Expression, Purification and Characterisation of Recombinant Chromatin Assembly Factor 1

Nikki Royle

Department of Biochemistry
University of Cambridge
Peterhouse

Dissertation for Doctorate of Philosophy in Biochemistry

Declaration

This dissertation and the experimental work herein is all my own, except where specifically indicated in the text. It has not been previously submitted for any qualification at any university.

Dedication

To my parents, for everything.

Abstract

Chromatin Assembly Factor 1 (CAF-1) is the only known replication dependant histone chaperone, responsible for the deposition of the histone H3/H4 tetramer onto DNA. Found in all eukaryotes, CAF-1 consists of three subunits, p150, p60 and p48. Since its identification work on CAF-1 has mainly focused on *in vivo* studies due to the lack of a reliable method to produce large quantities of recombinant protein for biochemical studies.

Herein the cloning, production and purification of the three subunits of recombinant CAF-1 is described. The proteins were expressed as complexes and individually in insect cells and *Escherichia coli*, optimised protocols are described for maximum protein recovery and purity. Constructs of p150 and p60 were also produced and used to analyse the binding regions and modes of both the p48 and p60 proteins to p150. It is shown that the two smaller subunits of CAF-1 do not interact in the absence of p150 and that the p150 subunit of CAF-1 acts as a scaffold for assembly of the complex, binding directly to both p48 and p60. The stoichiometry of the CAF-1 complex was also investigated and a basis for further work, including structural studies, discussed.

Acknowledgements

I would like to thank my supervisor Professor Ernest Laue for his ideas, help and support; Dr Natasha Murzina for long discussions and help with both practical problems and writing; Dr Stephen McLaughlin for his supervision in the lab and encouragement during my first year and for his help since leaving, both practical and with writing; Dr Sara Lejon, Andal Murthy, Dr Aleksandra Watson, Dr Wei Zhang and everyone else in the EDL and NMR labs for their help and encouragement. I would also like to thank Dr Zbigniew Pietras for returning the favour as my proofreader and pestering me for chapters when things were going slowly.

At EMBL Grenoble I would like to thank our collaborator Dr Imre Berger for all his help and kindness; Frederic Garzioni for his expertise and for training me in insect cell expression; Yan Ni for preparation of the vector containing the CAF-1 self-cleaving fusion protein and his help with molecular biology; Maxime Chaillet for his help solving protein purification problems and the rest of the Berger and Schaffitzel groups for making me so welcome on my visits.

I thank the BBSRC for my studentship; the Wellcome Trust for the funding for the research in Cambridge and P-CUBE for funding my time in Grenoble. Peterhouse has been a home to me during my time in Cambridge and has provided me with much needed moral support from friends in the same situation.

Finally I thank my parents for all their love and support over the years, without which I would not be who I am today.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Chromatin

The DNA content of one human cell is approximately 2 meters long. This has to be packed into a nucleus a few micrometres in diameter whilst remaining accessible for transcription and untangled to allow efficient replication. This dense, ordered packing is achieved by the formation of an ordered structure known as chromatin. Double stranded DNA is wound around a protein core, forming a nucleosome (Kornberg, 1977; Luger *et al.*, 1997). Nucleosomes are traditionally thought to then condense to form fibres that coil further and pack into a super-structure known as a chromosome, as illustrated in Figure 1.

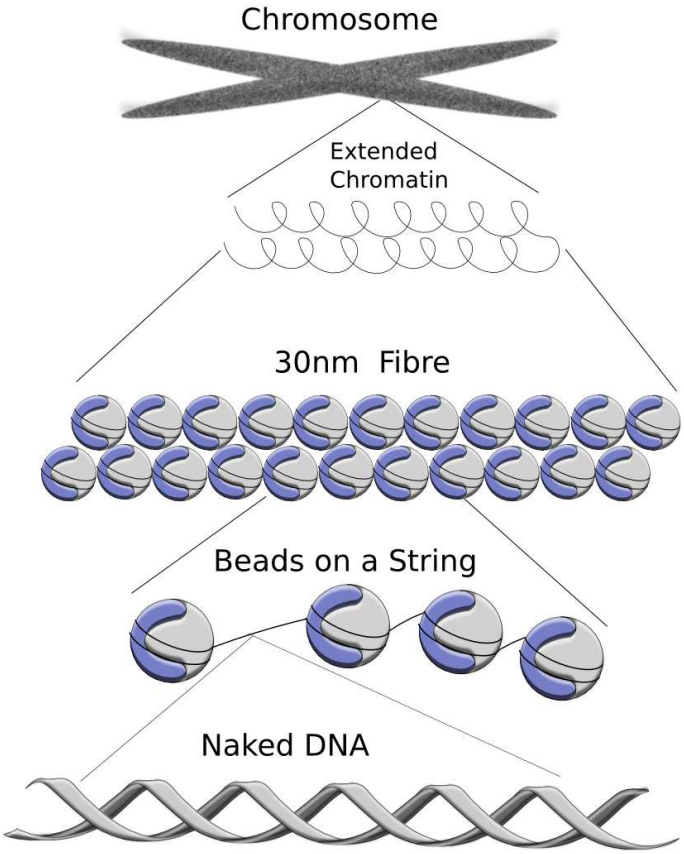


Figure 1: **Traditional View of Chromatin Packing**.
Naked DNA is wrapped around nucleosomes to form a loose structure with a beads-on-a-string structure. These are traditionally thought to pack closely to form a 30 nm fibre which winds up into a less well defined structure which packs further to form a chromosome.

In a nucleosome, 146 bp for DNA is packed around a protein octamer (the nucleosome core), consisting of four duplicated subunits, histones H2A, H2B, H3 and H4 (Figure 2). This core is assembled by the deposition onto naked DNA of a H3/H4 tetramer followed by two H2A/H2B dimers. The crystal structure of the nucleosome particle has been solved and shows a globular core surrounded by DNA with the N- and C-terminal 'tails' of the histones on the outside of this compact structure (Luger *et al.*, 1997), (as shown in Figure 2).

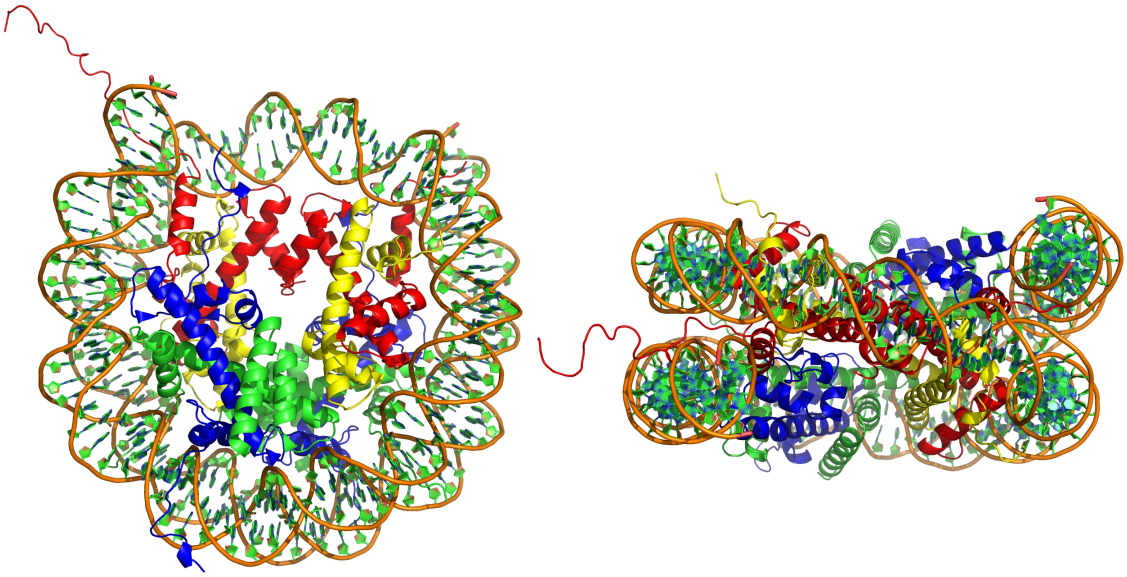


Figure 2: **Structure of the Nucleosome**.
Representation of the crystal structure of the nucleosome, showing DNA encircling a nucleosome core. DNA is shown as a double helix in orange and green. Histone H3 is red, H4 yellow, H2A blue, H2B green. Figure produced using PyMol (http://www.pymol.org/), PDB ID 1AOI (Luger *et al.*, 1997).

