30 cycles of PCR. Primers for the 5× GAL4-TK promoter were

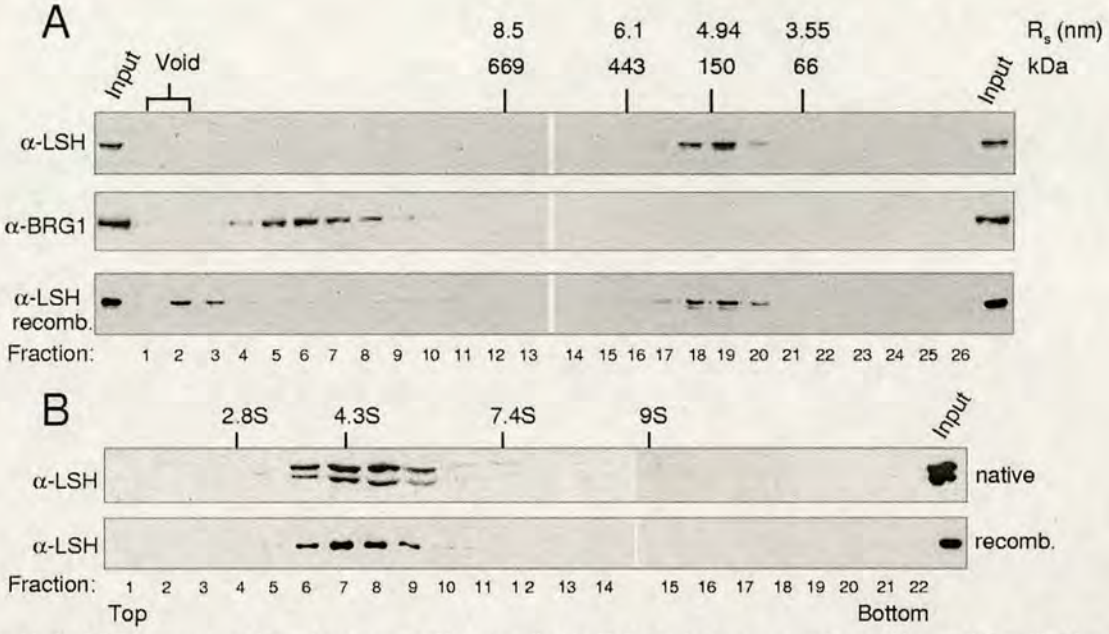


FIG. 1. Native LSH is a monomer in nuclear extracts of human cells. (A) HeLa nuclear extracts and recombinant LSH purified from insect cells were fractionated on a Superose 6 size exclusion column and run on sodium dodecyl sulfate-polyacrylamide gels. LSH and BRG1 proteins were detected in the collected fractions by Western blotting with appropriate antibodies. The molecular masses and Stokes radii (R_s) of marker proteins used to calibrate the column are indicated at the top. (B) HeLa nuclear extracts and recombinant LSH were fractionated in 5 to 20% sucrose gradients. LSH was detected in gradient fractions on Western blots with anti-LSH antibodies. Sedimentation coefficients of marker proteins are indicated above the blot.

5′AATTGCTCAACAGTATGAACATTTC and 3′CAATTGTTTTGTCACGA TCAAAGGA. Primers for the GAPDH (glyceraldehyde-3-phosphate dehydrogenase) promoter were 5′-GAGGCTGTGAGCTGGCTGTC and 3′-CAGAGC AGAGTAGCAAGAGCAAGG.

RESULTS

Native LSH cannot be detected in a large protein complex. Within the SNF2 family of ATPases, LSH is most closely related to the ISWI subfamily of chromatin-remodeling factors, which have diverse functions in transcriptional regulation (16, 20). Most SNF2 family members, including ISWI proteins, have intrinsic ATPase activity but usually associate with more than one additional polypeptide in vivo (8, 22). These additional subunits of chromatin-remodeling complexes either modulate the activity of SNF2 ATPases or target them to specific chromatin regions (22). Given the possibility that LSH may act as an accessibility factor for DNMTs and that mouse Lsh coimmunoprecipitates with Dnmt3a and Dnmt3b, we investigated whether LSH in human cells can be detected in a large protein complex. We fractionated HeLa nuclear extract by size exclusion chromatography and analyzed the elution profile of LSH by Western blotting. LSH appeared in four of the eluted fractions (fractions 17 to 20) with a peak corresponding to a molecular mass of about 150 kDa (Fig. 1A, top). In comparison, another chromatin-remodeling protein, BRG1, which is usually present within a 2-MDa complex (26, 36), eluted at a much higher molecular mass as expected (fractions 5 to 11) (Fig. 1A, middle). Since LSH eluted at an apparent molecular mass slightly greater than its theoretical mass of ~97 kDa, we considered the possibility that it may associate with a relatively small protein(s). However, in a similar experiment recombinant LSH purified from insect cells eluted with a profile very similar to that of the native LSH, suggesting that the vast majority of LSH in HeLa nuclear extracts does not associate with additional proteins (Fig. 1A, bottom). Nuclear extracts from other human cell lines, cells synchronized in different stages of the cell cycle (see Fig. S1 in the supplemental material), and mouse embryonic stem (ES) cells, as well as extracts prepared from nuclear pellets, containing most of the insoluble chromatin, showed identical elution of LSH at ~150 kDa (not shown).

The molecular mass of a native protein or a protein complex can be determined accurately by an equation derived by Siegel and Monty, which combines the Stokes radius (a hydrodynamic radius of a molecule freely tumbling in solution) calculated from size exclusion chromatography with the sedimentation coefficient determined by separation in sucrose gradient (35). Based on the elution profile of protein standards from size exclusion chromatography we calculated the Stokes radius of LSH to be 4.94 nm. To establish the sedimentation coefficient of LSH, we fractionated HeLa and mouse ES cell nuclear extracts on 5 to 20% sucrose gradients and detected LSH in the gradient fractions by Western blots (Fig. 1B). In these experiments the sedimentation coefficient of both human and mouse LSH was calculated to be 4.5S, relative to protein standards. Applying the calculated Stokes radius and sedimentation coefficient of LSH in the Siegel-Monty equation, we derived a molecular mass of 91.5 kDa. The calculated molecular mass of native LSH is close to the predicted theoretical mass of LSH monomer (97 kDa), confirming that the vast majority of LSH

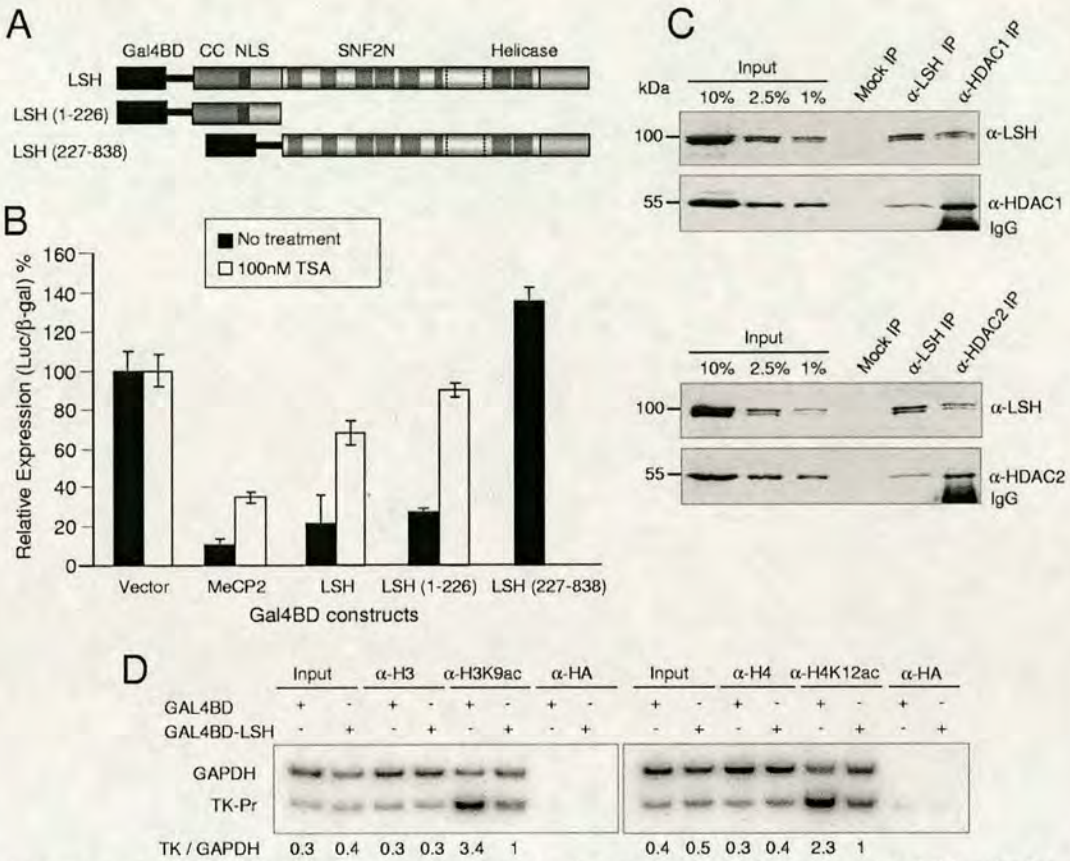


FIG. 2. LSH functions as an HDAC-dependent transcriptional repressor in vivo. (A) Schematic drawing of the full-length LSH and truncated LSH proteins fused to a GAL4BD. The functional domains of LSH, such as the coiled-coil domain (CC), the nuclear localization signal (NLS), and the eight conserved SNF2 motifs in the SNF2N and helicase domains, are indicated. (B) GAL4BD fusions of LSH were cotransfected into HCT116 cells with a luciferase reporter plasmid carrying GAL4 binding sites upstream of TK promoter and a control plasmid expressing β-galactosidase from an actin promoter. The relative expression of the reporters represents the ratio of luciferase to β-galactosidase products. MeCP2 was used as a control. The white bars represent experiments carried out in the presence of 100 nM TSA, which partially alleviates LSH- and MeCP2-mediated repression of the luciferase reporter. The error bars represent standard deviations of the means. (C) Coimmunoprecipitation experiments with anti-LSH, anti-HDAC1, and anti-HDAC2 antibodies from HCT116 nuclear extracts. Anti-HA antibodies were used in mock immunoprecipitations (IP) to control for nonspecific protein binding. (D) Chromatin immunoprecipitation with antibodies against acetylated H3-K9 and H4-K12. The TK promoter in cells transfected with GAL4BD-LSH is hypoacetylated compared to the GAPDH promoter used as an internal control. Anti-HA is a nonspecific control antibody.

protein in nuclear extracts of mouse and human cells is present as a free monomeric peptide.

LSH is an HDAC-dependent transcriptional repressor. Given that in fractionation experiments we failed to detect LSH within a large protein complex, we decided to explore whether LSH could interact transiently with other proteins. We initially intended to use a full-length GAL4 binding domain (GAL4BD)-tagged LSH in a yeast two-hybrid screen. However, we found that GAL4BD-LSH strongly repressed adenine and histidine genes in yeast reporter strains. The *Saccharomyces cerevisiae* genome encodes a protein, YFR038W, which shares 39% identity and 57% similarity with human LSH. Therefore, it was possible that the mammalian protein could mimic some of the protein-protein interactions of YFR038W that facilitate transcriptional repression. To investigate further whether LSH could act as a transcriptional repressor in mammalian cells, we cotransfected colorectal HCT116 cells with a plasmid expressing full-length LSH fused to GAL4BD and two reporter plasmids. The first reporter plasmid carried five GAL4 binding sites upstream of a thymidine kinase (TK) promoter driving the expression of the firefly luciferase gene, while the other control plasmid lacked GAL4 binding sites and expressed β-galactosidase from an actin promoter. The effect of LSH on transcription from the targeted and nontargeted reporter was measured as a ratio of luciferase to β-galactosidase expression. A GAL4BD-tagged transcriptional repression domain (TRD) of methyl-CpG binding protein MeCP2, which is known to strongly repress transcription in such assays, served as a control (21). In these experiments the full-length GAL4BD-LSH as well as GAL4BD-MeCP2 consistently reduced the expression of the luciferase reporter to about 20 to 25% and 10%, respectively, of the levels observed in cells transfected with an empty vector (Fig. 2B). This suggests that, like MeCP2, LSH can function as an efficient transcriptional repressor when targeted to a promoter of a reporter gene in human cells.

To investigate whether a specific domain of LSH was responsible for transcriptional silencing, we tested several LSH deletion constructs in the luciferase reporter assay. Interestingly, the N-terminal portion of LSH (amino acids 1 to 226), containing a predicted coiled-coil motif, was sufficient to repress the reporter gene to levels comparable to those observed with full-length LSH. A polypeptide corresponding to the SNF2 and helicase domains of LSH (amino acids 226 to 238), did not significantly affect the expression of the luciferase reporter (Fig. 2A and B). These experiments indicate that the 226-amino-acid coiled-coil region in the N terminus of LSH functions as a TRD and is sufficient for silencing of the luciferase reporter.

Since coiled-coil regions of many proteins are engaged in protein-protein interactions, we asked whether the TRD of LSH functionally cooperates with corepressor proteins such as HDACs. Consistent with this hypothesis, transcriptional repression by MeCP2, full-length LSH, and LSH(1–226) was partially alleviated when we repeated the luciferase reporter assays in the presence of the deacetylase inhibitor TSA (Fig. 2B). This indicates that transcriptional repression by LSH may operate through an interaction between the coiled-coil domain and HDACs. To explore this further, we asked whether HDACs coimmunoprecipitate with the endogenous LSH from HCT116 nuclear extracts. Antibodies against LSH but not control anti-GFP antibodies (mock immunoprecipitation) immunoprecipitated LSH and coimmunoprecipitated HDAC1 and HDAC2 (Fig. 2C). We could also detect LSH in reciprocal immunoprecipitations with anti-HDAC1 and anti-HDAC2 antibodies (Fig. 2C). Additionally, chromatin immunoprecipitations from cells carrying stably integrated copies of the luciferase reporter showed a threefold decrease of acetylated lysine 9 of histone H3 and a twofold decrease of acetylated lysine 12 of H4 at the TK promoter after the cells were transfected with GAL4BD-LSH (Fig. 2D). Further RNA interference knockdown experiments with HDAC1 and HDAC2 in HCT116 cells revealed that depletion of these two HDACs, but not HDAC3, could, similarly to TSA treatment, alleviate the repression of luciferase reporter (see Fig. S2 in the supplemental material). This indicates that HDAC1 and HDAC2 are essential for LSH-mediated repression.

Transcriptional repression by LSH requires DNMT1 and DNMT3B. Zhu et al. have recently reported that mouse Lsh coimmunoprecipitates with the de novo DNMTs Dnmt3a and Dnmt3b, but not with the maintenance DNMT Dnmt1, from extracts of mouse embryonic fibroblasts (45). Dnmt1, Dnmt3b, and Dnmt3a have also been shown to interact with each other and to coimmunoprecipitate and copurify with HDAC1 and HDAC2 (Fig. 3A) (9, 17, 31–33). Given all these complex interactions, we next investigated whether DNMTs contribute to LSH-mediated transcriptional repression. KO HCT116 cell lines that are genetically null for DNMT1, DNMT3B, DNMT1 and DNMT3B, or DNMT3A and DNMT3B have been generated by homologous recombination (15, 29, 30). Subsequently it was found that most DNMT1 KO cell lines, including the DNMT1/DNMT3B double-KO (DKO) line, express a truncated DNMT1 protein missing 150 amino acids of the N terminus, including the regions essential for binding to DNMT3A, DNMT3B, and PCNA (Fig. 3A; see Fig. 6A) (6, 37). To our surprise, neither GAL4BD-LSH nor GAL4BD-LSH(1–226) repressed the reporter luciferase gene in DNMT1 KO and DNMT3B KO HCT116 cells (Fig. 3B and C). The same was observed in DNMT1/DNMT3B DKO and DNMT3A/DNMT3B DKO cells (not shown). However, when we cotransfected Dnmt1-GFP with either GAL4BD-LSH or GAL4BD-LSH(1–226) and the reporter plasmids into DNMT1 KO cells, luciferase expression was reduced to levels comparable to those observed in wild-type HCT116 cells (Fig. 3B). Interestingly, a catalytically inactive Dnmt1 carrying a single point mutation, C1229W (34), also rescued LSH-mediated repression in DNMT1 KO cells as well as wild-type Dnmt1. This indicates that the enzymatic activity of DNMT1 is not required for transcriptional silencing of the luciferase reporter.

To determine which region of DNMT1 is required for LSH-mediated repression, we tested whether shorter DNMT1 constructs could restore the repression of the luciferase reporter in DNMT1 KO cells (Fig. 3A). DNMT1(1–1125)-GFP, which contains the known interaction sites with DNMT3A, DNMT3B, and HDACs but is lacking the C-terminal catalytic domain, could partially restore the repression by GAL4BD-LSH and GAL4BD-LSH(1–226), while the shorter proteins DNMT1(1–701)-GFP and DNMT1(1–250)-GFP did not (Fig. 3B). This suggests that a relatively large portion of the DNMT1 N terminus and perhaps some of the C-terminal amino acids are involved in protein-protein interactions that are crucial for LSH-mediated repression. Cotransfection of DNMT3B-GFP with GAL4BD-LSH or GAL4BD-LSH(1–226) into DNMT1 KO cells also led to a modest (~30%) decrease of transcription from the reporter gene (Fig. 3B). This indicates that elevated levels of DNMT3B are not sufficient to restore LSH-mediated repression in cells expressing low levels of truncated DNMT1.