Nucleosomes can then be capped by the linker histone H1 to make a chromatosome, engaging 167 bp of DNA in a more tightly wound structure. Electron microscopy shows that many chromatosomes incorporated over a strand of DNA resemble "Beads on a String", known as the 10 nm filament (Tremethick, 2007; Staynov, 2008). A higher order structure was also observed in early work, known as the 30 nm fibre, discussed in greater detail below. The 30 nm fibre was believed to coil further to form chromosomes, which are diffuse, loosely packed structures during growth phases and DNA replication but heavily structured, densely packed constructs during cell division (Staynov, 2008; Robinson and Rhodes, 2006).

1.1.1 The 30 nm Fibre

The compact 30 nm fibre has long been thought to be the first level of compaction for silent DNA. The structure and indeed, existence of the '30 nm fibre' is still under debate, despite being initially reported in the late '70s as it could not be crystallised due to the heterogeneous nature of chromatin with fibres of diameter 33 and 44 nm initially observed (Staynov, 2008).

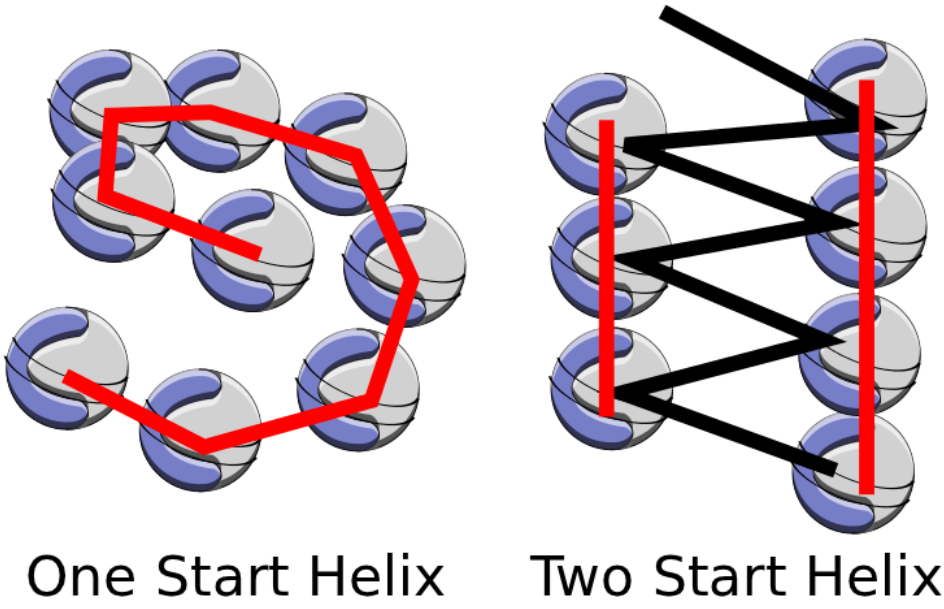


Figure 3: **Two Models of the 30 nm Fibre.** The one start model can have a number of arrangements of DNA within the solenoidal stack of nucleosomes, whereas the two start model has a zigzag arrangement, indicated by the thick black line. The red line indicates the stacking of nucleosomes, Based on a Figure from Staynov (2008)

There have been several models proposed for the structure of the 30 nm fibre, illustrated in Figure 3. These can be divided into two categories defined by the presence of one or two start helices. In the two start model the DNA zigzags across between two stacks of nucleosomes binding one in each stack alternately, with the result resembling the DNA double helix. In the one start helix the nucleosomes form a single solenoid, with the DNA running between histones sequentially in the simplest model, similar to a coiled spring. More complex one start models have the same helical nucleosome arrangement but with the DNA crossing the helix to bind the opposite rather than the sequential nucleosome. Chromatin arranged in both the one and two start models has been observed under different *in vitro* conditions (Robinson and Rhodes, 2006; Wu *et al.*, 2007).

Reconstitution of chromatin using strongly positioned nucleosomes (see section 1.2) allowed (Schalch *et al.*, 2005) to crystallise a tetranucleosome (four nucleosomes on a stretch of DNA). This structure was obtained using a high concentration of salt and lacks the H1 linker histone. The result appears to follow the two-start model. The one start helix arrangement was visualised by electron microscopy (Robinson *et al.*, 2006) of fibres formed in the presence of H1 and with variable lengths of DNA between histones.

To reconcile the discrepancy between these different results it was proposed that the structure of the fibre varies depending on the environment, with shorter linkers between nucleosomes resulting in more compact ribbons and the two start helix (Robinson and Rhodes, 2006).

Very recently, further data utilising SAXS and cryo-EM measurements has suggested that the 30 nm fibre is an artefact of chromatin purification in many human cells. A fibre of diameter around 30 nm, consistent with a model of a two-start helix was observed in the transcriptionally silent chicken erythrocyte cells. However 30 nm structures in mitotic HeLa cells were shown to be due to ribosomal aggregation rather than chromatin packing (Nishino *et al.*, 2012), indicating that the 30 nm fibre may not even exist in the majority of cells.

Given the most recent data, the 30 nm fibre appears to be a highly specific case seen in cells with no active chromatin or as an artefact of regularly spaced nucleosomal arrays formed *in vitro*. The compact fibre appears to be totally absent exist in at least one human cell line. The lack of this defined structure in potentially any transcriptionally active cell may explain why the finer details of the fibre have remained elusive for so long.

1.1.2 Heterochromatin

There are two major types of chromatin, identified initially by their different staining pattern with dyes such as DAPI (Zink *et al.*, 2003). Euchromatin is lightly staining, diffuse chromatin which is usually gene rich and undergoing active transcription. Heterochromatin is darkly staining, generally inactive, compact chromatin associated with gene repression. Euchromatin tends to be enriched in acetylated histones H3 and H4 along with methylation at H3 K4 (H3 lysine 4), whereas heterochromatin is hypoacetylated and contains methylated histones including H3 K9 and H3 K27 as well as methylated DNA (Noma *et al.*, 2001; Jenuwein and Allis, 2001; Tamaru, 2010). Histone variants and their modifications are discussed in detail in section 1.1.3.

There are two types of heterochromatin: constitutive and facultative. Constitutive heterochromatin is repressed in all of an organism's cells, generally containing highly repetitive sequences such as telomeres and centromeric regions, which interact with cohesin proteins

during cell division, as well as transposable elements, preventing them from destabilising the genome (Zeng *et al.*, 2010). Facultative chromatin can be silenced by repressor proteins and RNA, both of which recruit repressor complexes to induce histone modifications (Beisel and Paro, 2011). The best studied model of faculative chromatin formation is that of X-chromosome inactivation which prevents gene dosage problems in mammalian females(Beisel and Paro, 2011). Here repression is is the result of silencing via the Xist RNA, which spreads across the chromosome and recruits proteins such as the Polycomb complexes, causing enrichment of repressive marks such as trimethylation of H3 K27 and DNA methylation (Basu and Zhang, 2011) and repressive histone variants such as macroH2A (Costanzi and Pehrson, 1998).

Constitutive heterochromatin contains methylated DNA and methylated H3 K9 which recruits HP1, the major repressor protein. HP1 is found in almost all eukaryotes with the exception of *S. cerevisiae* which utilises a SIR protein based repression system instead (Kwon and Workman, 2011). HP1 (Swi6 in *S. pombe*) is critical for maintenance of heterochromatin and consists of three main isoforms in mammals: HPα which localises to heterochromatin, HP1β found in hetero and euchromatin and HP1γ mainly seen on euchromatin, where it fulfils functions other than heterochromatin formation. All the HP1 variants contain a chromodomain and a chromo shadow domain. The chromodomain binds methylated H3 K9 to localise the protein to heterochromatin and the chromo shadow domain provides a binding site for heterochromatin associated proteins and for self-association of HP1 (Kwon and Workman, 2011; Zeng *et al.*, 2010; Murzina *et al.*, 2008).