We next attempted to rescue the LSH-mediated repression of the luciferase reporter in DNMT3B KO cells. As in the case of DNMT1, when we cotransfected either GAL4BD-LSH or GAL4BD-LSH(1–226) and the reporter plasmids with either wild-type GFP-DNMT3B or catalytically inactive GFP-DNMT3B^{C640S}, the repression of the luciferase gene was largely restored (Fig. 3C). Interestingly, cotransfection of GAL4BD-LSH with DNMT1-GFP into DNMT3B KO cells could also reduce the expression of the luciferase reporter by approximately 45 to 50%, indicating that DNMT1, when overexpressed, could (although not very efficiently) compensate for the lack of DNMT3B (Fig. 3C). Collectively, these experiments demonstrate that both DNMT1 and DNMT3B are required for effective GAL4-LSH-mediated transcriptional silencing.

The interaction of LSH with HDAC1 and HDAC2 is lost in DNMT KO cells. Given that the repression by LSH was sensitive to TSA and that HDAC1 and HDAC2 coimmunoprecipitated with LSH from nuclear extracts of wild-type HCT116 cells, we next examined whether these interactions remain intact in DNMT KO cells, where GAL4BD-LSH was unable to silence the reporter gene. When we immunoprecipitated the endogenous LSH from either DNMT1 KO or DNMT3B KO extracts, we could not detect HDAC1 and HDAC2 in anti-LSH immunoprecipitations (Fig. 4A, B, C, and D, top panels; compare with Fig. 2C). Consistently, anti-HDAC1 and anti-HDAC2 antibodies efficiently immunoprecipitated HDACs but failed to coimmunoprecipitate LSH from extracts of DNMT1 and DNMT3B KO cells (Fig. 4A, B, C, and D, bottom

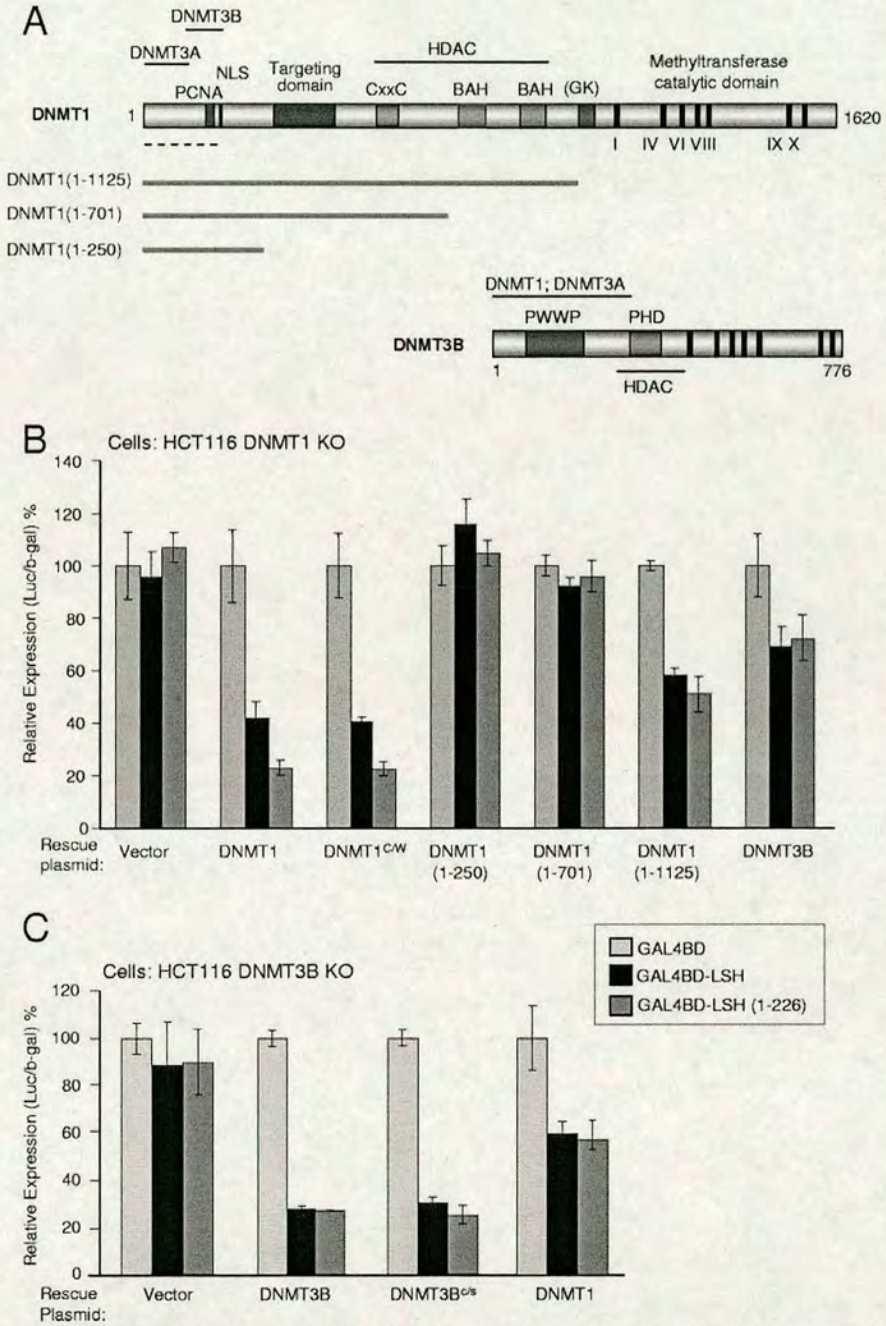


FIG. 3. Transcriptional repression by LSH requires DNMT3B and the N-terminal portion of DNMT1. (A) Schematic representation of DNMT1 and DNMT3B proteins with their functional domains. The cysteine-rich putative DNA binding CxxC domain, bromo-adjacent homeobox motifs (BAH), GK-rich repeats, the domain involved in targeting to replication foci, and the catalytic part of DNMT1 are indicated. Mapped interactions with DNMT3A, DNMT3B, PCNA, and HDAC1 and -2 are shown above the diagram. The dashed line indicates the portion of DNMT1 that has been spliced out in DNMT1 KO HCT116 cells with targeted disruption of the DNMT1 gene (6, 37). The DNMT3B protein contains a DNA binding PWWP motif, a PHD, domain and a conserved catalytic DNMT domain. The portions of DNMT3B interacting with DNMT1 and DNMT3A are indicated. (B) Neither full-length GAL4BD-LSH nor the TRD of LSH, GAL4-LSH(1–226), could silence the luciferase reporter in DNMT1 KO cells. Cotransfection of GAL4BD-LSH proteins together with wild-type GFP-tagged DNMT1, catalytically inactive DNMT1$^{C/W}$, and the N-terminal portion of DNMT1(1–1125) can rescue the repression of luciferase reporter in DNMT1 KO cells. Shorter N-terminal DNMT1 proteins [DNMT1(1–250) and DNMT1(1–701)] did not rescue the repression of the luciferase reporter gene. DNMT3B was used as an additional control. Error bars indicate standard deviations. (C) GAL4BD-LSH and GAL4BD-LSH(1–226) did not repress the luciferase reporter in DNMT3B KO cells. Cotransfection of LSH proteins with either GFP-DNMT3B or a catalytically inactive GFP-DNMT3B$^{C/S}$ restored the repression of the reporter to levels observed in wild-type HCT116 cells. Cotransfection of GAL4BD-LSH with GFP-DNMT1 also reduced the repression of luciferase in DNMT3B KO cells, although not as efficiently as DNMT3B.

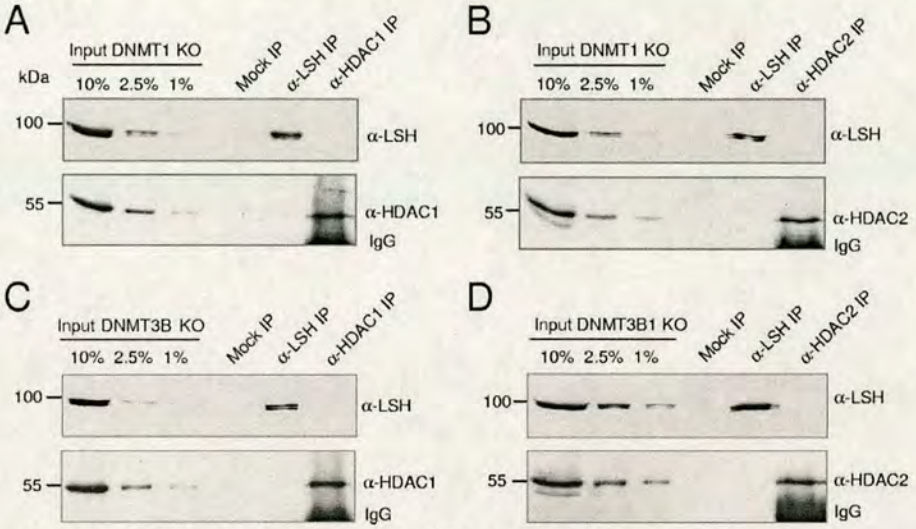


FIG. 4. The interactions of LSH with HDACs are lost in DNMT KO cells. (A and B) Endogenous LSH, HDAC1, and HDAC2 could be efficiently immunoprecipitated (IP) from extracts of DNMT1 KO cells. However, LSH could not be detected in HDAC immunoprecipitations, nor was HDAC1 or HDAC2 detected in LSH immunoprecipitations. IgG, immunoglobulin G. (C and D) LSH, HDAC1 and HDAC2 do not coimmunoprecipitate from extracts of DNMT3B KO cells. Anti-GFP antibodies were used as a control for nonspecific interactions.

panels). As the protein levels of LSH, HDAC1, and HDAC2 in KO cells did not differ significantly from those in the wild-type HCT116 cells, these experiments indicate that DNMT1 and/or DNMT3B could either directly or indirectly recruit HDAC1 and HDAC2 to LSH. Notably, the presence of both DNMTs is required to promote stable interactions of HDACs with LSH.

LSH coimmunoprecipitates with DNMT1 and DNMT3B. Our experiments thus far suggested that LSH participates in a protein complex (or complexes) that contains DNMT1, DNMT3B, HDAC1, and/or HDAC2. To investigate further whether we could detect an interaction between LSH and DNMT1, we cotransfected DNMT1 KO cells with GAL4DB-LSH and DNMT1-GFP and used anti-GAL4 and anti-GFP antibodies for immunoprecipitation experiments. Anti-GFP antibodies detected DNMT1-GFP in anti-GAL4 immunoprecipitations but not in control anti-HA immunoprecipitations, suggesting that GAL4BD-LSH and DNMT1-GFP interact with each other (Fig. 5A, top panel). In similar experiments we found that GAL4DB-LSH and DNMT3B-GFP cotransfected into DNMT3B KO cells also coimmunoprecipitate (Fig. 5A,

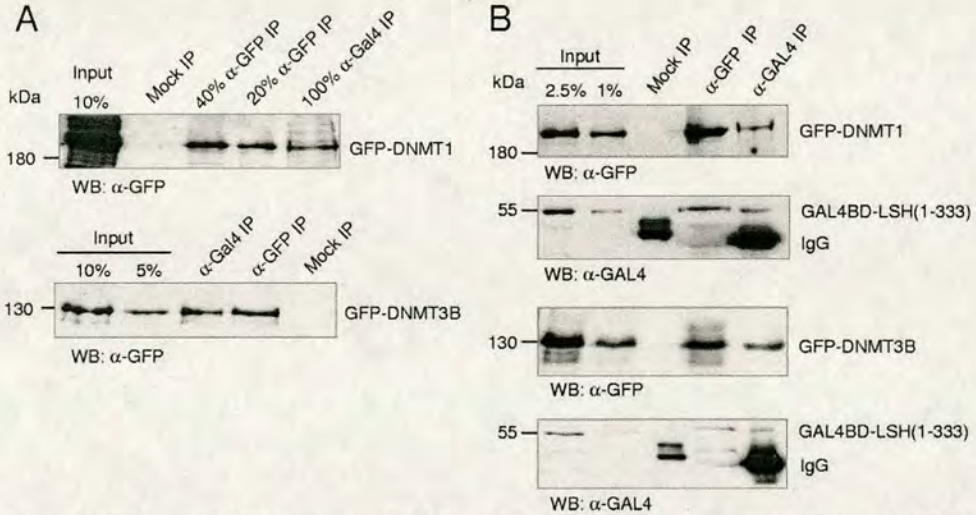


FIG. 5. The coiled-coil TRD domain of LSH interacts with DNMTs. (A) GAL4BD-LSH coimmunoprecipitated with GFP-DNMT1 when both proteins were coexpressed in DNMT1 KO cells. We could detect only about 20% of GFP-DNMT1 in the immunoprecipitations (IP) with anti-GAL4BD antibodies. GAL4BD-LSH coimmunoprecipitated more efficiently with GFP-DNMT3B when both proteins were coexpressed in DNMT3B KO cells. WB, Western blot. (B) GAL4BD-LSH(1–333) protein, containing the coiled-coil TRD domain of LSH, coimmunoprecipitates with GFP-DNMT1 and GFP-DNMT3B from extracts of DNMT1 and DNMT3B KO cells, respectively. Anti-HA antibodies (mock immunoprecipitation) were used as a control.

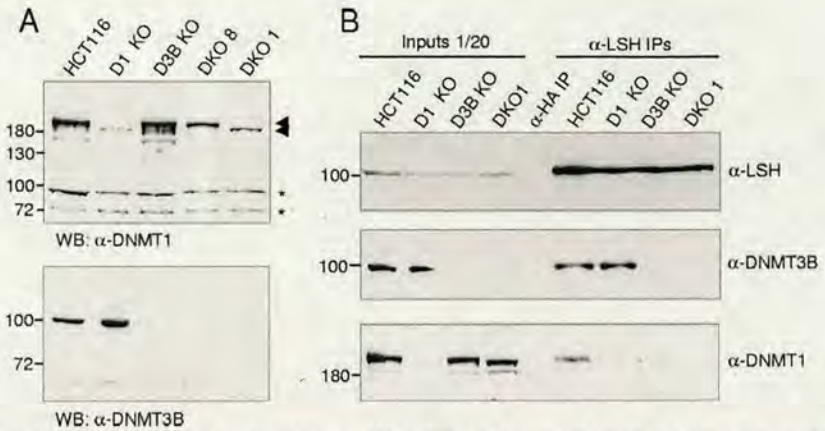


FIG. 6. The interaction of LSH with DNMT1 in vivo requires DNMT3B. (A) Western blots (WB) with antibodies against the C terminus of DNMT1 and the N terminus of DBMT3B on nuclear extracts of HCT116 and KO cell lines. Note that a truncated form of DNMT1 is detectable in DNMT1 KO (D1 KO2) cells as well as in DNMT1/DNMT3B DKO (DKO1) cells, while a second DKO cell line (DKO8) expresses full-length DNMT1. DNMT3B is detectable only in HCT116 and DNMT1 KO2 cells. In the top panel, the full-length and the truncated DNMT1s are indicated with arrowheads. The asterisks indicate nonspecific bands that serve as loading controls. (B) Anti-LSH antibodies efficiently immunoprecipitate LSH from nuclear extracts of HCT116 and KO cells. In identical anti-LSH immunoprecipitations (IP), DNMT3B immunoprecipitates with LSH from HCT116 and DNMT1 KO extracts, indicating that the N terminus of DNMT1 is not required for the interaction of DNMT3B with LSH. In contrast, DNMT1 coimmunoprecipitates with LSH only from HCT116 cells, suggesting that the presence of DNMT3B mediates the recruitment of DNMT1 to LSH.

bottom panel). Consistent with our reporter assays, when we cotransfected the cells with the coiled-coil TRD domain of LSH and either DNMT1-GFP or DNMT3B-GFP, we could detect LSH(1–333) coimmunoprecipitating with each of the two DNMTs (Fig. 5B). These experiments clearly demonstrate that the TRD of LSH is necessary and sufficient for the interaction of LSH with DNMT1 and DNMT3B in vivo.

DNMT3B is required for the recruitment of DNMT1 to LSH. As the luciferase reporter assays and coimmunoprecipitation experiments described above relied on overexpression of tagged proteins, we next examined whether the endogenous LSH interacts with DNMTs in HCT116, DNMT1, and DNMT3B KO cells. In agreement with other studies (6, 37), an antibody against the C terminus of DNMT1 detected a truncated DNMT1 protein in nuclear extracts of DNMT1 KO and DNMT1/DNMT3B DKO1 cells compared to the wild-type HCT116 and DNMT1/DNMT3B DKO8 cells (Fig. 6A, top panel). As observed by others, the truncated DNMT1 protein was significantly more abundant in one of the DKO cell lines (DKO1) than in the KO cell line, where DNMT1 was barely detectable (Fig. 6A, top panel). We did not detect DNMT3B protein either in DNMT3B KO or in any of the two DKO cell lines (Fig. 6A, bottom panel). We further used extracts from four of these cell lines, HCT116, DNMT1 KO, DNMT3B KO, and DKO1, to immunoprecipitate LSH and asked whether DNMT1 and DNMT3B could be detected in LSH immunoprecipitations (Fig. 6B, top panel). DNMT3B was present in anti-LSH immunoprecipitations from HCT116 cells, as expected, but it was also detectable in anti-LSH immunoprecipitations from DNMT1 KO cells (Fig. 6B, middle panel). However, DNMT1 coimmunoprecipitated with LSH only from extracts of wild-type HCT116 cells and not from extracts of DNMT3B KO or DKO1 cells (Fig. 6B, bottom panel). These results indicate that DNMT1 does not efficiently interact with LSH in the absence of DNMT3B (Fig. 6A). On the other hand, the presence of DNMT1 may not be required for the interaction of DNMT3B with LSH, given that approximately equal amounts of DNMT3B coimmunoprecipitate with LSH from HCT116 and DNMT1 KO cells expressing a truncated DNMT1 that does not contain the DNMT3B interaction domain (Fig. 3A and 6A). Taken together with the reporter luciferase assays, these immunoprecipitation experiments suggest that LSH may exist in a complex with DNMT3B with or without DNMT1. However, an LSH complex containing DNMT1, HDAC1, and HDAC2 must also include DNMT3B.