Once HP1α or β has bound to a H3 K9 methylation motif it can propagate the heterochromatic state. Studies utilising Swi6 have indicated that this is facilitated by self-association of HP1 with four HP1 monomers binding to a H3 K9 modified mononucleosome, leaving two chromoshadow domains free to recruit the next modified nucleosome and so causing the propagation of HP1 binding (Canzio *et al.*, 2011). HP1 is also capable of recruiting other chromatin associated proteins such as the methylase Suv39h1, which methylates H3 K9, inducing a feedback loop of HP1 recruitment (Maison and Almouzni, 2004; Probst *et al.*, 2009). This may also be important during replication of heterochromatin to allow reproduction of the repressive histone modifications.

Upon replication of heterochromatin the ISWI-type complex ACF (ATP-utilising chromatin assembly and remodelling factor, see section 1.2.5.2) is recruited to heterochromatin along with the replication machinery including the MCM complex (see section 1.4) to provide local destabilisation of the heterochromatin to allow passage of the replication fork (Collins *et al.*, 2002). Behind the replication fork, nucleosomes are replaced by the action of the CAF-1 (Chromatin Assembly Factor 1, the focus of this study) complex and other histone chaperones (See section 1.3). The p150 subunit of the CAF-1 complex is known to bind HP1 independently to the action of CAF-1 during replication (Murzina *et al.*, 2008). This interaction may recruit HP1 and therefore Suv39h1 to newly replicated heterochromatin, allowing re-establishment of epigenetic marks (Maison and Almouzni, 2004).

To prevent the spread of the heterochromatic state throughout the genome, proteins known as insulators form boundary sites. Proteins containing the JmjC domain such as DMM1 in *Neurospora* and Epe1 in *S. pombe* bind to chromatin and prevent further methylation and spread of HP1/Swi6 (Tamaru, 2010). The CTCF protein binds to cHS4 DNA elements and defines chromatin boundaries by looping out sections of DNA (Phillips and Corces, 2009). This can block the spread of heterochromatin and also prevents the interaction of transcription promoters and enhancers in active DNA (Gaszner and Felsenfeld, 2006).

1.1.3 Histones

There are two types of histone proteins, those that make up the nucleosome, i.e. the core histones (histones H2A, H2B, H3 and H4) and the linker histones, e.g. H1 and H5 which are involved in chromatin compaction. The linker histones function by binding to the nucleosome to stabilise the interaction between the DNA and core histones and allow packing to form higher order structures (Happel and Doenecke, 2009).

The core histones are highly charged, highly conserved proteins, only found in the cell bound to histone chaperones or DNA. Histones H3 and H4 are the first to be deposited onto DNA during chromatin formation, followed by H2A and H2B. These histones exist as dimers and

occasionally tetramers when bound to chaperones in the cell, forming an octamer containing two copies of each core histone (Luger *et al.*, 1997).

Whilst histones are highly conserved across species there are variant forms of some of them. Within the H3/H4 dimer, H4 is invariant whereas histone H3 has several common variants. Canonical histone H3, also known as H3.1 is the most abundant histone H3 type in replicating cells and is deposited on newly replicated DNA by CAF-1 (Tagami *et al.*, 2004). Histone H3.2 only differs from H3.1 at one amino acid and is also replication dependent but may have a role in chromatin based silencing (Hake and Allis, 2006). Histone H3.3 is the major H3 variant outside replication and is the most abundant histone in non-replicating cells (Henikoff and Ahmad, 2005). It differs from H3.1 by four residues and marks DNA undergoing transcription (Tagami *et al.*, 2004). During transcription H3.3 is retained on the daughter chromatin and is thought to act as an epigenetic marker of active chromatin. It is deposited by HIRA (HIstone Regulatory homolog A) after disruption of chromatin by the transcription machinery. Recently H3.3 has also been found to be enriched in silent chromatin, opening its role up to debate (Szenker *et al.*, 2011). The final major variant of H3 is CENP-A, found at centromeres and vital for stable centromere and kinetochore formation. It is thought to be deposited by the histone chaperone RbAp48 (Retinoblastoma Associated protein 48 kDa, discussed in more detail in section 1.5.2.3) in a replication independent manner (Hake and Allis, 2006; Smith, 2002; Furuyama *et al.*, 2006).

Within the H2A/H2B dimer, H2B does not have any major variants but several versions of H2A have been discovered. H2A.X is a damage associated histone deposited by INO80 (see section 1.2.5) and is found at sites of chromatin remodelling. H2A.X is phosphorylated at sites of double strand breaks and this event recruits the DNA repair machinery (Henikoff *et al.*, 2004; Talbert and Henikoff, 2010; Pusarla and Bhargava, 2005). The H2A.Z variant is essential in animals but not in *S. cerevisiae*, it is found at most gene promoters and has divergent roles in both activation and silencing of genes. It is added to the nucleosome by replacement of a H2A/H2B dimer by Swr1 (Henikoff and Ahmad, 2005). H2A.Z is enriched near transcription start sites where it defines nucleosome positioning as it is present in the +1 and -1 nucleosomes around the start site (Mito *et al.*, 2007; Guillemette *et al.*, 2005) (see section 1.2). The role of H2A.Z in gene silencing is thought to be a contribution

to heterochromatin packing via interactions with HP1 (Pusarla and Bhargava, 2005). The mechanism by which H2A.Z fulfils both these functions is currently under debate. Other variants of H2A include macroH2A, found in constitutive heterochromatin on the inactive X-chromosome and H2A-Bbd, a vertebrate specific variant associated with active chromatin (Henikoff and Ahmad, 2005).

Histones are subject to a variety of post-translational modifications, thought to be the basis of epigenetic inheritance, determining gene regulation. Known modifications include methylation, acetylation, phosphorylation, ubiquitination, sumoylation, ADP ribosylation, deamination and proline isomerisation, with the first three being the best studied (Kouzarides, 2007). A few examples of how modifications of histones can affect chromatin are discussed below.

The acetylation state of histones is mediated by histone acetyl transferases (HATs) and histone deacetylases (HDACs) which specifically modify lysine (K) residues in the core histones tails. This modification is generally correlated with activation of transcription. Some acetylated lysines function by recruiting other proteins via recognition by a bromodomain, however acetylation also results in a change in the overall charge of histones, which can loosen the structure of chromatin and increase DNA accessibility (Anderson *et al.*, 2002). The acetylation of histone H3 at K56 by Rtt109 in *S. cerevisiae*, seen on many newly synthesised histones, is correlated with activation of transcription (Williams *et al.*, 2008; Masumoto *et al.*, 2005) and DNA damage repair (Masumoto *et al.*, 2005). This modification is unusual as it is within the core of H3 and increases nucleosome breathing (Neumann *et al.*, 2009), which will allow easier access to DNA for the transcription and DNA repair machinery (see section 1.2.4). Modification of a more canonical site is that of histone H4 K16 where acetylation inhibits nucleosome compaction as well as preventing the action of the chromatin remodeller ACF (Shogren-Knaak *et al.*, 2006). Acetylation is also an important marker of histone age with newly synthesised histones acetylated at H4 K5 and H4 K12 (Sobel *et al.*, 1995), this modification is added by the HAT1 complex, consisting of the chaperone RbAp46 and the HAT1 acetyltransferase (Jasencakova *et al.*, 2010).

Methylation was originally regarded as a non-reversible mark of silencing. However with further investigation revealing the existence of demethylases, methylation has been shown

to have a mixed role in gene expression, which is dependent on position and environment (Bannister and Kouzarides, 2005). Methylation can occur at lysine and arginine residues with mono, di and tri methylation possible on lysine residues and mono, di-symmetric or di-asymmetric methylation on arginine residues. Recognition of methylated histone tails occurs via chromodomain, PHD and MBT containing proteins (Bártová *et al.*, 2008). Repressive methylation is exemplified on the inactive X-chromosome where trimethylation at H3 K27 and dimethylation at H3 K9 are common, along with DNA methylation. H3 K9 is known to recruit HP1, an essential component of heterochromatin (Chambeyron and Bickmore, 2004). The combination of acetylation at H3 K9 and dimethylation at H3 K4 are activating modification which causes chromatin decondensation allowing access by transcription machinery (Bártová *et al.*, 2008).