DNMT3B directly binds to LSH, while DNMT1 and HDACs do not. To explore whether any of the proteins that coimmunoprecipitate with LSH bind to LSH directly, we expressed and purified from *E. coli* two GST-tagged recombinant LSH polypeptides, designated LSH-N and LSH-C (Fig. 7A). LSH-N contains amino acids 1 to 503 of LSH and includes the N-terminal coiled-coil and the SNF2 domain. LSH-C corresponds to amino acids 248 to 883 and includes the SNF2 domain and the remainder of the LSH C terminus. We used these two proteins bound to glutathione-Sepharose beads or a control GST-GFP protein to pull down HA-tagged DNMT1 (amino acids 1 to 1125) and full-length DNMT3B in vitro translated in rabbit reticulocyte lysate. Neither LSH-N nor LSH-C could bind DNMT1 in these assays (Fig. 7B). In contrast, LSH-N but not LSH-C or GFP efficiently pulled down DNMT3B (Fig. 7B). As DNMT1 and DNMT3B are known to bind each other (17) and LSH did not coimmunoprecipitate with DNMT1 from extracts of DNMT3B-deficient cells, we asked whether LSH-N would be able to pull down DNMT1 when DNMT3B was present. To investigate this, we added increasing amounts of recombinant DNMT3B expressed and purified from insect cells to reticulocyte lysate containing in vitro-translated DNMT1 (Fig. 7C, bottom panel). We could detect increasing amounts of DNMT1 being pulled down by GST-LSH-N only when the baculovirus-produced DNMT3B

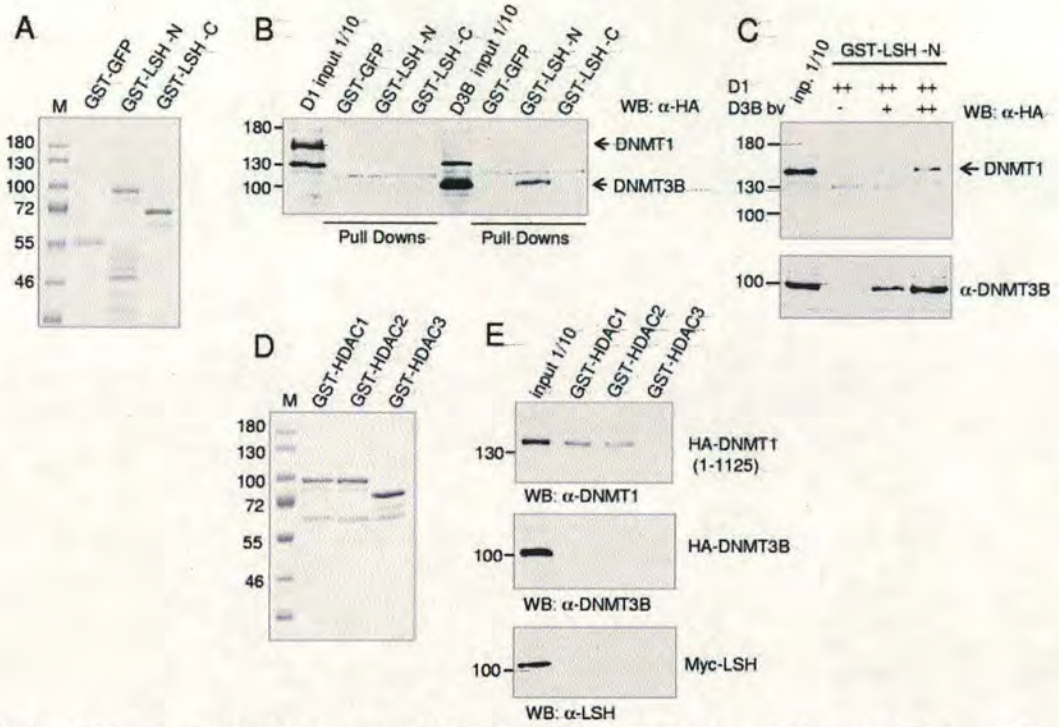


FIG. 7. LSH directly interacts with DNMT3B but not with DNMT1 or HDACs. (A) A Coomassie blue-stained gel shows GST-tagged purified GFP, LSH-N, and LSH-C proteins used in the pull-down assays (B and C). (B) The N terminus of LSH binds in vitro-translated DNMT3B. Neither the N terminus nor the C terminus of LSH pulls down in vitro-translated DNMT1 (amino acids 1 to 1125). WB, Western blot. (C) GST-LSH-N can pull down DNMT1 in the presence of recombinant DNMT3B (0.5 and 1 μg). (D) A Coomassie blue-stained gel shows purified GST-HDAC1, HDAC2, and HCAC3. (E) GST-HDAC1 and GST-HDAC2 pull down in vitro-translated DNMT1 but not DNMT3B or LSH.

was present (Fig. 7C). Consistent with our coimmunoprecipitation assays, these experiments imply that DNMT3B directly binds to the N terminus of LSH, while the interaction of DNMT1 with LSH in vitro and in vivo requires the presence of DNMT3B.

In order to examine whether HDAC1 and HDAC2 could directly bind to LSH, DNMT1, or DNMT3B, we expressed GST-tagged full-length HDAC1, HDAC2, and, as a control, HDAC3 in *E. coli* and bound them to glutathione-Sepharose (Fig. 7D). We next used the Sepharose-bound HDAC proteins to pull down in vitro-translated HA-tagged DNMT1(1–1125), full-length DNMT3B, and Myc-tagged LSH (Fig. 7E). In agreement with previous reports, we could detect DNMT1 bound to HDAC1 and HDAC2 but not to HDAC3 (Fig. 7E, top panel). However, neither DNMT3B nor LSH was pulled down by GST-HDAC1, -2, or -3 (Fig. 7E, middle and bottom panels). These in vitro experiments are consistent with our reporter assays and coimmunoprecipitation results suggesting that DNMT1 recruits HDAC1 and HDAC2 to the LSH-bound DNMT3B (Fig. 8).

DISCUSSION

The plant SNF2 family protein DDM1 and its mammalian homolog LSH were initially identified as proteins essential for the establishment of DNA methylation in vivo (5, 16). DDM1-deficient plants and mice with targeted disruption of the Lsh gene develop with dramatically reduced levels of methylated cytosine within their genomes and are defective in silencing of various transposable elements and a few specific genes (7, 12, 19). In mice, lack of Lsh is not essential for embryonic development but leads to postnatal death (5). Cytological studies

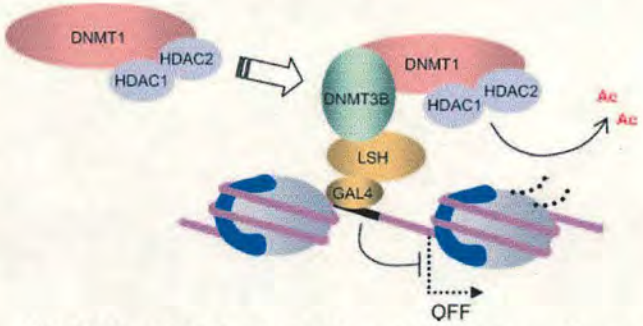


FIG. 8. A model of how the LSH-associated protein complex acts to repress transcription. Our experiments are consistent with a model where DNA-bound LSH recruits a complex that includes DNMT3B, DNMT1, HDAC1, and HDAC2. LSH-associated HDACs remove acetyl groups (Ac) from histone tails, generating deacetylated chromatin incompatible with transcriptional activation. LSH does not directly interact with DNMT1 and HDACs but requires DNMT3B for the assembly of the repressive complex. On the other hand, HDAC1 and HDAC2 require both DNMT1 and DNMT3B for association with the LSH complex. The order of these interactions explains why in cells expressing N-terminally truncated DNMT1, which does not bind DNMT3B, or in cells lacking DNMT3B, LSH-mediated repression is disrupted.

and experiments with ES cells have shown that although Lsh is present together with DNMT1 at replication foci during late S phase, Lsh is dispensable for maintenance of DNA methylation on replicating episomal plasmids (42, 45). It was further found that Lsh coimmunoprecipitates with de novo methyltransferases Dnmt3a and Dnmt3b and that cells treated with Lsh small interfering RNA have reduced de novo DNA methylation (45). Taken together, these studies led to the conclusion that Lsh is involved primarily in de novo DNA methylation but is dispensable for the maintenance of DNA methylation during DNA replication.

Despite the important role of Lsh in supporting efficient DNA methylation in mammalian cells, very little is known about how this protein interacts with the DNA methylation machinery. This lack of information may be due to difficulties in identifying proteins that interact with Lsh. Indeed, based on the observed molecular mass of the native LSH calculated from biochemical fractionation experiments, we found that the vast majority of the nuclear pool of LSH in human and mouse cells appears to consist of free monomeric LSH. Therefore, purification of a low-abundance LSH complex by traditional chromatography methods would not have been feasible. Additionally, we observed that the full-length LSH had a strong repressive effect when recruited to the promoters of reporter genes in yeast, further complicating the use of yeast two-hybrid screens for identification of LSH-interacting partners. Taking advantage of our discovery that LSH behaved as a transcriptional repressor in mammalian cells as well as in yeast, we used luciferase reporter assays as a tool to determine which region of LSH mediates transcriptional repression and to identify proteins that interact with LSH in vivo. By recruiting truncated LSH polypeptides to GAL binding sites upstream of the luciferase reporter gene driven by the TK promoter, we found that the N-terminal coiled-coil domain of LSH spanning amino acids 1 to 226 was necessary and sufficient for transcriptional silencing of the reporter. Therefore, it seemed plausible that this region of LSH, which we designated the TRD, interacts with corepressor proteins that modify chromatin into a transcriptionally nonpermissive state.

As LSH-mediated silencing of the luciferase reporter was sensitive to the deacetylase inhibitor TSA, we performed coimmunoprecipitation assays and found that two HDACs, HDAC1 and HDAC2, interact with LSH in vivo. Further, in vitro experiments revealed that neither HDAC1 nor HDAC2 directly binds to LSH, suggesting that they are recruited to LSH by other proteins. Given the previous reports that LSH coimmunoprecipitates with de novo DNMTs from extracts of mouse cells, we examined whether transcriptional silencing of the luciferase reporter required the presence of DNMTs. In human colorectal carcinoma HCT116 cells with targeted disruption of the DNMT3B gene (DNMT3B KO1) as well as cells expressing low levels of N-terminally truncated DNMT1 protein (DNMT1 KO2), we observed a significant reduction of LSH-mediated repression, indicating that these two DNMTs could be involved in the recruitment of HDAC1 and HDAC2 to LSH. Our further experiments confirmed that LSH does not coimmunoprecipitate with HDACs from extracts of DNMT3B KO1 and DNMT1 KO2 cells. We also observed that the endogenous LSH as well as GAL4BD-tagged LSH coimmunoprecipitated with DNMT1 and DNMT3B from extracts of HCT116 cells.

It has been previously reported that both DNMT1 and DNMT3B interact with HDAC1 and HDAC2 in vitro and in vivo (9, 10, 32). Therefore, it was possible that DNMT1 and DNMT3B bind to LSH independently of each other and promote the recruitment of HDACs to LSH-targeted chromatin. However, we found that although LSH coimmunoprecipitated with DNMT3B from DNMT1 KO cells, it did not interact with HDACs or the truncated DNMT1 in these cells and was unable to silence the luciferase reporter. Taken together, our data suggest that DNMT1-bound HDAC1 and HDAC2 are recruited to LSH indirectly via DNMT3B (Fig. 8). Our in vitro pull-down experiments further support the in vivo reporter assays and coimmunoprecipitation data. Interestingly, we observed that when DNMT1 was overexpressed in DNMT3B KO cells, it could, to some extent, rescue the transcriptional repression of luciferase reporter mediated by GAL4BD-LSH. As DNMT1 does not directly bind to LSH, it is possible that the recruitment of DNMT1 to LSH may occur via DNMT3A or some other protein. However, the first possibility seems unlikely, since we did not detect DNMT3A protein in any of the HCT116 wild-type or KO cell lines except DKO cells (not shown). Given that overexpression of DNMT3B in DNMT1 KO cells transfected with GAL4DB-LSH also reduced the expression of the luciferase reporter by about 30%, it is possible that DNMT3B can either function in transcriptional repression independently of the truncated DNMT1 or, to some extent but not very efficiently, interact with the C terminus of DNMT1. The second interpretation seems plausible, since the N-terminal portion of DNMT1(1–1125) did not fully rescue LSH-mediated repression in DNMT1 KO cells compared to the full-length DNMT1.

In summary, we have investigated whether and how LSH interacts with the DNA methylation machinery in human cells and found that the LSH-associated complex contains at least four proteins: DNMT3B, DNMT1, HDAC1, and HDAC2. We show that GAL4BD-mediated targeting of LSH to a TK promoter driving the expression of a reporter gene results in transcriptional silencing, which is dependent upon the recruitment of HDACs via DNMT3B and DNMT1. However, transcriptional repression and recruitment of DNMTs did not immediately result in DNA methylation, as bisulfite sequencing did not detect methylated CpGs at TK promoter sequences 4 days after GAL4BD-LSH and the reporter plasmid were cotransfected into cells (see Fig. S3 in the supplemental material). These observations are consistent with previous reports that DNA methylation follows rather than precedes transcriptional silencing established by negative transcriptional regulators, including chromatin-related mechanisms (23, 40). Thus, although LSH protein serves as a scaffold for assembly of the DNMT/HDAC complex, the primary function of the LSH complex may not be to methylate DNA but to establish deacetylated inactive chromatin. Transcriptional repression caused by deacetylation of histone tails by LSH-interacting HDAC1/2 can be viewed as an initial and, perhaps, reversible step in LSH-mediated heterochromatin formation. A longer-term association of LSH with specific loci and a persistently high local concentration of DNMT1 and DNMT3B may result in methylation of CpGs at these loci. The kinetics of DNA

methylation induced by the LSH-associated complex requires a more detailed investigation. Nevertheless, experiments with unmethylated episomal plasmids capable of replicating in mammalian cells indicate that Lsh-facilitated DNA methylation was observed weeks after the plasmids were introduced into these cells (45). Taken together, these data suggest that LSH cooperates with DNMTs and HDAC1/2 to act as a general transcriptional repressor in mammalian cells.

In mouse cells, Lsh as well as DNMT1 and DNMT3B are required for DNA methylation of pericentromeric major satellite repeats. Yan et al. have reported that in Lsh-deficient mouse embryonic fibroblasts, H3K4 dimethylation, a histone modification associated with transcriptionally permissive chromatin, accumulates at normally heterochromatic, Lsh-bound and H3K4-depleted pericentromeric sequences (43). Our experiments provide a mechanistic explanation for the observed increase of histone acetylation and other positive histone modifications at pericentric heterochromatin and other loci in $Lsh^{-/-}$ cells (41, 43).

It is still unclear whether and how LSH and/or the LSH-associated DNMT complex is recruited to chromatin to establish transcriptionally silenced chromatin and DNA methylation at specific loci. One could also envision that LSH continuously scans chromatin for regions that challenge the processivity of DNMT enzymes. How LSH recruitment is related to functional states of chromatin and what makes certain loci preferred targets for LSH would be intriguing questions to answer. Based on observations in plants that DNA methylation of transposable elements and related tandem repeats is dependent on small interfering RNA and DDM1, it has been suggested that small interfering RNA may guide DDM1 to establish heterochromatin at specific genomic locations (19). However, this proposed model still awaits experimental proof in both plants and in animal cells.

Thus far our preliminary experiments indicate that LSH binds to DNA and to linker DNA regions of reconstituted chromatin more efficiently than either DNMT3B or DNMT1 (data not shown) (27, 31). It is possible to envision that the binding properties of the LSH complex as a whole are determined by the additive affinities of individual proteins, i.e., LSH, DNMTs, and HDACs, for DNA and/or histones. In vitro reconstitution of the LSH complex and further investigation of how this complex interacts with nucleosomal DNA in vitro and in vivo may provide interesting insights.

ACKNOWLEDGMENTS

We thank Ken Sawin, Adrian Bird, Maria Vogelauer, and members of the Stancheva lab for reading the manuscript and helpful suggestions. We appreciate the help from colleagues who provided essential reagents.

This work was supported by a CRUK senior fellowship to Irina Stancheva, the EMBO Young Investigators Programme, and EU Network of Excellence "The Epigenome." Kevin Myant is a CRUK-funded Ph.D. student.

REFERENCES

1. **Bird, A.** 2002. DNA methylation patterns and epigenetic memory. Genes Dev. **16:**6–21.
2. **Brzeski, J., and A. Jerzmanowski.** 2003. Deficient in DNA methylation 1 (DDM1) defines a novel family of chromatin-remodeling factors. J. Biol. Chem. **278:**823–828.
3. **Chen, T., and E. Li.** 2004. Structure and function of eukaryotic DNA methyltransferases. Curr. Top. Dev. Biol. **60:**55–89.
4. **Chuang, L. S., H. I. Ian, T. W. Koh, H. H. Ng, G. Xu, and B. F. Li.** 1997. Human DNA-(cytosine-5) methyltransferase-PCNA complex as a target for p21WAF1. Science **277:**1996–2000.
5. **Dennis, K., T. Fan, T. Geiman, Q. Yan, and K. Muegge.** 2001. Lsh, a member of the SNF2 family, is required for genome-wide methylation. Genes Dev. **15:**2940–2944.
6. **Egger, G., S. Jeong, S. G. Escobar, C. C. Cortez, T. W. Li, Y. Saito, C. B. Yoo, P. A. Jones, and G. Liang.** 2006. Identification of DNMT1 (DNA methyltransferase 1) hypomorphs in somatic knockouts suggests an essential role for DNMT1 in cell survival. Proc. Natl. Acad. Sci. USA **103:**14080–14085.
7. **Fan, T., J. P. Hagan, S. V. Kozlov, C. L. Stewart, and K. Muegge.** 2005. Lsh controls silencing of the imprinted Cdkn1c gene. Development **132:**635–644.
8. **Flaus, A., D. M. Martin, G. J. Barton, and T. Owen-Hughes.** 2006. Identification of multiple distinct Snf2 subfamilies with conserved structural motifs. Nucleic Acids Res. **34:**2887–2905.
9. **Fuks, F., W. A. Burgers, A. Brehm, L. Hughes-Davies, and T. Kouzarides.** 2000. DNA methyltransferase Dnmt1 associates with histone deacetylase activity. Nat. Genet. **24:**88–91.
10. **Geiman, T. M., U. T. Sankpal, A. K. Robertson, Y. Zhao, and K. D. Robertson.** 2004. DNMT3B interacts with hSNF2H chromatin remodeling enzyme, HDACs 1 and 2, and components of the histone methylation system. Biochem. Biophys. Res. Commun. **318:**544–555.
11. **Gowher, H., C. J. Stockdale, R. Goyal, H. Ferreira, T. Owen-Hughes, and A. Jeltsch.** 2005. De novo methylation of nucleosomal DNA by the mammalian Dnmt1 and Dnmt3A DNA methyltransferases. Biochemistry **44:**9899–9904.
12. **Huang, J., T. Fan, Q. Yan, H. Zhu, S. Fox, H. J. Issaq, L. Best, L. Gangi, D. Munroe, and K. Muegge.** 2004. Lsh, an epigenetic guardian of repetitive elements. Nucleic Acids Res. **32:**5019–5028.
13. **Ishizuka, T., and M. A. Lazar.** 2003. The N-CoR/histone deacetylase 3 complex is required for repression by thyroid hormone receptor. Mol. Cell. Biol. **23:**5122–5131.
14. **Jackson-Grusby, L., C. Beard, R. Possemato, M. Tudor, D. Fambrough, G. Csankovszki, J. Dausman, P. Lee, C. Wilson, E. Lander, and R. Jaenisch.** 2001. Loss of genomic methylation causes p53-dependent apoptosis and epigenetic deregulation. Nat. Genet. **27:**31–39.
15. **Jair, K. W., K. E. Bachman, H. Suzuki, A. H. Ting, I. Rhee, R. W. Yen, S. B. Baylin, and K. E. Schuebel.** 2006. De novo CpG island methylation in human cancer cells. Cancer Res. **66:**682–692.
16. **Jeddeloh, J. A., T. L. Stokes, and E. J. Richards.** 1999. Maintenance of genomic methylation requires a SWI2/SNF2-like protein. Nat. Genet. **22:**94–97.
17. **Kim, G. D., J. Ni, N. Kelesoglu, R. J. Roberts, and S. Pradhan.** 2002. Co-operation and communication between the human maintenance and de novo DNA (cytosine-5) methyltransferases. EMBO J. **21:**4183–4195.
18. **Li, E., T. H. Bestor, and R. Jaenisch.** 1992. Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. Cell **69:**915–926.
19. **Lippman, Z., A. V. Gendrel, M. Black, M. W. Vaughn, N. Dedhia, W. R. McCombie, K. Lavine, V. Mittal, B. May, K. D. Kasschau, J. C. Carrington, R. W. Doerge, V. Colot, and R. Martienssen.** 2004. Role of transposable elements in heterochromatin and epigenetic control. Nature **430:**471–476.
20. **Meehan, R. R., S. Pennings, and I. Stancheva.** 2001. Lashings of DNA methylation, forkfuls of chromatin remodeling. Genes Dev. **15:**3231–3236.
21. **Nan, X., F. J. Campoy, and A. Bird.** 1997. MeCP2 is a transcriptional repressor with abundant binding sites in genomic chromatin. Cell **88:**471–481.
22. **Narlikar, G. J., H. Y. Fan, and R. E. Kingston.** 2002. Cooperation between complexes that regulate chromatin structure and transcription. Cell **108:**475–487.
23. **Ohm, J. E., K. M. McGarvey, X. Yu, L. Cheng, K. E. Schuebel, L. Cope, H. P. Mohammad, W. Chen, V. C. Daniel, W. Yu, D. M. Berman, T. Jenuwein, K. Pruitt, S. J. Sharkis, D. N. Watkins, J. G. Herman, and S. B. Baylin.** 2007. A stem cell-like chromatin pattern may predispose tumor suppressor genes to DNA hypermethylation and heritable silencing. Nat. Genet. **39:**237–242.
24. **Okano, M., D. W. Bell, D. A. Haber, and E. Li.** 1999. DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. Cell **99:**247–257.
25. **Okuwaki, M., and A. Verreault.** 2004. Maintenance DNA methylation of nucleosome core particles. J. Biol. Chem. **279:**2904–2912.
26. **Phelan, M. L., S. Sif, G. J. Narlikar, and R. E. Kingston.** 1999. Reconstitution of a core chromatin remodeling complex from SWI/SNF subunits. Mol. Cell **3:**247–253.
27. **Qiu, C., K. Sawada, X. Zhang, and X. Cheng.** 2002. The PWWP domain of mammalian DNA methyltransferase Dnmt3b defines a new family of DNA-binding folds. Nat. Struct. Biol. **9:**217–224.
28. **Raabe, E. H., L. Abdurrahman, G. Behbehani, and R. J. Arceci.** 2001. An SNF2 factor involved in mammalian development and cellular proliferation. Dev. Dyn. **221:**92–105.
29. **Rhee, I., K. E. Bachman, B. H. Park, K. W. Jair, R. W. Yen, K. E. Schuebel, H. Cui, A. P. Feinberg, C. Lengauer, K. W. Kinzler, S. B. Baylin, and B. Vogelstein.** 2002. DNMT1 and DNMT3b cooperate to silence genes in human cancer cells. Nature **416:**552–556.
30. **Rhee, I., K. W. Jair, R. W. Yen, C. Lengauer, J. G. Herman, K. W. Kinzler,**

B. Vogelstein, S. B. Baylin, and K. E. Schuebel. 2000. CpG methylation is maintained in human cancer cells lacking DNMT1. Nature **404:**1003–1007.
31. **Robertson, A. K., T. M. Geiman, U. T. Sankpal, G. L. Hager, and K. D. Robertson.** 2004. Effects of chromatin structure on the enzymatic and DNA binding functions of DNA methyltransferases DNMT1 and Dnmt3a in vitro. Biochem. Biophys. Res. Commun. **322:**110–118.
32. **Robertson, K. D., S. Ait-Si-Ali, T. Yokochi, P. A. Wade, P. L. Jones, and A. P. Wolffe.** 2000. DNMT1 forms a complex with Rb, E2F1 and HDAC1 and represses transcription from E2F-responsive promoters. Nat. Genet. **25:**338–342.
33. **Rountree, M. R., K. E. Bachman, and S. B. Baylin.** 2000. DNMT1 binds HDAC2 and a new co-repressor, DMAP1, to form a complex at replication foci. Nat. Genet. **25:**269–277.
34. **Schermelleh, L., F. Spada, H. P. Easwaran, K. Zolghadr, J. B. Margot, M. C. Cardoso, and H. Leonhardt.** 2005. Trapped in action: direct visualization of DNA methyltransferase activity in living cells. Nat. Methods **2:**751–756.
35. **Siegel, L. M., and K. J. Monty.** 1965. Determination of molecular weights and frictional ratios of macromolecules in impure systems: aggregation of urease. Biochem. Biophys. Res. Commun. **19:**494–499.
36. **Sif, S., A. J. Saurin, A. N. Imbalzano, and R. E. Kingston.** 2001. Purification and characterization of mSin3A-containing Brg1 and hBrm chromatin remodeling complexes. Genes Dev. **15:**603–618.
37. **Spada, F., A. Haemmer, D. Kuch, U. Rothbauer, L. Schermelleh, E. Kremmer, T. Carell, G. Langst, and H. Leonhardt.** 2007. DNMT1 but not its interaction with the replication machinery is required for maintenance of DNA methylation in human cells. J. Cell Biol. **176:**565–571.
38. **Suetake, I., J. Miyazaki, C. Murakami, H. Takeshima, and S. Tajima.** 2003. Distinct enzymatic properties of recombinant mouse DNA methyltransferases Dnmt3a and Dnmt3b. J. Biochem. (Tokyo) **133:**737–744.
39. **Sun, L. Q., D. W. Lee, Q. Zhang, W. Xiao, E. H. Raabe, A. Meeker, D. Miao, D. L. Huso, and R. J. Arceci.** 2004. Growth retardation and premature aging phenotypes in mice with disruption of the SNF2-like gene, PASG. Genes Dev. **18:**1035–1046.
40. **Widschwendter, M., H. Fiegl, D. Egle, E. Mueller-Holzner, G. Spizzo, C. Marth, D. J. Weisenberger, M. Campan, J. Young, I. Jacobs, and P. W. Laird.** 2007. Epigenetic stem cell signature in cancer. Nat. Genet. **39:**157–158.
41. **Xi, S., H. Zhu, H. Xu, A. Schmidtmann, T. M. Geiman, and K. Muegge.** 2007. Lsh controls Hox gene silencing during development. Proc. Natl. Acad. Sci. USA **104:**14366–14371.
42. **Yan, Q., E. Cho, S. Lockett, and K. Muegge.** 2003. Association of Lsh, a regulator of DNA methylation, with pericentromeric heterochromatin is dependent on intact heterochromatin. Mol. Cell. Biol. **23:**8416–8428.
43. **Yan, Q., J. Huang, T. Fan, H. Zhu, and K. Muegge.** 2003. Lsh, a modulator of CpG methylation, is crucial for normal histone methylation. EMBO J. **22:**5154–5162.
44. **Zhang, X., and H. Bremer.** 1995. Control of the Escherichia coli rrnB P1 promoter strength by ppGpp. J. Biol. Chem. **270:**11181–11189.
45. **Zhu, H., T. M. Geiman, S. Xi, Q. Jiang, A. Schmidtmann, T. Chen, E. Li, and K. Muegge.** 2006. Lsh is involved in de novo methylation of DNA. EMBO J. **25:**335–345.

6. Reference List

Aalfs, J.D., Narlikar, G.J., and Kingston, R.E. (2001). Functional differences between the human ATP-dependent nucleosome remodeling proteins BRG1 and SNF2H. The Journal of Biological Chemistry *276*, 34270-34278.

Aapola, U., Lyle, R., Krohn, K., Antonarakis, S.E., and Peterson, P. (2001). Isolation and initial characterization of the mouse Dnmt3l gene. Cytogenetics and Cell Genetics *92*, 122-126.

Abrams, E., Neigeborn, L., and Carlson, M. (1986). Molecular analysis of SNF2 and SNF5, genes required for expression of glucose-repressible genes in Saccharomyces cerevisiae. Molecular and Cellular Biology *6*, 3643-3651.

Anderson, J.D., Thastrom, A., and Widom, J. (2002). Spontaneous Access of Proteins to Buried Nucleosomal DNA Target Sites Occurs via a Mechanism That Is Distinct from Nucleosome Translocation. Molecular and Cellular Biology *22,* 7147-7157.

Antequera, F., Boyes, J., and Bird, A. (1990). High levels of de novo methylation and altered chromatin structure at CpG islands in cell lines. Cell *62*, 503-514.

Auble, D.T., Hansen, K.E., Mueller, C.G., Lane, W.S., Thorner, J., and Hahn, S. (1994). Mot1, a global repressor of RNA polymerase II transcription, inhibits TBP binding to DNA by an ATP-dependent mechanism. Genes & Development *8*, 1920-1934.

Auble, D.T., and Steggerda, S.M. (1999). Testing for DNA tracking by MOT1, a SNF2/SWI2 protein family member. Molecular and Cellular Biology *19*, 412-423.

Bachman, K.E., Rountree, M.R., and Baylin, S.B. (2001). Dnmt3a and Dnmt3b are transcriptional repressors that exhibit unique localization properties to heterochromatin. The Journal of Biological Chemistry *276*, 32282-32287.

Bacolla, A., Pradhan, S., Roberts, R.J., and Wells, R.D. (1999). Recombinant human DNA (cytosine-5) methyltransferase. II. Steady-state kinetics reveal allosteric

activation by methylated dna. The Journal of Biological Chemistry *274*, 33011-33019.

Bagchi, A., Papazoglu, C., Wu, Y., Capurso, D., Brodt, M., Francis, D., Bredel, M., Vogel, H., and Mills, A.A. (2007). CHD5 is a tumor suppressor at human 1p36. Cell *128*, 459-475.

Bannister, A.J., Zegerman, P., Partridge, J.F., Miska, E.A., Thomas, J.O., Allshire, R.C., and Kouzarides, T. (2001). Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. Nature *410*, 120-124.

Bao, Y., and Shen, X. (2007). INO80 subfamily of chromatin remodeling complexes. Mutation Research *618*, 18-29.

Bestor, T., Laudano, A., Mattaliano, R., and Ingram, V. (1988). Cloning and sequencing of a cDNA encoding DNA methyltransferase of mouse cells. The carboxyl-terminal domain of the mammalian enzymes is related to bacterial restriction methyltransferases. Journal of Molecular Biology *203*, 971-983.

Bestor, T.H. (2000). The DNA methyltransferases of mammals. Human Molecular Genetics *9*, 2395-2402.

Biegel, J.A., Zhou, J.Y., Rorke, L.B., Stenstrom, C., Wainwright, L.M., and Fogelgren, B. (1999). Germ-line and acquired mutations of INI1 in atypical teratoid and rhabdoid tumors. Cancer Research *59*, 74-79.

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

Bird, A., Taggart, M., Frommer, M., Miller, O.J., and Macleod, D. (1985). A fraction of the mouse genome that is derived from islands of nonmethylated, CpG-rich DNA. Cell *40*, 91-99.

Bird, A.P. (1986). CpG-rich islands and the function of DNA methylation. Nature *321*, 209-213.

Borgione, E., Sturnio, M., Spalletta, A., Angela Lo Giudice, M., Castiglia, L., Galesi, O., Ragusa, A., and Fichera, M. (2003). Mutational analysis of the ATRX gene by

DGGE: a powerful diagnostic approach for the ATRX syndrome. Human Mutation *21*, 529-534.

Bork, P., and Koonin, E.V. (1993). An expanding family of helicases within the 'DEAD/H' superfamily. Nucleic Acids Research *21*, 751-752.

Borsani, G., Tonlorenzi, R., Simmler, M.C., Dandolo, L., Arnaud, D., Capra, V., Grompe, M., Pizzuti, A., Muzny, D., Lawrence, C., *et al.* (1991). Characterization of a murine gene expressed from the inactive X chromosome. Nature *351*, 325-329.

Bostick, M., Kim, J.K., Esteve, P.O., Clark, A., Pradhan, S., and Jacobsen, S.E. (2007). UHRF1 plays a role in maintaining DNA methylation in mammalian cells. Science *317*, 1760-1764.

Bourc'his, D., and Bestor, T.H. (2004). Meiotic catastrophe and retrotransposon reactivation in male germ cells lacking Dnmt3L. Nature *431*, 96-99.

Bowen, N.J., Fujita, N., Kajita, M., and Wade, P.A. (2004). Mi-2/NuRD: multiple complexes for many purposes. Biochimica et Biophysica Acta *1677*, 52-57.

Boyer, L.A., Langer, M.R., Crowley, K.A., Tan, S., Denu, J.M., and Peterson, C.L. (2002). Essential role for the SANT domain in the functioning of multiple chromatin remodeling enzymes. Molecular Cell *10*, 935-942.

Brockdorff, N., Ashworth, A., Kay, G.F., Cooper, P., Smith, S., McCabe, V.M., Norris, D.P., Penny, G.D., Patel, D., and Rastan, S. (1991). Conservation of position and exclusive expression of mouse Xist from the inactive X chromosome. Nature *351*, 329-331.

Brockdorff, N., Ashworth, A., Kay, G.F., McCabe, V.M., Norris, D.P., Cooper, P.J., Swift, S., and Rastan, S. (1992). The product of the mouse Xist gene is a 15 kb inactive X-specific transcript containing no conserved ORF and located in the nucleus. Cell *71*, 515-526.