Phosphorylation of serine and threonine residues is mediated by various kinases, including touseled-like kinase (tlk) which is involved in phosphorylation of H3 S10 (Li *et al.*, 2001). Generally considered to be an activatory mark, phosphorylated tails are recognised by 14-3-3 domain containing proteins (Macdonald *et al.*, 2005). Phosphorylation changes the charge on the histone molecule and may have a direct impact on chromatin condensation. The role of this modification in chromatin is complex with phosphorylation of serine 10 required for chromatin condensation during segregation and cell division (Wei *et al.*, 1999) and activation of certain genes (Kouzarides, 2007). Phosphorylation of H3 S10 is inhibited by the repressive H3 K9 methylation mark (Rea *et al.*, 2000). Conversely where present H3 S10 phosphorylation reduces the binding of HP1 to H3 K9 (Fischle *et al.*, 2005), which destabilises heterochromatin, possibly allowing access to S-phase proteins.

All these modifications are essential for the maintenance of correct gene expression and therefore when DNA is replicated the marks on histones must also be retained to maintain gene expression patterns and prevent aberrant protein expression.

1.2 Nucleosome Positioning

The overall pattern of nucleosomes on DNA is not simply a repeating unit along the whole molecule. In recent years, genome-wide studies of nucleosome locations have confirmed previous results and trends in nucleosome binding patterns.

The position of nucleosomes was mapped in *S. cerevisiae* (Yuan *et al.*, 2005; Lee *et al.*, 2007) using micrococcal nuclease (MNase) to digest linker DNA and microarrays to identify DNA sequences that were protected by nucleosomes (i.e. those that remained after digestion). This technique has been refined and Solexa sequencing has allowed the investigation of larger genomes such as human cell lines (Schones *et al.*, 2008). Whilst the data obtained is reproducible between methods, the reliance on MNase to generate mononucleosomes is known to introduce a bias towards terminal A/T bases cleavage sites and some promoters are thought to be unstably bound to histones under experimental conditions (Arya *et al.*, 2010). Although other methods are still required to confirm these observations, there are general trends, discussed below.

The nucleosome content of DNA is non-uniform. Nucleosomes on a length of DNA can be either well positioned, protecting a specific strand of DNA from degradation by MNase or "fuzzy", where a less specific protective effect is observed. In *S. cerevisiae* nearly 70% of nucleosomes are well positioned (Lee *et al.*, 2007) with the percentage seen to be lower in multicellular organisms (Johnson *et al.*, 2006; Schones *et al.*, 2008; Mavrich *et al.*, 2008b).

Specific regions of DNA exhibit characteristic nucleosome positioning patterns. Telomeric DNA tends to be nucleosome depleted, whereas centromeres exhibit higher occupancy. Coding regions of DNA tend to be highly occupied by nucleosomes, with the exception of the most highly transcribed genes such as the ribosomal subunits (Yuan *et al.*, 2005; Lee *et al.*, 2007; Bernstein *et al.*, 2004).

The pattern of nucleosomes in the promoter region of a gene tends to indicate the basal level of transcription. Activity transcribed gene promoters contain a nucleosome free region followed by a very strongly positioned nucleosome, known as the +1 nucleosome. The exact

position of this nucleosome in relation to the transcription start site varies between organisms and may be related to the composition of the transcription machinery (Mavrich *et al.*, 2008a; Yuan *et al.*, 2005). Stress response genes such as those controlled by TATA boxes, (which are usually repressed) do not contain this nucleosome free region and require other proteins to remove the nucleosomes to allow binding of the transcription machinery and activation of the gene (Yuan *et al.*, 2005).

Initial observations of the periodicity of histones led to the mathematical prediction that with one boundary point and a set linker length one would see a string of positioned nucleosomes with increased "fuzziness" as distance from the boundary point increases(Kornberg and Stryer, 1988). This is the case with many nucleosomal arrangements in the genome, exemplified by the +1 nucleosome at genetic promoters, indicating that many nucleosomes are not positioned by specific elements, rather by their interactions with the surrounding nucleosomes (Yuan *et al.*, 2005). More recently, strongly positioned nucleosomes have been found at the 5' end of some genes and are thought to fulfil similar functions to those at the 3' end (Mavrich *et al.*, 2008b). This specific nucleosome positioning is thought to depend on several factors which work in concert to position nucleosomes, discussed below.

1.2.1 DNA Sequence and Modifications

Whilst nucleosomes will bind to any section of DNA, they do have different affinities for different nucleotide sequences. Work reconstituting nucleosomes on synthetic DNA showed that long runs of poly A/T have a lower affinity for nucleosomes compared to G/C or mixed elements. In the latter cases, it is thought that the sequences are rendered less flexible, which makes wrapping around nucleosomes unfavourable and generally destabilises nucleosome binding. This is thought to contribute to creating the nucleosome free region seen at constitutively active gene promoters (Nelson *et al.*, 1987; Segal and Widom, 2009a).

The sequence of the DNA bound to the nucleosome has also been shown to affect its position. This is due to the requirement for DNA to bend sharply around the nucleosome, with the grooves in the double helix compressed on the inside of the turn (Travers *et al.*, 2009).

The presence of an A/T dimer every 10 bp, seen in many nucleosomes, creates a flexible sequence that easily winds around the nucleosome and centres the nucleosome on an AT dimer (Thaström *et al.*, 1999; Johnson *et al.*, 2006). The preference for a flexible DNA sequence is also seen in the binding of histone H1, indicating a role in higher order chromatin structure as well as nucleosome positioning (Cui and Zhurkin, 2009). This specification is thought to contribute to strongly positioned nucleosomes such as the +1 nucleosome.

This reduction of DNA flexibility is also seen with DNA methylation (Nathan and Crothers, 2002), although as proteins containing MBDs (methyl binding domains) such as MeCP2 are known to bind to methylated DNA (Lewis *et al.*, 1992), this modification may have a secondary effect at reducing histone occupancy.

1.2.2 Histones and Higher Order Structure

The histone composition of the nucleosome is thought to influence DNA binding (Arya *et al.*, 2010). However nucleosomes containing H2A.Z have been observed to position in a similar manner to normal H2A containing nucleosomes (Mito *et al.*, 2007). The details of this and any specific sequence preferences remain to be investigated. It is possible that the modifications seen on histones will have some effect on the nucleosome spacing, however as these modifications function to recruit effector proteins, it is more likely that any change in positioning is due to protein binding (Segal, 2008).

It is also possible that the 3-dimensional structure of chromatin influences the histone arrangement, however dissecting cause from effect is particularly difficult here, as the distance between nucleosomes may influence the 3D structure as well as *vice versa*, making interpretation of results difficult.

1.2.3 DNA Binding Proteins

Nucleosomes are not the only found protein bound to DNA. Transcription factors often wrap around the DNA helix sterically excluding nucleosomes. Many transcription factor binding

sites occur in the nucleosome free region of gene promoters (Yuan *et al.*, 2005; Segal and Widom, 2009a), although some are known to bind to regions normally associated with nucleosomes. To do this they must access DNA normally bound to nucleosomes, moving or displacing them in response to cell signalling pathways, which can drastically change the pattern of nucleosome binding within a cell following gene activation (Schones *et al.*, 2008).

Once a transcription factor has bound to a previously nucleosomal region of DNA, it can facilitate the binding of further factors. For example the *S. cerevisiae* transcription factor RAP1 is known to be required for activation of HIS4 (involved in the amino acid biosynthetic pathway), by the associated transcription factor GCN4 upon amino acid starvation. The binding of RAP1 creates a nucleosome free region which then allows GCN4 to bind, when alone neither is sufficient to activate the gene. RAP1 is found throughout the yeast genome where it is thought to fulfil a similar function with other transcription factors and is therefore known as a general regulatory factor (Yarragudi *et al.*, 2004; Yu and Morse, 1999; Bernstein *et al.*, 2007). Remodellers are required for the correct positioning of nucleosomes at most promoters, indicating that intrinsic positioning sequences alone are not sufficient (Korber and Hörz, 2004).