Brockdorff, N., and Duthie, S.M. (1998). X chromosome inactivation and the Xist gene. Cell Mol Life Science *54*, 104-112.

Brown, C.J., Lafreniere, R.G., Powers, V.E., Sebastio, G., Ballabio, A., Pettigrew, A.L., Ledbetter, D.H., Levy, E., Craig, I.W., and Willard, H.F. (1991). Localization

of the X inactivation centre on the human X chromosome in Xq13. Nature *349*, 82-84.

Brzeski, J., and Jerzmanowski, A. (2003). Deficient in DNA methylation 1 (DDM1) defines a novel family of chromatin-remodeling factors. The Journal of Biological Chemistry *278*, 823-828.

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

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

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.

Cardoso, C., Timsit, S., Villard, L., Khrestchatisky, M., Fontes, M., and Colleaux, L. (1998). Specific interaction between the XNP/ATR-X gene product and the SET domain of the human EZH2 protein. Human Molecular Genetics *7*, 679-684.

Chedin, F., Lieber, M.R., and Hsieh, C.L. (2002). The DNA methyltransferase-like protein DNMT3L stimulates de novo methylation by Dnmt3a. Proceedings of the National Academy of Sciences of the United States of America *99*, 16916-16921.

Chen, R.Z., Pettersson, U., Beard, C., Jackson-Grusby, L., and Jaenisch, R. (1998). DNA hypomethylation leads to elevated mutation rates. Nature *395*, 89-93.

Chen, T., Tsujimoto, N., and Li, E. (2004). The PWWP domain of Dnmt3a and Dnmt3b is required for directing DNA methylation to the major satellite repeats at pericentric heterochromatin. Molecular and Cellular Biology *24*, 9048-9058.

Chen, T., Ueda, Y., Dodge, J.E., Wang, Z., and Li, E. (2003). Establishment and maintenance of genomic methylation patterns in mouse embryonic stem cells by Dnmt3a and Dnmt3b. Molecular and Cellular Biology *23*, 5594-5605.

Chen, Z.X., Mann, J.R., Hsieh, C.L., Riggs, A.D., and Chedin, F. (2005). Physical and functional interactions between the human DNMT3L protein and members of the de novo methyltransferase family. Journal of Cellular Biochemistry *95*, 902-917.

Cheng, S.W., Davies, K.P., Yung, E., Beltran, R.J., Yu, J., and Kalpana, G.V. (1999). c-MYC interacts with INI1/hSNF5 and requires the SWI/SNF complex for transactivation function. Nature Genetics *22*, 102-105.

Cheng, X., and Roberts, R.J. (2001). AdoMet-dependent methylation, DNA methyltransferases and base flipping. Nucleic Acids Research *29,* 3784-3795.

Chi, T. (2004). A BAF-centred view of the immune system. Nature Reviews Immunology *4*, 965-977.

Chmiel, N.H., Rio, D.C., and Doudna, J.A. (2006). Distinct contributions of KH domains to substrate binding affinity of Drosophila P-element somatic inhibitor protein. RNA *12*, 283-291.

Chuang, L.S., Ian, H.I., Koh, T.W., Ng, H.H., Xu, G., and Li, B.F. (1997). Human DNA-(cytosine-5) methyltransferase-PCNA complex as a target for p21WAF1. Science *227,* 1996-2000.

Clapier, C.R., Langst, 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.

Corona, D.F., and Tamkun, J.W. (2004). Multiple roles for ISWI in transcription, chromosome organization and DNA replication. Biochimica et Biophysica Acta *1677*, 113-119.

Corona, D.F.V., Längst, G., Clapier, C.R., Bonte, E.J., Ferrari, S., Tamkun, J.W., and Becker, P.B. (1999). ISWI Is an ATP-Dependent Nucleosome Remodeling Factor. Molecular Cell *3*, 239-245.

Cote, J., Quinn, J., Workman, J.L., and Peterson, C.L. (1994). Stimulation of GAL4 derivative binding to nucleosomal DNA by the yeast SWI/SNF complex. Science *265*, 53-60.

Csankovszki, G., Nagy, A., and Jaenisch, R. (2001). Synergism of Xist RNA, DNA methylation, and histone hypoacetylation in maintaining X chromosome inactivation. The Journal of Cell Biology *153*, 773-784.

Davis, J.L., Kunisawa, R., and Thorner, J. (1992). A presumptive helicase (MOT1 gene product) affects gene expression and is required for viability in the yeast Saccharomyces cerevisiae. Molecular and Cellular Biology *12*, 1879-1892.

Deb-Rinker, P., Ly, D., Jezierski, A., Sikorska, M., and Walker, P.R. (2005). Sequential DNA methylation of the Nanog and Oct-4 upstream regions in human NT2 cells during neuronal differentiation. The Journal of Biological Chemistry *280*, 6257-6260.

Decristofaro, M.F., Betz, B.L., Rorie, C.J., Reisman, D.N., Wang, W., and Weissman, B.E. (2001). Characterization of SWI/SNF protein expression in human breast cancer cell lines and other malignancies. Journal of Cellular Physiology *186*, 136-145.

Denissenko, M.F., Chen, J.X., Tang, M.S., and Pfeifer, G.P. (1997). Cytosine methylation determines hot spots of DNA damage in the human P53 gene. Proceedings of the National Academy of Sciences of the United States of America *94*, 3893-3898.

Denissenko, M.F., Pao, A., Tang, M., and Pfeifer, G.P. (1996). Preferential formation of benzo[a]pyrene adducts at lung cancer mutational hotspots in P53. Science *274*, 430-432.

Dennis, K., Fan, T., Geiman, T., Yan, Q., and Muegge, K. (2001). Lsh, a member of the SNF2 family, is required for genome-wide methylation. Genes & Development *15*, 2940-2944.

Deuring, R., Fanti, L., Armstrong, J.A., Sarte, M., Papoulas, O., Prestel, M., Daubresse, G., Verardo, M., Moseley, S.L., Berloco, M., *et al.* (2000). The ISWI chromatin-remodeling protein is required for gene expression and the maintenance of higher order chromatin structure in vivo. Molecular Cell *5*, 355-365.

Dignam, J.D., Lebovitz, R.M., and Roeder, R.G. (1983). Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. Nucleic Acids Research *11*, 1475-1489.

Dillingham, M.S., Wigley, D.B., and Webb, M.R. (2002). Direct measurement of single-stranded DNA translocation by PcrA helicase using the fluorescent base analogue 2-aminopurine. Biochemistry *41*, 643-651.

Ding, Z., Liang, J., Lu, Y., Yu, Q., Songyang, Z., Lin, S.Y., and Mills, G.B. (2006). A retrovirus-based protein complementation assay screen reveals functional AKT1-binding partners. Proceedings of the National Academy of Sciences of the United States of America *103*, 15014-15019.

Dong, A., Yoder, J.A., Zhang, X., Zhou, L., Bestor, T.H., and Cheng, X. (2001). Structure of human DNMT2, an enigmatic DNA methyltransferase homolog that displays denaturant-resistant binding to DNA. Nucleic Acids Research *29*, 439-448.

Dorigo, B., Schalch, T., Bystricky, K., and Richmond, T.J. (2003). Chromatin fiber folding: requirement for the histone H4 N-terminal tail. Journal of Molecular Biology *327*, 85-96.

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.

Durr, H., Korner, C., Muller, 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.

Eberharter, A., Langst, G., and Becker, P.B. (2004). A nucleosome sliding assay for chromatin remodeling factors. Methods in Enzymology *377*, 344-353.

Edayathumangalam, R.S., Weyermann, P., Dervan, P.B., Gottesfeld, J.M., and Luger, K. (2005). Nucleosomes in solution exist as a mixture of twist-defect states. Journal of Molecular Biology *345*, 103-114.

Eden, A., Gaudet, F., Waghmare, A., and Jaenisch, R. (2003). Chromosomal Instability and Tumors Promoted by DNA Hypomethylation. Science *300,* 455-456.

Egger, G., Jeong, S., Escobar, S.G., Cortez, C.C., Li, T.W., Saito, Y., Yoo, C.B., Jones, P.A., and Liang, G. (2006). Identification of DNMT1 (DNA methyltransferase 1) hypomorphs in somatic knockouts suggests an essential role for DNMT1 in cell survival. Proceedings of the National Academy of Sciences of the United States of America *103*, 14080-14085.

Ehrlich, M., Sanchez, C., Shao, C., Nishiyama, R., Kehrl, J., Kuick, R., Kubota, T., and Hanash, S.M. (2008). ICF, an immunodeficiency syndrome: DNA methyltransferase 3B involvement, chromosome anomalies, and gene dysregulation. Autoimmunity *41,* 253 - 271.

Eisen, J.A., Sweder, K.S., and Hanawalt, P.C. (1995). Evolution of the SNF2 family of proteins: subfamilies with distinct sequences and functions. Nucleic Acids Research *23*, 2715-2723.

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.

Esteller, M., Fraga, M.F., Guo, M., Garcia-Foncillas, J., Hedenfalk, I., Godwin, A.K., Trojan, J., Vaurs-Barriere, C., Bignon, Y.J., Ramus, S., *et al.* (2001). DNA methylation patterns in hereditary human cancers mimic sporadic tumorigenesis. Human Molecular Genetics *10*, 3001-3007.

Esteller, M. Herman, J.G. (2002). Cancer as an epigenetic disease: DNA methylation and chromatin alterations in human tumours. Journal of Pathology *196,* 1-7.

Fan, H.Y., Trotter, K.W., Archer, T.K., and Kingston, R.E. (2005a). Swapping function of two chromatin remodeling complexes. Molecular Cell *17*, 805-815.

Fan, T., Hagan, J.P., Kozlov, S.V., Stewart, C.L., and Muegge, K. (2005b). Lsh controls silencing of the imprinted Cdkn1c gene. Development *132*, 635-644.

Fan, T., Yan, Q., Huang, J., Austin, S., Cho, E., Ferris, D., and Muegge, K. (2003). Lsh-deficient murine embryonal fibroblasts show reduced proliferation with signs of abnormal mitosis. Cancer Research *63*, 4677-4683.

Feinberg, A.P. (1988). Alterations in DNA methylation in colorectal polyps and cancer. Progress in Clinical and Biological Research *279*, 309-317.

Feinberg, A.P., and Vogelstein, B. (1983). Hypomethylation distinguishes genes of some human cancers from their normal counterparts. Nature *301*, 89-92.

Feldman, N., Gerson, A., Fang, J., Li, E., Zhang, Y., Shinkai, Y., Cedar, H., and Bergman, Y. (2006). G9a-mediated irreversible epigenetic inactivation of Oct-3/4 during early embryogenesis. Nature Cell Biology *8*, 188-194.

Filion, G.J., Zhenilo, S., Salozhin, S., Yamada, D., Prokhortchouk, E., and Defossez, P.A. (2006). A family of human zinc finger proteins that bind methylated DNA and repress transcription. Molecular and Cellular Biology *26*, 169-181.

Fischle, W., Tseng, B.S., Dormann, H.L., Ueberheide, B.M., Garcia, B.A., Shabanowitz, J., Hunt, D.F., Funabiki, H., and Allis, C.D. (2005). Regulation of HP1-chromatin binding by histone H3 methylation and phosphorylation. Nature *438*, 1116-1122.

Fitzgerald, D.J., DeLuca, C., Berger, I., Gaillard, H., Sigrist, R., Schimmele, K., and Richmond, T.J. (2004). Reaction cycle of the yeast Isw2 chromatin remodeling complex. The EMBO Journal *23*, 3836-3843.

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. (2004). Mechanisms for ATP-dependent chromatin remodelling: farewell to the tuna-can octamer? Current Opinion in Genetics & Development *14*, 165-173.

Flaus, A., and Richmond, T.J. (1998). Positioning and stability of nucleosomes on MMTV 3'LTR sequences. Journal of Molecular Biology *275*, 427-441.

Fraga, M.F., Ballestar, E., Paz, M.F., Ropero, S., Setien, F., Ballestar, M.L., Heine-Suner, D., Cigudosa, J.C., Urioste, M., Benitez, J., *et al.* (2005). Epigenetic differences arise during the lifetime of monozygotic twins. Proceedings of the National Academy of Sciences of the United States of America *102*, 10604-10609.

Fuks, F., Burgers, W.A., Brehm, A., Hughes-Davies, L., and Kouzarides, T. (2000). DNA methyltransferase Dnmt1 associates with histone deacetylase activity. Nature genetics *24*, 88-91.

Fuks, F., Burgers, W.A., Godin, N., Kasai, M., and Kouzarides, T. (2001). Dnmt3a binds deacetylases and is recruited by a sequence-specific repressor to silence transcription. The EMBO Journal *20*, 2536-2544.

Fuks, F., Hurd, P.J., Deplus, R., and Kouzarides, T. (2003). The DNA methyltransferases associate with HP1 and the SUV39H1 histone methyltransferase. Nucleic Acids Research *31*, 2305-2312.

Gal-Yam, E.N., Jeong, S., Tanay, A., Egger, G., Lee, A.S., and Jones, P.A. (2006). Constitutive nucleosome depletion and ordered factor assembly at the GRP78 promoter revealed by single molecule footprinting. PLoS Genetics *2*, e160.

Gama-Sosa, M.A., Slagel, V.A., Trewyn, R.W., Oxenhandler, R., Kuo, K.C., Gehrke, C.W., and Ehrlich, M. (1983). The 5-methylcytosine content of DNA from human tumors. Nucleic Acids Research *11*, 6883-6894.

Gardiner-Garden, M., and Frommer, M. (1987). CpG islands in vertebrate genomes. Journal of Molecular Biology *196*, 261-282.

Gaudet, F., Hodgson, J.G., Eden, A., Jackson-Grusby, L., Dausman, J., Gray, J.W., Leonhardt, H., and Jaenisch, R. (2003). Induction of Tumors in Mice by Genomic Hypomethylation. Science *300,* 489-492.

Gavin, I., Horn, P.J., and Peterson, C.L. (2001). SWI/SNF chromatin remodeling requires changes in DNA topology. Molecular Cell *7*, 97-104.

Ge, Y.Z., Pu, M.T., Gowher, H., Wu, H.P., Ding, J.P., Jeltsch, A., and Xu, G.L. (2004). Chromatin targeting of de novo DNA methyltransferases by the PWWP domain. The Journal of Biological Chemistry *279*, 25447-25454.

Geiman, T.M., Durum, S.K., and Muegge, K. (1998). Characterization of gene expression, genomic structure, and chromosomal localization of Hells (Lsh). Genomics *54*, 477-483.

Geiman, T.M., and Muegge, K. (2000). Lsh, an SNF2/helicase family member, is required for proliferation of mature T lymphocytes. Proceedings of the National Academy of Science USA *97*, 4772-4777.

Geiman, T.M., Sankpal, U.T., Robertson, A.K., Zhao, Y., Zhao, Y., and Robertson, K.D. (2004a). DNMT3B interacts with hSNF2H chromatin remodeling enzyme, HDACs 1 and 2, and components of the histone methylation system. Biochemical and Biophysical Research Communications *318*, 544-555.

Geiman, T.M., Sankpal, U.T., Robertson, A.K., Zhao, Y., Zhao, Y., and Robertson, K.D. (2004b). DNMT3B interacts with hSNF2H chromatin remodeling enzyme, HDACs 1 and 2, and components of the histone methylation system. Biochemical and Biophysical Research Communications *318*, 544-555.

Geiman, T.M., Tessarollo, L., Anver, M.R., Kopp, J.B., Ward, J.M., and Muegge, K. (2001). Lsh, a SNF2 family member, is required for normal murine development. Biochimica et Biophysica Acta *1526*, 211-220.

Gendrel, A.V., Lippman, Z., Yordan, C., Colot, V., and Martienssen, R.A. (2002). Dependence of heterochromatic histone H3 methylation patterns on the Arabidopsis gene DDM1. Science *297*, 1871-1873.

Gibbons, R. (2006). Alpha thalassaemia-mental retardation, X linked. Orphanet Journal of Rare Diseases *1*, 15.

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., Pellagatti, A., Garrick, D., Wood, W.G., Malik, N., Ayyub, H., Langford, C., Boultwood, J., Wainscoat, J.S., and Higgs, D.R. (2003). Identification of acquired somatic mutations in the gene encoding chromatin-remodeling factor ATRX in the alpha-thalassemia myelodysplasia syndrome (ATMDS). Nature genetics *34*, 446-449.

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 alpha-thalassemia (ATR-X syndrome). Cell *80*, 837-845.

Gilbert, D.M. (2002). Replication timing and transcriptional control: beyond cause and effect. Current Opinion in Cell Biology *14*, 377-383.

Goldmark, J.P., Fazzio, T.G., Estep, P.W., Church, G.M., and Tsukiyama, T. (2000). The Isw2 chromatin remodeling complex represses early meiotic genes upon recruitment by Ume6p. Cell *103*, 423-433.

Goll, M.G., Kirpekar, F., Maggert, K.A., Yoder, J.A., Hsieh, C.-L., Zhang, X., Golic, K.G., Jacobsen, S.E., and Bestor, T.H. (2006). Methylation of tRNAAsp by the DNA Methyltransferase Homolog Dnmt2. Science *311,* 395-398.

Gowher, H., and Jeltsch, A. (2001). Enzymatic properties of recombinant Dnmt3a DNA methyltransferase from mouse: the enzyme modifies DNA in a non-processive manner and also methylates non-CpG [correction of non-CpA] sites. Journal of Molecular Biology *309*, 1201-1208.

Gowher, H., and Jeltsch, A. (2002). Molecular enzymology of the catalytic domains of the Dnmt3a and Dnmt3b DNA methyltransferases. The Journal of Biological Chemistry *277*, 20409-20414.

Gowher, H., Leismann, O., and Jeltsch, A. (2000). DNA of Drosophila melanogaster contains 5-methylcytosine. The EMBO Journal *19*, 6918-6923.

Gowher, H., Liebert, K., Hermann, A., Xu, G., and Jeltsch, A. (2005a). Mechanism of stimulation of catalytic activity of Dnmt3A and Dnmt3B DNA-(cytosine-C5)-methyltransferases by Dnmt3L. The Journal of Biological Chemistry *280*, 13341-13348.

Gowher, H., Stockdale, C.J., Goyal, R., Ferreira, H., Owen-Hughes, T., and Jeltsch, A. (2005b). De novo methylation of nucleosomal DNA by the mammalian Dnmt1 and Dnmt3A DNA methyltransferases. Biochemistry *44*, 9899-9904.

Grewal, S.I.S., and Jia, S. (2007). Heterochromatin revisited. Nature Reviews Genetics *8*, 35-46.

Guccione, E., Bassi, C., Casadio, F., Martinato, F., Cesaroni, M., Schuchlautz, H., Luscher, B., and Amati, B. (2007). Methylation of histone H3R2 by PRMT6 and H3K4 by an MLL complex are mutually exclusive. Nature *449*, 933-937.

Guerrini, R., Shanahan, J.L., Carrozzo, R., Bonanni, P., Higgs, D.R., and Gibbons, R.J. (2000). A nonsense mutation of the ATRX gene causing mild mental retardation and epilepsy. Annals of Neurology *47*, 117-121.

Hamiche, A., Kang, J.-G., Dennis, C., Xiao, H., and Wu, C. (2001). Histone tails modulate nucleosome mobility and regulate ATP-dependent nucleosome sliding by NURF. Proceedings of the National Academy of Sciences USA *98,* 14316-14321.

Hamiche, A., Sandaltzopoulos, R., Gdula, D.A., and Wu, C. (1999). ATP-dependent histone octamer sliding mediated by the chromatin remodeling complex NURF. Cell *97*, 833-842.

Handa, V., and Jeltsch, A. (2005). Profound flanking sequence preference of Dnmt3a and Dnmt3b mammalian DNA methyltransferases shape the human epigenome. Journal of Molecular Biology *348*, 1103-1112.

Hansen, R.S., Stoger, R., Wijmenga, C., Stanek, A.M., Canfield, T.K., Luo, P., Matarazzo, M.R., D'Esposito, M., Feil, R., Gimelli, G., *et al.* (2000). Escape from gene silencing in ICF syndrome: evidence for advanced replication time as a major determinant. Human Molecular Genetics *9*, 2575-2587.

Hassan, A.H., Prochasson, P., Neely, K.E., Galasinski, S.C., Chandy, M., Carrozza, M.J., and Workman, J.L. (2002). Function and selectivity of bromodomains in anchoring chromatin-modifying complexes to promoter nucleosomes. Cell *111*, 369-379.

Hata, K., Okano, M., Lei, H., and Li, E. (2002). Dnmt3L cooperates with the Dnmt3 family of de novo DNA methyltransferases to establish maternal imprints in mice. Development *129*, 1983-1993.

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.

He, X., Fan, H.Y., Narlikar, G.J., and Kingston, R.E. (2006). Human ACF1 alters the remodeling strategy of SNF2h. The Journal of Biological Chemistry *281*, 28636-28647.

Heitz, E. (1928). Das heterochromatin der moose. I Jahrb Wiss Botanik *69*, 762-818.

Herman, J.G., Umar, A., Polyak, K., Graff, J.R., Ahuja, N., Issa, J.P., Markowitz, S., Willson, J.K., Hamilton, S.R., Kinzler, K.W., *et al.* (1998). Incidence and functional consequences of hMLH1 promoter hypermethylation in colorectal carcinoma. Proceedings of the National Academy of Sciences of the United States of America *95*, 6870-6875.

Hermann, A., Gowher, H., and Jeltsch, A. (2004). Biochemistry and biology of mammalian DNA methyltransferases. Cellular and Molecular Life Sciences *61*, 2571-2587.

Hermann, A., Schmitt, S., and Jeltsch, A. (2003). The human Dnmt2 has residual DNA-(cytosine-C5) methyltransferase activity. The Journal of Biological Chemistry *278*, 31717-31721.

Hirota, T., Lipp, J.J., Toh, B.H., and Peters, J.M. (2005). Histone H3 serine 10 phosphorylation by Aurora B causes HP1 dissociation from heterochromatin. Nature *438*, 1176-1180.

Hsieh, C.L. (1999). In vivo activity of murine de novo methyltransferases, Dnmt3a and Dnmt3b. Molecular and Cellular Biology *19*, 8211-8218.

Huang, J., Fan, T., Yan, Q., Zhu, H., Fox, S., Issaq, H.J., Best, L., Gangi, L., Munroe, D., and Muegge, K. (2004). Lsh, an epigenetic guardian of repetitive elements. Nucleic Acids Research *32*, 5019-5028.

Hung, M.-S., Karthikeyan, N., Huang, B., Koo, H.-C., Kiger, J., and Shen, C.K.J. (1999). Drosophila proteins related to vertebrate DNA (5-cytosine) methyltransferases. Proceedings of the National Academy of Science USA. *96,* 11940-11945.

Illingworth, R., Kerr, A., Desousa, D., Jorgensen, H., Ellis, P., Stalker, J., Jackson, D., Clee, C., Plumb, R., Rogers, J., *et al.* (2008). A novel CpG island set identifies tissue-specific methylation at developmental gene loci. PLoS Biology *6*, e22.

Ishizuka, T., and Lazar, M.A. (2003). The N-CoR/histone deacetylase 3 complex is required for repression by thyroid hormone receptor. Molecular and Cellular Biology *23*, 5122-5131.

Issa, J.P. (2000). CpG-island methylation in aging and cancer. Current topics in Microbiology and Immunology *249*, 101-118.

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.

Jackson, J.P., Lindroth, A.M., Cao, X., and Jacobsen, S.E. (2002). Control of CpNpG DNA methylation by the KRYPTONITE histone H3 methyltransferase. Nature *416*, 556-560.

Jackson, M., Krassowska, A., Gilbert, N., Chevassut, T., Forrester, L., Ansell, J., and Ramsahoye, B. (2004). Severe global DNA hypomethylation blocks differentiation and induces histone hyperacetylation in embryonic stem cells. Molecular and Cellular Biology *24*, 8862-8871.

Jacobs, S.A., and Khorasanizadeh, S. (2002). Structure of HP1 Chromodomain Bound to a Lysine 9-Methylated Histone H3 Tail. Science *295,* 2080-2083.

Jair, K.W., Bachman, K.E., Suzuki, H., Ting, A.H., Rhee, I., Yen, R.W., Baylin, S.B., and Schuebel, K.E. (2006). De novo CpG island methylation in human cancer cells. Cancer Research *66*, 682-692.

Jarvis, C.D., Geiman, T., Vila-Storm, M.P., Osipovich, O., Akella, U., Candeias, S., Nathan, I., Durum, S.K., and Muegge, K. (1996). A novel putative helicase produced in early murine lymphocytes. Gene *169*, 203-207.

Jaskelioff, M., Gavin, I.M., Peterson, C.L., and Logie, C. (2000). SWI-SNF-mediated nucleosome remodeling: role of histone octamer mobility in the persistence of the remodeled state. Molecular and Cellular Biology *20*, 3058-3068.

Jeddeloh, J.A., Stokes, T.L., and Richards, E.J. (1999). Maintenance of genomic methylation requires a SWI2/SNF2-like protein. Nature Genetics *22*, 94-97.

Jeong-Heon Lee, S.R.L.H.D.G.S. (2004). Histone deacetylase activity is required for embryonic stem cell differentiation. Genesis *38,* 32-38.

Jia, D., Jurkowska, R.Z., Zhang, X., Jeltsch, A., and Cheng, X. (2007). Structure of Dnmt3a bound to Dnmt3L suggests a model for de novo DNA methylation. Nature *449*, 248-251.

Jones, P.A., and Baylin, S.B. (2002). The fundamental role of epigenetic events in cancer. Nature Reviews Genetics *3*, 415-428.

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.

Kakutani, T. (1997). Genetic characterization of late-flowering traits induced by DNA hypomethylation mutation in Arabidopsis thaliana. Plant Journal *12*, 1447-1451.

Kakutani, T., Jeddeloh, J.A., Flowers, S.K., Munakata, K., and Richards, E.J. (1996). Developmental abnormalities and epimutations associated with DNA hypomethylation mutations. Proceedings of the National Academy of Sciences of the United States of America *93*, 12406-12411.

Kakutani, T., Jeddeloh, J.A., and Richards, E.J. (1995). Characterization of an Arabidopsis thaliana DNA hypomethylation mutant. Nucleic Acids Research *23*, 130-137.

Kaneda, M., Okano, M., Hata, K., Sado, T., Tsujimoto, N., Li, E., and Sasaki, H. (2004). Essential role for de novo DNA methyltransferase Dnmt3a in paternal and maternal imprinting. Nature *429*, 900-903.

Kareta, M.S., Botello, Z.M., Ennis, J.J., Chou, C., and Chedin, F. (2006). Reconstitution and mechanism of the stimulation of de novo methylation by human DNMT3L. The Journal of Biological Chemistry *281*, 25893-25902.

Kiianitsa, K., Solinger, J.A., and Heyer, W.-D. (2002). Rad54 Protein Exerts Diverse Modes of ATPase Activity on Duplex DNA Partially and Fully Covered with Rad51 Protein. Journal of Biological Chemistry *277,* 46205-46215.

Kim, G.D., Ni, J., Kelesoglu, N., Roberts, R.J., and Pradhan, S. (2002). Co-operation and communication between the human maintenance and de novo DNA (cytosine-5) methyltransferases. The EMBO Journal *21*, 4183-4195.

Kirmizis, A., Santos-Rosa, H., Penkett, C.J., Singer, M.A., Vermeulen, M., Mann, M., Bahler, J., Green, R.D., and Kouzarides, T. (2007). Arginine methylation at histone H3R2 controls deposition of H3K4 trimethylation. Nature *449*, 928-932.

Klimasauskas, S., Kumar, S., Roberts, R.J., and Cheng, X. (1994). HhaI methyltransferase flips its target base out of the DNA helix. Cell *76*, 357-369.

Klose, R.J., and Bird, A.P. (2004). MeCP2 behaves as an elongated monomer that does not stably associate with the Sin3a chromatin remodeling complex. The Journal of Biological Chemistry *279*, 46490-46496.

Knudson, A.G., Jr. (1971). Mutation and cancer: statistical study of retinoblastoma. Proceedings of the National Academy of Sciences of the United States of America *68*, 820-823.

Kouzarides, T. (2007). Chromatin Modifications and Their Function. Cell *128*, 693-705.

Kumar, S., Cheng, X., Klimasauskas, S., Mi, S., Posfai, J., Roberts, R.J., and Wilson, G.G. (1994). The DNA (cytosine-5) methyltransferases. Nucleic Acids Research *22*, 1-10.

Kunert, N., Marhold, J., Stanke, J., Stach, D., and Lyko, F. (2003). A Dnmt2-like protein mediates DNA methylation in Drosophila. Development *130,* 5083-5090.

Laird, C.D., Pleasant, N.D., Clark, A.D., Sneeden, J.L., Hassan, K.M., Manley, N.C., Vary, J.C., Jr., Morgan, T., Hansen, R.S., and Stoger, R. (2004). Hairpin-bisulfite PCR: assessing epigenetic methylation patterns on complementary strands of individual DNA molecules. Proceedings of the National Academy of Sciences of the United States of America *101*, 204-209.

Lander, E.S., Linton, L.M., Birren, B., Nusbaum, C., Zody, M.C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W., *et al.* (2001). Initial sequencing and analysis of the human genome. Nature *409*, 860-921.

Langst, G., and Becker, P.B. (2001). Nucleosome mobilization and positioning by ISWI-containing chromatin-remodeling factors. Journal of Cell Science *114*, 2561-2568.

Langst, 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.

Laurent, B.C., Treich, I., and Carlson, M. (1993). The yeast SNF2/SWI2 protein has DNA-stimulated ATPase activity required for transcriptional activation. Genes & Development *7*, 583-591.

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.

Le Guezennec, X., Vermeulen, M., Brinkman, A.B., Hoeijmakers, W.A., Cohen, A., Lasonder, E., and Stunnenberg, H.G. (2006). MBD2/NuRD and MBD3/NuRD, two distinct complexes with different biochemical and functional properties. Molecular and Cellular Biology *26*, 843-851.

Lee, D.W., Zhang, K., Ning, Z.Q., Raabe, E.H., Tintner, S., Wieland, R., Wilkins, B.J., Kim, J.M., Blough, R.I., and Arceci, R.J. (2000). Proliferation-associated SNF2-like gene (PASG): a SNF2 family member altered in leukemia. Cancer Research *60*, 3612-3622.

Lei, H., Oh, S.P., Okano, M., Juttermann, R., Goss, K.A., Jaenisch, R., and Li, E. (1996). De novo DNA cytosine methyltransferase activities in mouse embryonic stem cells. Development *122*, 3195-3205.

Leonhardt, H., Page, A.W., Weier, H.U., and Bestor, T.H. (1992). A targeting sequence directs DNA methyltransferase to sites of DNA replication in mammalian nuclei. Cell *71*, 865-873.

Lewis, R., Durr, H., Hopfner, K.P., and Michaelis, J. (2008). Conformational changes of a Swi2/Snf2 ATPase during its mechano-chemical cycle. Nucleic Acids Research *36*, 1881-1890.

Li, D., and Roberts, R. (2001). WD-repeat proteins: structure characteristics, biological function, and their involvement in human diseases. Cellular and Molecular Life Science *58*, 2085-2097.

Li, E., Beard, C., and Jaenisch, R. (1993). Role for DNA methylation in genomic imprinting. Nature *366*, 362-365.

Li, E., Bestor, T.H., and Jaenisch, R. (1992). Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. Cell *69*, 915-926.

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

Li, J., Langst, G., and Grummt, I. (2006b). NoRC-dependent nucleosome positioning silences rRNA genes. The EMBO Journal *25*, 5735-5741.

Li, J.Y., Pu, M.T., Hirasawa, R., Li, B.Z., Huang, Y.N., Zeng, R., Jing, N.H., Chen, T., Li, E., Sasaki, H., *et al.* (2007a). Synergistic function of DNA methyltransferases Dnmt3a and Dnmt3b in the methylation of Oct4 and Nanog. Molecular and Cellular Biology *27*, 8748-8759.

Li, X., Zhang, X.P., Solinger, J.A., Kiianitsa, K., Yu, X., Egelman, E.H., and Heyer, W.D. (2007b). Rad51 and Rad54 ATPase activities are both required to modulate Rad51-dsDNA filament dynamics. Nucleic Acids Research *35*, 4124-4140.

Lia, G., Praly, E., Ferreira, H., Stockdale, C., Tse-Dinh, Y.C., Dunlap, D., Croquette, V., Bensimon, D., and Owen-Hughes, T. (2006). Direct observation of DNA distortion by the RSC complex. Molecular Cell *21*, 417-425.

Lickert, H., Takeuchi, J.K., Von Both, I., Walls, J.R., McAuliffe, F., Adamson, S.L., Henkelman, R.M., Wrana, J.L., Rossant, J., and Bruneau, B.G. (2004). Baf60c is essential for function of BAF chromatin remodelling complexes in heart development. Nature *432*, 107-112.

Lindroth, A.M., Cao, X., Jackson, J.P., Zilberman, D., McCallum, C.M., Henikoff, S., and Jacobsen, S.E. (2001). Requirement of CHROMOMETHYLASE3 for maintenance of CpXpG methylation. Science *292*, 2077-2080.

Lippman, Z., Gendrel, A.V., Black, M., Vaughn, M.W., Dedhia, N., McCombie, W.R., Lavine, K., Mittal, V., May, B., Kasschau, K.D., *et al.* (2004). Role of transposable elements in heterochromatin and epigenetic control. Nature *430*, 471-476.

Liu, Y., Oakeley, E.J., Sun, L., and Jost, J.P. (1998). Multiple domains are involved in the targeting of the mouse DNA methyltransferase to the DNA replication foci. Nucleic Acids Research *26*, 1038-1045.

Lock, L.F., Melton, D.W., Caskey, C.T., and Martin, G.R. (1986). Methylation of the mouse hprt gene differs on the active and inactive X chromosomes. Molecular and Cellular Biology *6*, 914-924.

Logie, C., Tse, C., Hansen, J.C., and Peterson, C.L. (1999). The Core Histone N-Terminal Domains Are Required for Multiple Rounds of Catalytic Chromatin Remodeling by the SWI/SNF and RSC Complexes. Biochemistry *38,* 2514-2522.

Lorch, Y., Zhang, M., and Kornberg, R.D. (1999). Histone octamer transfer by a chromatin-remodeling complex. Cell *96*, 389-392.

Lowary, P.T., and Widom, J. (1998). New DNA sequence rules for high affinity binding to histone octamer and sequence-directed nucleosome positioning. Journal of Molecular Biology *276*, 19-42.