In addition to transcription factors the binding of RNA polymerase II (PolII) causes a change in nucleosome arrangement. Binding to promoter regions induces a downstream shift for the +1 nucleosome (Schones *et al.*, 2008; Mavrich *et al.*, 2008a). It is thought that some components of the polymerase pre-initiation complex such as TATA binding protein bind to DNA dissociated during nucleosome 'breathing' and bend the DNA which then perturbs re-binding of the nucleosome (Godde *et al.*, 1995). This is a feedback mechanism as the initial presence of PolII on the DNA is likely to be dependent on nucleosome positioning and modifications (Kireeva *et al.*, 2002).

1.2.4 Accessibility of Nucleosomal DNA

The movement of nucleosomes on the DNA, hiding and revealing sequences, is thought to occur by three main methods: nucleosomal breathing, histone turnover and remodelling of

chromatin by ATP-dependent enzymes.

1.2.4.1 Breathing

The spontaneous partial unwrapping of the DNA from nucleosomes as a consequence of thermal dynamics is a process known as 'breathing' (Blossey and Schiessel, 2011). The DNA partially dissociates from the nucleosome, where it can to bind to transcription factors before it is re-captured by the nucleosome, without the need for ATP-dependent remodelling (Segal and Widom, 2009b).

This breathing mechanism is also thought to create DNA loops, where the nucleosome re-associates with DNA further along the sequence. This has the potential to allow nucleosome translocation without dissociation from the DNA during processes such as RNA transcription permitting the nucleosome to potentially process along the DNA (Anderson *et al.*, 2002).

1.2.4.2 Turnover

Histones are removed and replaced on DNA as a consequence of transcription and replication. Histone chaperones and some remodelling proteins, as discussed below, are capable of moving nucleosomes between strands of DNA as well as sliding them along. This provides another route for access to the underlying DNA and also a route for removal of histone modifications. Studies in *S. cerevisia*e demonstrated that different regions of DNA have different rates of turnover with the highest rates being found in promoter sequences, particularly those under active transcription (Dion *et al.*, 2007; Lee *et al.*, 2004). For example, the *S. cerevisia*e PHO5 promoter loses nucleosomes upon transcription activation after hyperacetylation of the histone tails. It then gains histones from the soluble pool on reassembly rather than sliding nucleosomes around the DNA. This process requires the histone chaperones ASF1 and HIR1 as well as SWI/SNF remodellers and results in more strongly positioned nucleosomes than what would be expected from random reassembly based on DNA sequence alone (Reinke and Hörz, 2004).

Turnover is also high at chromatin boundaries. It is thought that this removes histone modifications, preventing the spread of active or repressed states over the whole chromosome (Mito *et al.*, 2007). The rates of turnover are observed to be lowest in coding regions, and it is possible that histones are briefly disrupted by the passage RNA polymerase but are returned to the same section of DNA rather than being removed and replaced (Dion *et al.*, 2007).

All of these factors mentioned above have an influence on nucleosome positioning and gene accessibility. Current models favour a dynamically regulated system with a mixture of factors contributing to the overall nucleosomal architecture of the cell (Segal and Widom, 2009b). Chromatin remodelling proteins and histone chaperones both play a key role in maintaining the chromatin structure of the cell.

1.2.5 Chromatin Remodellers

When a cell changes its transcription pattern, only a tiny percentage of nucleosomes are moved at the promoters of specific genes (Shivaswamy *et al.*, 2008). Nucleosomes can be translocated or removed from their binding site by remodelling factors. There are several families of remodelling proteins, all of which contain an ATPase subunit and are subcatagorised further based on the presence of additional subunits and domains. The major families include ISWI, SWI/SNF, CHD and INO80 (Blossey and Schiessel, 2011).

Remodellers tend to be recruited to chromatin by specific histone modifications and their function is often modulated by accessory proteins to the complex which can be tissue and time specific. The general mechanism by which the remodellers move nucleosomes along the DNA is thought to be by the generation of a loop of DNA, which moves around the nucleosome, dissociating a few residues at a time, minimising the energetic cost (Lusser and Kadonaga, 2003; Cairns, 2007).

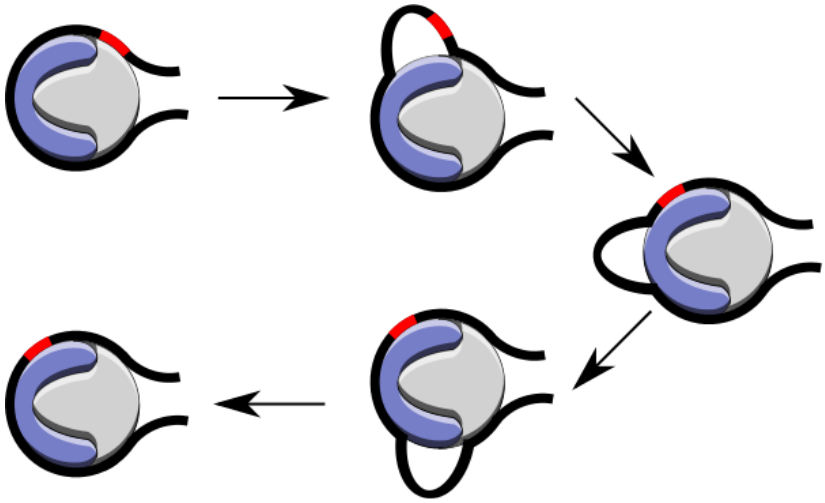


Figure 4: **Model of loop based nucleosome movement along DNA.** The red piece of DNA is translocated around the nucleosome by the 'looping out' of DNA, mediated by a remodeller protein (not shown). Based on information from (Gangaraju and Bartholomew, 2007; Längst *et al.*, 1999)

1.2.5.1 SWI/SNF Family

The SWI/SNF (SWItch/Sucrose Non-Fermentable) family of chromatin remodellers were the first to be discovered. They are characterised by an ATPase domain along with a bromodomain, which binds histones acetylated at lysine residues. SWI/SNF type remodellers are generally associated with activation of genes (Blossey and Schiessel, 2011), tend to randomise linker distances of previously ordered nucleosome arrays (Cairns, 2007) and have the ability to evict nucleosomes by moving them between strands of DNA (Phelan *et al.*, 2000; Lorch *et al.*, 1999).

There are two common SWI/SNF-type complexes. In yeast, they are known as SWI/SNF and RSC (Remodels the Structure of Chromatin). The mammalian equivalents are BAF and PBAF, respectively. These are large multi-subunit complexes with *S. cerevisiae* SWI/SNF containing around 11 subunits and RSC around 17, the exact composition has not been elucidated and appears to vary between cell types (Cairns, 2005; Gangaraju and Bartholomew, 2007).

All SWI/SNF complexes include the catalytic Swi2/Snf2 (hBRM or BRG1 in mammals) subunit and in higher organisms, tissue specific subunits such as BRACA1 which are thought

to program the complex (Lusser and Kadonaga, 2003). SWI/SNF binds transiently to any nucleosome near a promoter when recruited by transcription activators. However, where histones have been acetylated by proteins such as SAGA and NuA4, the presence of SWI/SNF is sustained by the interaction of the bromodomain of SWI/SNF with the histone tails and causes more substantial remodelling of nucleosomes, allowing the access of further transcription regulating proteins (Hassan *et al.*, 2001). Other roles of SWI/SNF have also been examined: the Brm subunit can interact with RNA polymerase II, reducing the rate at which it transcribes DNA, allowing alternative splicing through less optimal sites (Batsché *et al.*, 2006). A role in silencing has also been examined as SWI/SNF is required for the silencing of rRNA and at telomeres. However the exact mechanism of action has not yet been identified (Dror and Winston, 2004).

The RSC complex is more abundant and efficient than SWI/SNF, it contains the Sth1 ATPase and is an essential complex for development (Tang *et al.*, 2010; Lorch *et al.*, 1998, 1999). RSC is required to position nucleosomes that flank the nucleosome free regions of gene promoters (Wippo *et al.*, 2011; Hartley and Madhani, 2009) and is also involved in the loading of cohesins onto sister chromatid arms (Huang and Laurent, 2004).

Both SWI/SNF and RSC are involved in the repair of DNA double strand breaks. During homologous recombination they are recruited at different points in the process, with SWI/SNF involved in strand invasion as part of the recombination process and RSC required after synapsis to finish repair, which may be due to a requirement for cohesin recruitment at the end of homologous recombination (Chai *et al.*, 2005).

1.2.5.2 ISWI Family

The generally repressive ISWI-type complexes contain SANT and SLIDE domains which bind histone tails and DNA respectively (Cairns, 2005; Blossey and Schiessel, 2011). They generally function by sliding nucleosomes into evenly spaced arrays without eviction (Cairns, 2005). As with all remodellers, ISWI complexes are primarily recruited by histone post translational modifications, with ISWI complexes containing a variety of recognition do-

mains, including bromodomains, which recognise acetylated histone H3 tails and PHD domains, which recognise tri-methylated lysines, also on H3 (Lusser and Kadonaga, 2003).

The mechanism by which ISWI remodellers choose where to move a nucleosome to is thought to be based on DNA sequence, with different remodellers acting in different ways. For example the ACF complex has been seen to move nucleosomes to a short (40 bp) sequence with high curvature (Rippe *et al.*, 2007).

In yeast, there are two major ISWI-type proteins: Isw1 and Isw2. Both proteins move histones into intergenic regions, against DNA sequence based positioning preferences (Wippo *et al.*, 2011). Isw1 in general moves nucleosomes onto coding regions and has the ability to disrupt established nucleosomal arrays whereas Isw2 tends to remodel promoter regions. Isw1 has two variant complexes: Isw1a and Isw1b. Both Isw1 complexes contain the same ISWI activity but are seen to bind to different regions of DNA and nucleosomal subunits, with Isw1a having a greater nucleosome sliding activity and Isw1b more efficiently spacing nucleosomes over long stretches of DNA (Tsukiyama *et al.*, 1999; Vary *et al.*, 2003).

In multicellular organisms the variety of ISWI complexes is increased, due to their need to interact with various different proteins. *D. melanogaster* contains only one protein with ISWI activity but serves different purposes in the NURF, ACF and CHRAC complexes. Humans have two major ISWI subunits: Snf2H and Snf2L2, which are involved in a variety of cellular processes including DNA replication, damage repair, nucleosome replacement and centromere formation (Erdel and Rippe, 2011).

The role of the accessory proteins in these ISWI complexes is to determine function, both by targeting and regulating the activity of the ISWI subunit. In the absence of accessory proteins, *D. melanogaster* ISWI catalyses the sliding of nucleosomes towards the ends of DNA. However when ISWI is contained within complexes such as NURF, ACF or CHRAC, the nucleosomes are moved towards the middle of DNA fragments (Eberharter *et al.*, 2001; Vary *et al.*, 2003).

The various multicellular ISWI complexes have different roles in the cell. NURF (Nucleosome remodelling factor), the first ISWI complex to be identified (Tsukiyama and Wu,

1995) is a ternary complex involved in both transcription activation and repression. It consists of ISWI (Snf2L in humans), NURF-55 (RbAp46/48 in humans), BPTF/Nurf301 and NURF-38 (Barak *et al.*, 2003). RbAp48 is a histone chaperone protein found in many chromatin associated complexes (see section 1.5.2.3). BPTF is responsible for control of ISWI activity and targeting of the complex to tri-methylated H3 K4 (marker of transcription start sites) and acetylated H4 K16 *via* its PHD (Plant HomeoDomain) finger and bromodomain. A complex of BPTF and ISWI is sufficient for nucleosome sliding (Xiao *et al.*, 2001). NURF-38 is an inorganic pyrophosphatase, and although this activity is dispensable for chromatin remodelling, it may function in the wider context of NURF's action, possibly helping to drive RNA transcription reactions to completion (Gdula *et al.*, 1998). The binding of transcription factors such as Gal4 to DNA enhances NURF's activity, allowing it to slide nucleosomes over a larger area, revealing RNA polymerase binding sites (Mizuguchi *et al.*, 1997; Hamiche *et al.*, 1999; Kang *et al.*, 2002; Gangaraju and Bartholomew, 2007).

The targeting function of the BPTF domain of NURF is similar to that of the Acf1 subunit seen in the ACF and CHRAC (Chromatin accessibility complex) complexes. Both ACF and CHRAC utilise ISWI/Snf2h to remodel repressive chromatin, forming regularly spaced histone arrays (Eberharter *et al.*, 2001; Gangaraju and Bartholomew, 2007). ACF consists of ISWI and Acf1 and is known to associate with HP1 and increase the affinity of HP1 for methylated H3 K9, targeting both ACF and HP1 to heterochromatic regions (Eskeland *et al.*, 2007). CHRAC consists of both ACF subunits with CHRAC-14 and 15, which have histone-like folds and facilitate nucleosome sliding under conditions where the ACF complex alone is not sufficient (Corona *et al.*, 2000; Kukimoto *et al.*, 2004). The Acf1 targeting protein is also multifunctional: as well as targeting its complex to chromatin, it acts as a binding scaffold for other proteins, modulates the activity of ISWI by increasing ATP turnover and causing nucleosome sliding towards the centre of DNA rather than towards ends (Varga-Weisz *et al.*, 1997; Eberharter *et al.*, 2001; Kukimoto *et al.*, 2004; Längst *et al.*, 1999). ACF/CHRAC is involved in DNA repair, as is the WICH complex.

WICH (WSTF–ISWI complex) is an activatory complex, targeted through its interaction with PCNA (proliferating cell nuclear antigen, see section 1.4.2.1) to newly replicated DNA, where it maintains an open chromatin structure and is involved in post-replication assembly

of chromatin. It consists of Snf2h and WSTF (Williams Syndrome Transcription Factor), which is similar to Acf1 in function (Poot *et al.*, 2004), with the addition of a tyrosine kinase activity that phosphorylates H2A.X Y142 at double strand breaks (Xiao *et al.*, 2009).

As well as the above complexes acting on general chromatin, other ISWI complexes have more specialised activities. The RSF (remodelling and spacing factor) ISWI complex, consisting of Rsf1 and the Snf2h ISWI, it is involved in the maintenance of centromeric chromatin by facilitating the addition and spacing CENP-A containing nucleosomes after deposition of the CENP-A containing nucleosomes (Perpelescu *et al.*, 2009), and increasing DNA accessibility during transcription initiation (LeRoy *et al.*, 1998). The Snf2h ISWI subunit loads cohesins via interaction with the cohesin hRad21 subunit (Hakimi *et al.*, 2002). The ISWI family includes many functional complexes with a general theme of moving histones into ordered arrays, using accessory proteins to modify and target the various complexes to their sites of action.

A further, mixed remodelling factor has been purified (Kitagawa *et al.*, 2003).. WINAC contains the Brg1/hBrm ATPase subunit of SWI/SNF, several BAF subunits from SWI/SNF and RSC, the WSTF subunit seen in the ISWI-type WICH complex and the p150 subunit of the replication dependent CAF-1 complex. This suggests that WINAC is localised to the replication fork and indicates that although these complexes were isolated individually that there may be more interplay between complex subunits in the cell than previously thought

1.2.5.3 CHD and INO80 Families

The CHD (Chromodomain-helicase DNA binding) family contains one or more chromodomains in addition to the ATPase domain (Delmas *et al.*, 1993). The chromodomain is seen in many heterochromatin associated proteins such as HP1, however CHD-1 has both positive and negative effects on repression, with a preference for moving nucleosomes onto low-affinity AT-tracts (see section 1.2). CHD-1 is located to very active sites of replication in Drosophila but in contrast to SWI/SNF has not been seen to remove nucleosomes from large areas of chromatin (Stokes and Perry, 1995; Stokes *et al.*, 1996; Tran *et al.*, 2000).