Luckow, V.A., Lee, S.C., Barry, G.F., and Olins, P.O. (1993). Efficient generation of infectious recombinant baculoviruses by site-specific transposon-mediated insertion of foreign genes into a baculovirus genome propagated in Escherichia coli. Journal of Virology *67*, 4566-4579.

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

Luger, K., Rechsteiner, T.J., and Richmond, T.J. (1999). Preparation of nucleosome core particle from recombinant histones. Methods in Enzymology *304*, 3-19.

Lusser, A., Urwin, D.L., and Kadonaga, J.T. (2005). Distinct activities of CHD1 and ACF in ATP-dependent chromatin assembly. Nature Structural & Molecular Biology *12*, 160-166.

Lyko, F., Ramsahoye, B.H., and Jaenisch, R. (2000). DNA methylation in Drosophila melanogaster. Nature *408*, 538-540.

Lyko, F., Ramsahoye, B.H., Kashevsky, H., Tudor, M., Mastrangelo, M.A., Orr-Weaver, T.L., and Jaenisch, R. (1999). Mammalian (cytosine-5) methyltransferases cause genomic DNA methylation and lethality in Drosophila. Nature Genetics *23*, 363-366.

Maison, C., and Almouzni, G. (2004). HP1 and the dynamics of heterochromatin maintenance. Nature Reviews Molecular and Cellular Biology *5*, 296-304.

Mayer-Jung, C., Moras, D., and Timsit, Y. (1997). Effect of cytosine methylation on DNA-DNA recognition at CpG steps. Journal of Molecular Biology *270*, 328-335.

McDowell, T.L., Gibbons, R.J., Sutherland, H., O'Rourke, D.M., Bickmore, W.A., Pombo, A., Turley, H., Gatter, K., Picketts, D.J., Buckle, V.J., *et al.* (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 of the United States of America *96*, 13983-13988.

Meersseman, G., Pennings, S., and Bradbury, E.M. (1992). Mobile nucleosomes--a general behavior. The EMBO Journal *11*, 2951-2959.

Merlo, A., Herman, J.G., Mao, L., Lee, D.J., Gabrielson, E., Burger, P.C., Baylin, S.B., and Sidransky, D. (1995). 5' CpG island methylation is associated with transcriptional silencing of the tumour suppressor p16/CDKN2/MTS1 in human cancers. Nature Medicine *1*, 686-692.

Miura, A., Yonebayashi, S., Watanabe, K., Toyama, T., Shimada, H., and Kakutani, T. (2001). Mobilization of transposons by a mutation abolishing full DNA methylation in Arabidopsis. Nature *411*, 212-214.

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.

Moore, L.E., Pfeiffer, R.M., Poscablo, C., Real, F.X., Kogevinas, M., Silverman, D., Garcia-Closas, R., Chanock, S., Tardon, A., Serra, C., *et al.* (2008). Genomic DNA hypomethylation as a biomarker for bladder cancer susceptibility in the Spanish Bladder Cancer Study: a case-control study. The Lancet Oncology *9*, 359-366.

Moreau, J.L., Lee, M., Mahachi, N., Vary, J., Mellor, J., Tsukiyama, T., and Goding, C.R. (2003). Regulated displacement of TBP from the PHO8 promoter in vivo requires Cbf1 and the Isw1 chromatin remodeling complex. Molecular Cell *11*, 1609-1620.

Moshkin, Y.M., Mohrmann, L., van Ijcken, W.F., and Verrijzer, C.P. (2007). Functional differentiation of SWI/SNF remodelers in transcription and cell cycle control. Molecular and Cellular Biology *27*, 651-661.

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. The Journal of Biological Chemistry *275*, 7648-7655.

Mutskov, V., and Felsenfeld, G. (2004). Silencing of transgene transcription precedes methylation of promoter DNA and histone H3 lysine 9. The EMBO Journal *23*, 138-149.

Myohanen, S.K., Baylin, S.B., and Herman, J.G. (1998). Hypermethylation can selectively silence individual p16ink4A alleles in neoplasia. Cancer Research *58*, 591-593.

Nakagawa, H., Nuovo, G.J., Zervos, E.E., Martin, E.W., Jr., Salovaara, R., Aaltonen, L.A., and de la Chapelle, A. (2001). Age-related hypermethylation of the 5' region of MLH1 in normal colonic mucosa is associated with microsatellite-unstable colorectal cancer development. Cancer Research *61*, 6991-6995.

Nakayama, J.-i., Rice, J.C., Strahl, B.D., Allis, C.D., and Grewal, S.I.S. (2001). Role of Histone H3 Lysine 9 Methylation in Epigenetic Control of Heterochromatin Assembly. Science *292*, 110-113.

Nan, X., Campoy, F.J., and Bird, A. (1997). MeCP2 is a transcriptional repressor with abundant binding sites in genomic chromatin. Cell *88*, 471-481.

Nan, X., Hou, J., Maclean, A., Nasir, J., Lafuente, M.J., Shu, X., Kriaucionis, S., and Bird, A. (2007). Interaction between chromatin proteins MECP2 and ATRX is disrupted by mutations that cause inherited mental retardation. Proceedings of the National Academy of Science USA *104* 2709-2714.

Nan, X., Ng, H.H., Johnson, C.A., Laherty, C.D., Turner, B.M., Eisenman, R.N., and Bird, A. (1998). Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex. Nature *393*, 386-389.

Narlikar, G.J., Fan, H.Y., and Kingston, R.E. (2002). Cooperation between complexes that regulate chromatin structure and transcription. Cell *108*, 475-487.

Neely, K.E., Hassan, A.H., Brown, C.E., Howe, L., and Workman, J.L. (2002). Transcription activator interactions with multiple SWI/SNF subunits. Molecular and Cellular Biology *22*, 1615-1625.

Nimura, K., Ishida, C., Koriyama, H., Hata, K., Yamanaka, S., Li, E., Ura, K., and Kaneda, Y. (2006). Dnmt3a2 targets endogenous Dnmt3L to ES cell chromatin and induces regional DNA methylation. Genes Cells *11*, 1225-1237.

Nishioka, K., Chuikov, S., Sarma, K., Erdjument-Bromage, H., Allis, C.D., Tempst, P., and Reinberg, D. (2002). Set9, a novel histone H3 methyltransferase that facilitates transcription by precluding histone tail modifications required for heterochromatin formation. Genes & Development *16*, 479-489.

Ohm, J.E., McGarvey, K.M., Yu, X., Cheng, L., Schuebel, K.E., Cope, L., Mohammad, H.P., Chen, W., Daniel, V.C., Yu, W., *et al.* (2007). A stem cell-like chromatin pattern may predispose tumor suppressor genes to DNA hypermethylation and heritable silencing. Nature Genetics *39*, 237-242.

Okabe, I., Bailey, L.C., Attree, O., Srinivasan, S., Perkel, J.M., Laurent, B.C., Carlson, M., Nelson, D.L., and Nussbaum, R.L. (1992). Cloning of human and bovine homologs of SNF2/SWI2: a global activator of transcription in yeast S. cerevisiae. Nucleic Acids Research *20*, 4649-4655.

Okano, M., Bell, D.W., Haber, D.A., and Li, E. (1999). DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. Cell *99*, 247-257.

Okano, M., Xie, S., and Li, E. (1998a). Cloning and characterization of a family of novel mammalian DNA (cytosine-5) methyltransferases. Nature Genetics *19*, 219-220.

Okano, M., Xie, S., and Li, E. (1998b). Dnmt2 is not required for de novo and maintenance methylation of viral DNA in embryonic stem cells. Nucleic Acids Research *26*, 2536-2540.

Ooi, S.K., Qiu, C., Bernstein, E., Li, K., Jia, D., Yang, Z., Erdjument-Bromage, H., Tempst, P., Lin, S.P., Allis, C.D., *et al.* (2007). DNMT3L connects unmethylated lysine 4 of histone H3 to de novo methylation of DNA. Nature *448*, 714-717.

Osterod, M., Larsen, E., Le Page, F., Hengstler, J.G., Van Der Horst, G.T., Boiteux, S., Klungland, A., and Epe, B. (2002). A global DNA repair mechanism involving the Cockayne syndrome B (CSB) gene product can prevent the in vivo accumulation of endogenous oxidative DNA base damage. Oncogene *21*, 8232-8239.

Owen-Hughes, T., Utley, R.T., Cote, J., Peterson, C.L., and Workman, J.L. (1996). Persistent site-specific remodeling of a nucleosome array by transient action of the SWI/SNF complex. Science *273*, 513-516.

Pena, P.V., Davrazou, F., Shi, X., Walter, K.L., Verkhusha, V.V., Gozani, O., Zhao, R., and Kutateladze, T.G. (2006). Molecular mechanism of histone H3K4me3 recognition by plant homeodomain of ING2. Nature *442*, 100-103.

Pfeifer, G.P., Tang, M., and Denissenko, M.F. (2000). Mutation hotspots and DNA methylation. Current Topics in Microbiology and Immunology *249*, 1-19.

Platero, J.S., Hartnett, T., and Eissenberg, J.C. (1995). Functional analysis of the chromo domain of HP1. The EMBO Journal *14*, 3977-3986.

Poot, R.A., Dellaire, G., Hulsmann, B.B., Grimaldi, M.A., Corona, D.F., Becker, P.B., Bickmore, W.A., and Varga-Weisz, P.D. (2000). HuCHRAC, a human ISWI chromatin remodelling complex contains hACF1 and two novel histone-fold proteins. The EMBO Journal *19*, 3377-3387.

Pradhan, S., Bacolla, A., Wells, R.D., and Roberts, R.J. (1999). Recombinant human DNA (cytosine-5) methyltransferase. I. Expression, purification, and comparison of de novo and maintenance methylation. The Journal of Biological Chemistry *274*, 33002-33010.

Prendergast, G.C., Lawe, D., and Ziff, E.B. (1991). Association of Myn, the murine homolog of max, with c-Myc stimulates methylation-sensitive DNA binding and ras cotransformation. Cell *65*, 395-407.

Prendergast, G.C., and Ziff, E.B. (1991). Methylation-sensitive sequence-specific DNA binding by the c-Myc basic region. Science *251*, 186-189.

Prokhortchouk, A., Hendrich, B., Jorgensen, H., Ruzov, A., Wilm, M., Georgiev, G., Bird, A., and Prokhortchouk, E. (2001). The p120 catenin partner Kaiso is a DNA methylation-dependent transcriptional repressor. Genes & Development *15*, 1613-1618.

Quinn, J., Fyrberg, A.M., Ganster, R.W., Schmidt, M.C., and Peterson, C.L. (1996). DNA-binding properties of the yeast SWI/SNF complex. Nature *379*, 844-847.

Raabe, E.H., Abdurrahman, L., Behbehani, G., and Arceci, R.J. (2001). An SNF2 factor involved in mammalian development and cellular proliferation. Developmental Dynamics *221*, 92-105.

Ramsahoye, B.H., Biniszkiewicz, D., Lyko, F., Clark, V., Bird, A.P., and Jaenisch, R. (2000a). Non-CpG methylation is prevalent in embryonic stem cells and may be mediated by DNA methyltransferase 3a. Proceedings of the National Academy of Sciences of the United States of America *97*, 5237-5242.

Raschle, M., Van Komen, S., Chi, P., Ellenberger, T., and Sung, P. (2004). Multiple interactions with the Rad51 recombinase govern the homologous recombination function of Rad54. The Journal of Biological Chemistry *279*, 51973-51980.

Reisman, D.N., Sciarrotta, J., Wang, W., Funkhouser, W.K., and Weissman, B.E. (2003). Loss of BRG1/BRM in human lung cancer cell lines and primary lung cancers: correlation with poor prognosis. Cancer Research *63*, 560-566.

Rhee, I., Bachman, K.E., Park, B.H., Jair, K.W., Yen, R.W., Schuebel, K.E., Cui, H., Feinberg, A.P., Lengauer, C., Kinzler, K.W., *et al.* (2002). DNMT1 and DNMT3b cooperate to silence genes in human cancer cells. Nature *416*, 552-556.

Rhee, I., Jair, K.W., Yen, R.W., Lengauer, C., Herman, J.G., Kinzler, K.W., Vogelstein, B., Baylin, S.B., and Schuebel, K.E. (2000). CpG methylation is maintained in human cancer cells lacking DNMT1. Nature *404*, 1003-1007.

Richmond, T.J., and Davey, C.A. (2003a). The structure of DNA in the nucleosome core. Nature *423*, 145-150.

Richmond, T.J., and Davey, C.A. (2003b). The structure of DNA in the nucleosome core. Nature *423*, 145-150.

Riggs, A.D., Xiong, Z., Wang, L., and LeBon, J.M. (1998). Methylation dynamics, epigenetic fidelity and X chromosome structure. Novartis Foundation symposium *214*, 214-225; discussion 225-232.

Robertson, A.K., Geiman, T.M., Sankpal, U.T., Hager, G.L., and Robertson, K.D. (2004). Effects of chromatin structure on the enzymatic and DNA binding functions of DNA methyltransferases DNMT1 and Dnmt3a in vitro. Biochemical and Biophysical Research Communications *322*, 110-118.

Robertson, K.D., Ait-Si-Ali, S., Yokochi, T., Wade, P.A., Jones, P.L., and Wolffe, A.P. (2000). DNMT1 forms a complex with Rb, E2F1 and HDAC1 and represses transcription from E2F-responsive promoters. Nature Genetics *25*, 338-342.

Ross, M.T., Grafham, D.V., Coffey, A.J., Scherer, S., McLay, K., Muzny, D., Platzer, M., Howell, G.R., Burrows, C., Bird, C.P., *et al.* (2005). The DNA sequence of the human X chromosome. Nature *434*, 325-337.

Rountree, M.R., Bachman, K.E., and Baylin, S.B. (2000). DNMT1 binds HDAC2 and a new co-repressor, DMAP1, to form a complex at replication foci. Nature Genetics *25*, 269-277.

Sado, T., Fenner, M.H., Tan, S.S., Tam, P., Shioda, T., and Li, E. (2000). X inactivation in the mouse embryo deficient for Dnmt1: distinct effect of hypomethylation on imprinted and random X inactivation. Developmental Biology *225*, 294-303.

Saha, A., Wittmeyer, J., and Cairns, B.R. (2002). Chromatin remodeling by RSC involves ATP-dependent DNA translocation. Genes & Development *16*, 2120-2134.

Saha, A., Wittmeyer, J., and Cairns, B.R. (2005). Chromatin remodeling through directional DNA translocation from an internal nucleosomal site. Nature Structural & Molecular Biology *12*, 747-755.

Santoro, R., and Grummt, I. (2005). Epigenetic mechanism of rRNA gene silencing: temporal order of NoRC-mediated histone modification, chromatin remodeling, and DNA methylation. Molecular and Cellular Biology *25*, 2539-2546.

Santoro, R., Li, J., and Grummt, I. (2002). The nucleolar remodeling complex NoRC mediates heterochromatin formation and silencing of ribosomal gene transcription. Nature Genetics *32*, 393-396.

Sarraf, S.A., and Stancheva, I. (2004). Methyl-CpG binding protein MBD1 couples histone H3 methylation at lysine 9 by SETDB1 to DNA replication and chromatin assembly. Molecular Cell *15*, 595-605.

Schalch, T., Duda, S., Sargent, D.F., and Richmond, T.J. (2005). X-ray structure of a tetranucleosome and its implications for the chromatin fibre. Nature *436*, 138-141.

Schermelleh, L., Haemmer, A., Spada, F., Rosing, N., Meilinger, D., Rothbauer, U., Cardoso, M.C., and Leonhardt, H. (2007). Dynamics of Dnmt1 interaction with the replication machinery and its role in postreplicative maintenance of DNA methylation. Nucleic Acids Research *35*, 4301-4312.

Schermelleh, L., Spada, F., Easwaran, H.P., Zolghadr, K., Margot, J.B., Cardoso, M.C., and Leonhardt, H. (2005). Trapped in action: direct visualization of DNA methyltransferase activity in living cells. Nature Methods *2*, 751-756.

Schlesinger, Y., Straussman, R., Keshet, I., Farkash, S., Hecht, M., Zimmerman, J., Eden, E., Yakhini, Z., Ben-Shushan, E., Reubinoff, B.E., *et al.* (2007). Polycomb-mediated methylation on Lys27 of histone H3 pre-marks genes for de novo methylation in cancer. Nature Genetics *39*, 232-236.

Sevenet, N., Sheridan, E., Amram, D., Schneider, P., Handgretinger, R., and Delattre, O. (1999). Constitutional mutations of the hSNF5/INI1 gene predispose to a variety of cancers. American Journal of Human Genetics *65*, 1342-1348.

Sharif, J., Muto, M., Takebayashi, S., Suetake, I., Iwamatsu, A., Endo, T.A., Shinga, J., Mizutani-Koseki, Y., Toyoda, T., Okamura, K., *et al.* (2007). The SRA protein Np95 mediates epigenetic inheritance by recruiting Dnmt1 to methylated DNA. Nature *450*, 908-912.

Shen, J.C., Rideout, W.M., 3rd, and Jones, P.A. (1994). The rate of hydrolytic deamination of 5-methylcytosine in double-stranded DNA. Nucleic Acids Research *22*, 972-976.

Shen, X., Mizuguchi, G., Hamiche, A., and Wu, C. (2000). A chromatin remodelling complex involved in transcription and DNA processing. Nature *406*, 541-544.

Shen, X., Ranallo, R., Choi, E., and Wu, C. (2003). Involvement of actin-related proteins in ATP-dependent chromatin remodeling. Molecular Cell *12*, 147-155.