Additionally, the CHD family member Hrp1 has been shown to be involved in transcription termination, in contrast to the ISWI family which is involved in initiation events (Alén *et al.*, 2002). CHD family proteins are found in the Mi-2 and NURD complexes which are associated with histone deacetylation and both contain the RbAp46/48 histone chaperones and histone deacetylases (Zhang *et al.*, 1999; Wade *et al.*, 1998).

The INO80 family of remodellers differ from the other families as they are 'split' ATPases. As with other chromatin remodellers, these 14-15 subunit complexes are involved in all aspects of chromatin remodelling, replication, repair and transcription (Gangaraju and Bartholomew, 2007) The *S. cerevisiae* INO80 family protein SWR1 has been shown to replace the H2A/H2B dimer with H2A.Z/H2B dimer (Htz1/H2B in yeast) in a replication independent manner (Mizuguchi *et al.*, 2004).

The central theme of chromatin remodellers is the utilisation of one piece of central machinery - the catalytic subunit - to fulfil a variety of roles with the addition of modifying factors which are either tissue or function specific to target the complex and modulate its activity. There is a high degree of functional redundancy between remodellers, indicating that whilst they are seen to be discrete complexes they have overlapping roles within the cells (Tsukiyama *et al.*, 1999; Havas *et al.*, 2001). Remodellers are involved in all aspects of a nucleosome's life on DNA, from assisting in the incorporation into DNA after deposition (discussed further in section 1.3), moving nucleosomes to sites based on epigenetic patterns to eviction from DNA.

Histone chaperones, in contrast to chromatin remodellers, guide histones around the cell as well as functioning in eviction and deposition, they are discussed below.

1.3 Histone Chaperones

Histones are highly charged proteins and prone to aggregation. To prevent non-specific, deleterious aggregation, cells have developed a system of chaperones which bind histones when they are not attached to DNA, guiding their placement and preventing unwanted interactions. Histone chaperones have both histone binding and assembly properties, they generally chaperone a distinct histone pair (H2A/H2B or H3/H4) and many are known to deposit histones directly onto DNA as well as facilitating the interaction of histones with modifying proteins. Some histone chaperones have remodelling activity or interact with chromatin remodellers, allowing a dynamic system of histone assembly, sliding and removal to provide access to regions of DNA for transcription and replication.

1.3.1 H2A/H2B Chaperones

H2A/H2B dimers are found bound to Nap1 in the cytoplasm and during nuclear import. Generally cytoplasmic, Nap1 and its cargo relocate to the nucleus during replication, providing a fresh pool of histones for assembly (Ito *et al.*, 1996; Ransom *et al.*, 2010; Zlatanova *et al.*, 2007). Nap1 from *S. cerevisiae* has also been seen to act as a nucleosome remodeller *in vitro*, where the removal of H2A/H2B allows histone sliding, although the human equivalent does not exhibit this activity. Human Nap1 appears to act just as a chaperone and can exchange the histone variant H2A.Z/H2B for canonical H2A/H2B (Park *et al.*, 2005). NAP1 has also been seen to chaperone linker histones and can deposit H3/H4 onto DNA and is capable of assembling nucleosomes *in vitro*, although its role in nucleosome assembly *in vivo* remains under investigation (Zlatanova *et al.*, 2007).

FACT, a dimer of hSpt16 and SSRP1, is also seen to chaperone H2A/H2B during DNA replication and transcription, where it is thought to aid the progression of the MCM helicase or RNA polymerase by removing nucleosome 'blocks' ahead of the replication fork, acting as both a chaperone and a remodeller (Rocha and Verreault, 2008; Tan *et al.*, 2006; Belotserkovskaya *et al.*, 2003).

Histone chaperones are thought to work utilising thermodynamic principals for nucleosome assembly. In the case of yeast NAP1 it has been shown that tetrameric H3/H4 has a higher affinity for DNA than for NAP1 (Andrews *et al.*, 2008). NAP1 has a slightly higher affinity for H2A/H2B and prevents unwanted interactions between these dimers and DNA, encouraging nucleosome formation (Andrews *et al.*, 2010).

1.3.2 H3/H4 Chaperones

The histone H3/H4 dimer has three major chaperones, ASF-1, HIRA and CAF-1. ASF1 is the major chaperone in eukaryotic cells, sequestering the H3/H4 dimer in the cytoplasm, in a similar manner to Nap1, and allowing presentation of the histone dimer to modifying enzymes such as Rtt109 and Hat1. The CAF-1 and HIRA chaperones have complementary roles. CAF-1, discussed in detail in section 1.5 deposits H3.1/H4 in a replication dependent manner, whereas HIRA is replication independent and deposits H3.3/H4 during nucleosome turnover outside replication (Tagami *et al.*, 2004).

1.3.2.1 ASF1

Originally identified as an S-phase expressed gene that de-repressed silenced loci in yeast following over-expression (Le *et al.*, 1997) Anti Silencing Factor 1 (ASF1, also known as CIA1) (Munakata *et al.*, 2000) chaperones all H3/H4 dimers not associated with DNA within a cell (Kaufman *et al.*, 1997; English *et al.*, 2005). The complex of ASF1/H3/H4 has some intrinsic chromatin assembly activity (Tyler *et al.*, 1999), but appears to act mainly as a histone chaperone for further complexes. ASF1 is essential in animals as it is required for chromatin assembly during DNA replication, transcription and after double strand breaks (Chen *et al.*, 2008).

ASF1 is involved in a variety of histone processing steps. It presents the newly synthesised H3/H4 dimer to Rtt109 (p300/CBP in animals) and HAT1 for acetylation at H3 K56 and H4 K12 respectively (Avvakumov *et al.*, 2011; Miller *et al.*, 2008; Parthun *et al.*, 1996). ASF1

is thought to be involved in both chromatin disassembly, removing parental histones ahead of the replication/translation fork and assembly after the fork passes as it interacts with the MCM helicase and RF-C, found at the front of the replication fork (See section 1.4.2.1) (Groth *et al.*, 2007b; Sanematsu *et al.*, 2006; Adkins and Tyler, 2004) as well as both CAF-1 p60 subunit and HIRA, found behind (Malay *et al.*, 2008; Linger and Tyler, 2005; Eitoku *et al.*, 2008). ASF1 has been shown to bind to the site of the H3/H3 interface in the H3/H4 tetramer which would allow it to disrupt the nucleosome during replication (Natsume *et al.*, 2007).

ASF1 binds a H3/H4 dimer (English *et al.*, 2006). The H3/H4 dimer is seen bound to the side of one ASF1 molecule at the opposite side to helix-1 of histone H4 seen bound in the p46 and p55 crystal structures (Murzina *et al.*, 2008; Song *et al.*, 2008). This indicates a mechanism that would allow both ASF1 and p48 to interact with the histone dimer concurrently. It is thought that ASF1 brings the histone H3/H4 dimer to sites of DNA replication where it passes the histones to an assembly chaperone, such as HIRA or CAF-1 which then deposits the histones onto the DNA (Avvakumov *et al.*, 2011; Corpet and Almouzni, 2009; Mello *et al.*, 2002; Ray-Gallet *et al.*, 2007).

1.3.2.2 HIRA

HIRA deposits histone H3.3/H4 dimers onto DNA in a replication independent manner. The HIRA complex consists of HIRA, ASF1, CABIN1, UBN1 and UBN2 (Rai *et al.*, 2011) with HIRA having homology to Hir1 and Hir2 complexes in *S. cerevisiae*. The HIRA protein functions as a scaffold to recruit the other complex members, (along with ASF1), which is thought to provide the histones for the complex, binding *via* HIRA's C-terminal tail (Tang *et al.*, 2006). The other subunits are also recruited through their interaction with HIRA (Rai *et al.*, 2011). The interaction of HIRA with histones can be modulated by modifications: phosphorylation of histone H4 S47 promotes the interaction of the ASF1 bound H3.3/H4 dimer with HIRA whilst decreasing the affinity for CAF-1 (Kang *et al.*, 2011).