Shi, X., Hong, T., Walter, K.L., Ewalt, M., Michishita, E., Hung, T., Carney, D., Pena, P., Lan, F., Kaadige, M.R., *et al.* (2006). ING2 PHD domain links histone H3 lysine 4 methylation to active gene repression. Nature *442*, 96-99.

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.

Siegel, L.M., and Monty, K.J. (1965). Determination of Molecular Weights and Frictional Ratios of Macromolecules in Impure Systems: Aggregation of Urease. Biochemical and Biophysical Research Communications *19*, 494-499.

Silva, A.J., Ward, K., and White, R. (1993). Mosaic methylation in clonal tissue. Developmental Biology *156*, 391-398.

Singer, T., Yordan, C., and Martienssen, R.A. (2001). Robertson's Mutator transposons in A. thaliana are regulated by the chromatin-remodeling gene Decrease in DNA Methylation (DDM1). Genes & Development *15*, 591-602.

Singleton, M.R., Dillingham, M.S., and Wigley, D.B. (2007). Structure and mechanism of helicases and nucleic acid translocases. Annual Review of Biochemistry *76*, 23-50.

Spada, F., Haemmer, A., Kuch, D., Rothbauer, U., Schermelleh, L., Kremmer, E., Carell, T., Langst, G., and Leonhardt, H. (2007). DNMT1 but not its interaction with the replication machinery is required for maintenance of DNA methylation in human cells. The Journal of Cell Biology *176*, 565-571.

Sprouse, R.O., Brenowitz, M., and Auble, D.T. (2006). Snf2/Swi2-related ATPase Mot1 drives displacement of TATA-binding protein by gripping DNA. The EMBO Journal *25*, 1492-1504.

Stopka, T., and Skoultchi, A.I. (2003). The ISWI ATPase Snf2h is required for early mouse development. Proceedings of the National Academy of Sciences of the United States of America *100*, 14097-14102.

Stratagene (1998). λ *select-cII* mutation detection system for Big Blue rodents handbook, Version 028001 edn.

Strohner, R., Nemeth, A., Jansa, P., Hofmann-Rohrer, U., Santoro, R., Langst, G., and Grummt, I. (2001). NoRC--a novel member of mammalian ISWI-containing chromatin remodeling machines. The EMBO Journal *20*, 4892-4900.

Strohner, R., Wachsmuth, M., Dachauer, K., Mazurkiewicz, J., Hochstatter, J., Rippe, K., and Langst, G. (2005). A 'loop recapture' mechanism for ACF-dependent nucleosome remodeling. Nature Structural & Molecular Biology *12*, 683-690.

Strunnikova, M., Schagdarsurengin, U., Kehlen, A., Garbe, J.C., Stampfer, M.R., and Dammann, R. (2005). Chromatin inactivation precedes de novo DNA methylation during the progressive epigenetic silencing of the RASSF1A promoter. Molecular and Cellular Biology *25*, 3923-3933.

Studitsky, V.M., Clark, D.J., and Felsenfeld, G. (1994). A histone octamer can step around a transcribing polymerase without leaving the template. Cell *76*, 371-382.

Suetake, I., Miyazaki, J., Murakami, C., Takeshima, H., and Tajima, S. (2003). Distinct enzymatic properties of recombinant mouse DNA methyltransferases Dnmt3a and Dnmt3b. Journal of Biochemistry *133*, 737-744.

Suetake, I., Shinozaki, F., Miyagawa, J., Takeshima, H., and Tajima, S. (2004). DNMT3L stimulates the DNA methylation activity of Dnmt3a and Dnmt3b through a direct interaction. The Journal of Biological Chemistry *279*, 27816-27823.

Sun, L.Q., Lee, D.W., Zhang, Q., Xiao, W., Raabe, E.H., Meeker, A., Miao, D., Huso, D.L., and Arceci, R.J. (2004). Growth retardation and premature aging phenotypes in mice with disruption of the SNF2-like gene, PASG. Genes & Development *18*, 1035-1046.

Suzuki, M.M., Kerr, A.R.W., De Sousa, D., and Bird, A. (2007). CpG methylation is targeted to transcription units in an invertebrate genome. Genome *17,* 625-631.

Takai, D., and Jones, P.A. (2003). The CpG island searcher: a new WWW resource. In silico biology *3*, 235-240.

Takebayashi, S., Tamura, T., Matsuoka, C., and Okano, M. (2007). Major and essential role for the DNA methylation mark in mouse embryogenesis and stable association of DNMT1 with newly replicated regions. Molecular and Cellular Biology *27*, 8243-8258.

Takeshima, H., Suetake, I., Shimahara, H., Ura, K., Tate, S., and Tajima, S. (2006). Distinct DNA methylation activity of Dnmt3a and Dnmt3b towards naked and nucleosomal DNA. Journal of Biochemistry *139*, 503-515.

Tamaru, H., and Selker, E.U. (2001). A histone H3 methyltransferase controls DNA methylation in Neurospora crassa. Nature *414*, 277-283.

Tang, J., Wu, S., Liu, H., Stratt, R., Barak, O.G., Shiekhattar, R., Picketts, D.J., and Yang, X. (2004). A novel transcription regulatory complex containing death domain-associated protein and the ATR-X syndrome protein. The Journal of Biological Chemistry *279*, 20369-20377.

Tong, J.K., Hassig, C.A., Schnitzler, G.R., Kingston, R.E., and Schreiber, S.L. (1998). Chromatin deacetylation by an ATP-dependent nucleosome remodelling complex. Nature *395*, 917-921.

Toyota, M., Ahuja, N., Ohe-Toyota, M., Herman, J.G., Baylin, S.B., and Issa, J.P. (1999). CpG island methylator phenotype in colorectal cancer. Proceedings of the National Academy of Sciences of the United States of America *96*, 8681-8686.

Tremethick, D.J. (2007). Higher-Order Structures of Chromatin: The Elusive 30 nm Fiber. Cell *128*, 651-654.

Troelstra, C., van Gool, A., de Wit, J., Vermeulen, W., Bootsma, D., and Hoeijmakers, J.H. (1992). ERCC6, a member of a subfamily of putative helicases, is involved in Cockayne's syndrome and preferential repair of active genes. Cell *71*, 939-953.

Trojer, P., and Reinberg, D. (2007). Facultative Heterochromatin: Is There a Distinctive Molecular Signature? Molecular Cell *28*, 1-13.

Tsukiyama, T., Daniel, C., Tamkun, J., and Wu, C. (1995). ISWI, a member of the SWI2/SNF2 ATPase family, encodes the 140 kDa subunit of the nucleosome remodeling factor. Cell *83*, 1021-1026.

Tsukiyama, T., and Wu, C. (1995). Purification and properties of an ATP-dependent nucleosome remodeling factor. Cell *83*, 1011-1020.

Tweedie, S., Ng, H.H., Barlow, A.L., Turner, B.M., Hendrich, B., and Bird, A. (1999). Vestiges of a DNA methylation system in Drosophila melanogaster? Nature genetics *23*, 389-390.

Vakoc, C.R., Mandat, S.A., Olenchock, B.A., and Blobel, G.A. (2005). Histone H3 Lysine 9 Methylation and HP1[gamma] Are Associated with Transcription Elongation through Mammalian Chromatin. Molecular Cell *19*, 381-391.

Vallender, E.J., Pearson, N.M., and Lahn, B.T. (2005). The X chromosome: not just her brother's keeper. Nature Genetics *37*, 343-345.

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.

Versteege, I., Sevenet, N., Lange, J., Rousseau-Merck, M.F., Ambros, P., Handgretinger, R., Aurias, A., and Delattre, O. (1998). Truncating mutations of hSNF5/INI1 in aggressive paediatric cancer. Nature *394*, 203-206.

Vilkaitis, G., Suetake, I., Klimasauskas, S., and Tajima, S. (2005). Processive methylation of hemimethylated CpG sites by mouse Dnmt1 DNA methyltransferase. The Journal of Biological Chemistry *280*, 64-72.

Villard, L., and Fontes, M. (2002). Alpha-thalassemia/mental retardation syndrome, X-Linked (ATR-X, MIM #301040, ATR-X/XNP/XH2 gene MIM #300032). European Journal of Human Genetics *10*, 223-225.

Vongs, A., Kakutani, T., Martienssen, R.A., and Richards, E.J. (1993). Arabidopsis thaliana DNA methylation mutants. Science *260*, 1926-1928.

Wain, E.M., Mitchell, T.J., Russell-Jones, R., and Whittaker, S.J. (2005). Fine mapping of chromosome 10q deletions in mycosis fungoides and sezary syndrome: identification of two discrete regions of deletion at 10q23.33-24.1 and 10q24.33-25.1. Genes Chromosomes Cancer *42*, 184-192.

Wang, H.-B., and Zhang, Y. (2001). Mi2, an auto-antigen for dermatomyositis, is an ATP-dependent nucleosome remodeling factor. Nucleic Acids Research *29,* 2517-2521.

Wassmann, K., Liberal, V., and Benezra, R. (2003). Mad2 phosphorylation regulates its association with Mad1 and the APC/C. The EMBO Journal *22*, 797-806.

Weber, M., Hellmann, I., Stadler, M.B., Ramos, L., Paabo, S., Rebhan, M., and Schubeler, D. (2007). Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome. Nature Genetics *39*, 457-466.

Webster, K.E., O'Bryan, M.K., Fletcher, S., Crewther, P.E., Aapola, U., Craig, J., Harrison, D.K., Aung, H., Phutikanit, N., Lyle, R., *et al.* (2005). Meiotic and epigenetic defects in Dnmt3L-knockout mouse spermatogenesis. Proceedings of the National Academy of Sciences of the United States of America *102*, 4068-4073.

Weisenberger, D.J., Siegmund, K.D., Campan, M., Young, J., Long, T.I., Faasse, M.A., Kang, G.H., Widschwendter, M., Weener, D., Buchanan, D., *et al.* (2006). CpG island methylator phenotype underlies sporadic microsatellite instability and is tightly associated with BRAF mutation in colorectal cancer. Nature Genetics *38*, 787-793.

Whitehouse, I., Flaus, A., Cairns, B.R., White, M.F., Workman, J.L., and Owen-Hughes, T. (1999). Nucleosome mobilization catalysed by the yeast SWI/SNF complex. Nature *400*, 784-787.

Whitehouse, I., Stockdale, C., Flaus, A., Szczelkun, M.D., and Owen-Hughes, T. (2003). Evidence for DNA translocation by the ISWI chromatin-remodeling enzyme. Molecular and Cellular Biology *23*, 1935-1945.

Widschwendter, M., Fiegl, H., Egle, D., Mueller-Holzner, E., Spizzo, G., Marth, C., Weisenberger, D.J., Campan, M., Young, J., Jacobs, I., *et al.* (2007). Epigenetic stem cell signature in cancer. Nature Genetics *39*, 157-158.

Wolf, S.F., Jolly, D.J., Lunnen, K.D., Friedmann, T., and Migeon, B.R. (1984). Methylation of the hypoxanthine phosphoribosyltransferase locus on the human X chromosome: implications for X-chromosome inactivation. Proceedings of the National Academy of Sciences of the United States of America *81*, 2806-2810.

Wong, A.K., Shanahan, F., Chen, Y., Lian, L., Ha, P., Hendricks, K., Ghaffari, S., Iliev, D., Penn, B., Woodland, A.M., *et al.* (2000). BRG1, a component of the SWI-SNF complex, is mutated in multiple human tumor cell lines. Cancer Research *60*, 6171-6177.

Wu, J.C., and Santi, D.V. (1985). On the mechanism and inhibition of DNA cytosine methyltransferases. Progress in Clinical and Biological Research *198*, 119-129.

Wu, J.C., and Santi, D.V. (1987). Kinetic and catalytic mechanism of HhaI methyltransferase. The Journal of Biological Chemistry *262*, 4778-4786.

Wysocka, J., Swigut, T., Xiao, H., Milne, T.A., Kwon, S.Y., Landry, J., Kauer, M., Tackett, A.J., Chait, B.T., Badenhorst, P., *et al.* (2006). A PHD finger of NURF couples histone H3 lysine 4 trimethylation with chromatin remodelling. Nature *442*, 86-90.

Wyszynski, M.W., Gabbara, S., Kubareva, E.A., Romanova, E.A., Oretskaya, T.S., Gromova, E.S., Shabarova, Z.A., and Bhagwat, A.S. (1993). The cysteine conserved among DNA cytosine methylases is required for methyl transfer, but not for specific DNA binding. Nucleic Acids Research *21,* 295-301.

Xi, S., Zhu, H., Xu, H., Schmidtmann, A., Geiman, T.M., and Muegge, K. (2007). Lsh controls Hox gene silencing during development. Proceedings of the National Academy of Sciences of the United States of America *104*, 14366-14371.

Xie, S., Wang, Z., Okano, M., Nogami, M., Li, Y., He, W.W., Okumura, K., and Li, E. (1999). Cloning, expression and chromosome locations of the human DNMT3 gene family. Gene *236*, 87-95.

Xu, G.L., Bestor, T.H., Bourc'his, D., Hsieh, C.L., Tommerup, N., Bugge, M., Hulten, M., Qu, X., Russo, J.J., and Viegas-Pequignot, E. (1999). Chromosome instability and immunodeficiency syndrome caused by mutations in a DNA methyltransferase gene. Nature *402*, 187-191.

Xue, Y., Gibbons, R., Yan, Z., Yang, D., McDowell, T.L., Sechi, S., Qin, J., Zhou, S., Higgs, D., and Wang, W. (2003). The ATRX syndrome protein forms a chromatin-remodeling complex with Daxx and localizes in promyelocytic leukemia nuclear bodies. Proceedings of the National Academy of Sciences of the United States of America *100*, 10635-10640.

Yan, Q., Cho, E., Lockett, S., and Muegge, K. (2003a). Association of Lsh, a regulator of DNA methylation, with pericentromeric heterochromatin is dependent on intact heterochromatin. Molecular and Cellular Biology *23*, 8416-8428.

Yan, Q., Huang, J., Fan, T., Zhu, H., and Muegge, K. (2003b). Lsh, a modulator of CpG methylation, is crucial for normal histone methylation. The EMBO Journal *22*, 5154-5162.

Yan, Z., Cui, K., Murray, D.M., Ling, C., Xue, Y., Gerstein, A., Parsons, R., Zhao, K., and Wang, W. (2005). PBAF chromatin-remodeling complex requires a novel specificity subunit, BAF200, to regulate expression of selective interferon-responsive genes. Genes & Development *19*, 1662-1667.

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.

Yano, M., Ouchida, M., Shigematsu, H., Tanaka, N., Ichimura, K., Kobayashi, K., Inaki, Y., Toyooka, S., Tsukuda, K., Shimizu, N., *et al.* (2004). Tumor-specific exon creation of the HELLS/SMARCA6 gene in non-small cell lung cancer. International Journal of Cancer *112*, 8-13.

Yoon, H.G., Chan, D.W., Reynolds, A.B., Qin, J., and Wong, J. (2003). N-CoR mediates DNA methylation-dependent repression through a methyl CpG binding protein Kaiso. Molecular Cell *12*, 723-734.

You, X., Nguyen, A.W., Jabaiah, A., Sheff, M.A., Thorn, K.S., and Daugherty, P.S. (2006). Intracellular protein interaction mapping with FRET hybrids. Proceedings of the National Academy of Sciences of the United States of America *103*, 18458-18463.

Yu, L., Liu, C., Vandeusen, J., Becknell, B., Dai, Z., Wu, Y.Z., Raval, A., Liu, T.H., Ding, W., Mao, C., *et al.* (2005). Global assessment of promoter methylation in a mouse model of cancer identifies ID4 as a putative tumor-suppressor gene in human leukemia. Nature Genetics *37*, 265-274.

Zhang, Y., LeRoy, G., Seelig, H.P., Lane, W.S., and Reinberg, D. (1998). The dermatomyositis-specific autoantigen Mi2 is a component of a complex containing histone deacetylase and nucleosome remodeling activities. Cell *95*, 279-289.

Zhang, Y., Smith, C.L., Saha, A., Grill, S.W., Mihardja, S., Smith, S.B., Cairns, B.R., Peterson, C.L., and Bustamante, C. (2006a). DNA Translocation and Loop Formation Mechanism of Chromatin Remodeling by SWI/SNF and RSC. Molecular Cell *24*, 559-568.

Zhang, Y., Smith, C.L., Saha, A., Grill, S.W., Mihardja, S., Smith, S.B., Cairns, B.R., Peterson, C.L., and Bustamante, C. (2006b). DNA translocation and loop formation mechanism of chromatin remodeling by SWI/SNF and RSC. Molecular Cell *24*, 559-568.

Zhou, Y., Santoro, R., and Grummt, I. (2002). The chromatin remodeling complex NoRC targets HDAC1 to the ribosomal gene promoter and represses RNA polymerase I transcription. The EMBO Journal *21*, 4632-4640.

Zhu, H., Geiman, T.M., Xi, S., Jiang, Q., Schmidtmann, A., Chen, T., Li, E., and Muegge, K. (2006). Lsh is involved in de novo methylation of DNA. The EMBO Journal *25*, 335-345.

Zofall, M., Persinger, J., Kassabov, S.R., and Bartholomew, B. (2006). Chromatin remodeling by ISW2 and SWI/SNF requires DNA translocation inside the nucleosome. Nature Structural & Molecular Biology *13*, 339-346.