HIRA has been shown to posess several functions beyond chromatin assembly in mammals. The Hir complex in *S. cerevisiae* acts as a transcriptional repressor of genes such as histones

H2A and H2B, is involved in the formation of yeast heterochromatin and also activates other genes via recruitment of the SWI/SNF remodelling complex (Green *et al.*, 2005; Prochasson *et al.*, 2005). In mammals, HIRA has been shown to silence expression of histones H3.1 and H3.2 (Rocha and Verreault, 2008). In *S. pombe* the HIRA homologue Hip1B is involved in silencing genes and retrotransposons (Rai *et al.*, 2011).

HIRA is a seven beta-propeller protein and interacts with the similarly structured chaperone RbAp48 which allows recruitment of histone deacetylases HDAC1 and 2 (Ahmad *et al.*, 2004) and HP1 (Yamane *et al.*, 2011). These interactions indicate a repressive role for HIRA in multicellular organisms.

The multiple roles seen for HIRA in the cell illustrate the complex interplay between chaperones and other components of the chromatin assembly and modification system. These all combine to assemble chromatin in either a replication dependent or independent manner, as described below.

1.4 Assembling Chromatin

The assembly of chromatin is the result of a network of histone chaperones, chromatin remodellers and enzymes that modify this structure. Chromatin is assembled either after DNA replication or independent of replication, after events such as transcription or during histone exchange. In general, histone chaperones provide and deposit the histones onto DNA and a combination of intrinsic signals and remodelling proteins order nucleosome positioning, whilst enzymes recruited by the assembly complex and existing histone modifications reconstitute the pattern of modifications. The exact mechanism by which this is done is still under debate although some pathways are known for the different forms of assembly.

1.4.1 Replication Independent Assembly

Replication independent assembly encompasses a variety of events. Histone exchange is thought to happen on a localised scale but the pathways have not been well characterised, it is thought some exchanges are in response to signals such as DNA damage (Jin *et al.*, 2005; Das and Tyler, 2012). However the major driver of histone replacement and assembly outside of replication is the response to transcription.

During transcription the RNA polymerase complex requires the movement of nucleosomes to allow passage along and reading of DNA. Both the H2A/H2B dimer and the H3/H4 dimer have the potential to be replaced during these events (Jackson, 1990) and this can result in a mixture of old and new histones incorporated into a single nucleosome, possibly making the re-establishment of epigenetic marks easier (Xu *et al.*, 2010) as well as incorporation of histone H3.3, marking actively transcribed chromatin.

There are two methods by which RNA polymerase is thought to negotiate the nucleosome. Firstly the nucleosome is not totally disassembled: RNA polymerase can negotiate nucleosomes with the loss of one dimer of H2A/H2B and no movement of the nucleosome, described in *in vitro* systems (Kireeva *et al.*, 2002). This may account for the higher rate of exchange observed for the H2A/H2B dimer compared with H3/H4 (Jackson, 1990). There are

several mechanisms by which the loss of the H2A/H2b dimer is thought to be accomplished. The whole nucleosome presents a block to the passage of RNA polymerase II which has been shown to be removed by the addition of FACT, which allows transcription elongation without the need for ATP hydrolysis (Orphanides *et al.*, 1998). The chromatin remodeller RSC has also been shown to remove this block on recognition of acetylated histones, requiring hydrolysis of ATP and indicating a mechanism by which acetylation increases gene expression (Carey *et al.*, 2006).

More actively transcribed genes show greater loss of the H3/H4 tetramer than those that are less activly transcribed. It is thought that whilst RNA polymerase II can negotiate a reduced nucleosome, (where increased speed is required) the whole nucleosome is disassembled and reassembled behind the complex (Das and Tyler, 2012). Nucleosome eviction is also mediated by a variety of complexes. The remodeller RSC, in the presence of the histone chaperone Nap1 is capable of complete, ATP-mediated disassembly of the nucleosome (Lorch *et al.*, 2005). The remodeller Chd1 interacts with HIRA during *D. melanogaster* development to incorporate H3.3 into chromatin (Konev *et al.*, 2007). ASF1 is known to mediate both the eviction and deposition of H3/H4 (Del Rosario and Pemberton, 2008; Schwabish and Struhl, 2006). HIRA (HIR1) and ASF1 appear to act in concert as in *S. cerevisiae* where ASF1 re-incorporates existing histones onto the DNA and HIR1 adds newly synthesised histones (Kim *et al.*, 2007).

Given the high rate of retention of old nucleosomes by the DNA it has been suggested that the nucleosome components are localised to the site of eviction by interaction with the transcribed RNA. H2A/H2B have a high affinity for RNA (20 times its affinity for DNA) and on eviction by Nap1 may be temporarily deposited onto RNA before reassembly behind the transcription complex. This also provides a way to localise H3/H4 to the fork if it is evicted despite its lower affinity for RNA, through binding to H2A/H2B (Levchenko and Jackson, 2004). The interaction between histone chaperones and chromatin remodellers allows gene transcription to take place in a controlled manner independent of DNA replication.

Other replication independent assembly events that have been studied indicate a close association between chaperones and remodellers. The the addition of CENP-A to nucleosomes

near the centromere is a two-step process where deposited CENP-A interacts only weakly with the nucleosome until it is remodelled by RSF into a robust complex (Perpelescu *et al.*, 2009). In vitro studies have also shown that, in concert with NAP1, the ACF remodelling factor is required for the efficient deposition of histones onto DNA (Ito *et al.*, 1999). Histones are deposited as a non-nucleosomal intermediate complex by NAP1, thought to be the nucleosome proteins without the DNA properly positioned, requiring the action of a chromatin remodeller to form the final stable complex utilising an ATP driven process (Torigoe *et al.*, 2011). It is likely that a similar process is seen at all sites of chromatin assembly as remodelling proteins have been shown to be involved in all aspects of nucleosome turnover.

1.4.2 Replication Coupled Assembly

1.4.2.1 DNA Replication

Replication of DNA is a highly regulated process. To prevent multiple rounds of replication from the same point, cells 'licence' the many origins of DNA replication in a cell during the G1 growth phase. This involves formation of the pre-replication complex by the ORC (origin recognition complex) which recruits Cdc6 and Cdt1 followed by MCM helicase to form the licensing complex, which remains on origins of replication until S-phase. This step is heavily regulated by a combination of inhibitor protein and modifications to prevent the re-initiation of fired origins (Li and Jin, 2010).

The pre-replication complex is then activated during S-phase in a step thought to involve phosphorylation of the complex by the kinases Cdc7-Dbf4 and CDK2 (Talbert and Henikoff, 2010), reconfiguration of the MCM helicase and recruitment of other replication factors including Cdc45, GINS and DNA polymerase α. This complex then contains the helicase and primase required for DNA synthesis (Diffley, 2011).

During DNA synthesis, DNA polymerase α acts as a primase, generating a RNA primer followed by a stretch of DNA. This primer is then bound by RF-C which loads PCNA, a sliding clamp protein discussed below, onto the primer and displaces polymerase α. The loss

of polymerase α allows the binding of DNA polymerase ε, an enzyme with strong helicase and proofreading activities which can then accurately synthesise DNA along the leading strand template (Burgess *et al.*, 2010).

DNA polymerases can only synthesise DNA in a 5'-3' direction, where the leading strand processes in this direction, but the lagging strand requires synthesis in the opposite direction to the replication fork. To do this DNA polymerase α repeatedly synthesises short primers which are each recognised as before by RF-C, permitting the loading of PCNA and recruitment of DNA polymerase δ. This last polymerase proofreads the primer and acts as the elongation polymerase. To remove the multiple sections of RNA template either FEN1 or Dnase1 is recruited. Once synthesis has finished DNA ligase joins the synthesised fragments together (Burgers, 2009). Importantly, PCNA remains on the DNA after synthesis and couples DNA replication to chromatin assembly (Naryzhny *et al.*, 2005; Shibahara and Stillman, 1999).

PCNA is a homotrimeric sliding clamp protein (Krishna *et al.*, 1994) (Figure5) which increases the processivity of DNA polymerase along DNA. Loaded by RF-C onto the DNA primer template in the presence of ATP, it is conserved in structure across all species. PCNA acts as a 'landing pad' to recruit proteins associated with various aspects of the DNA replication and repair processes via a conserved motif known as a PIP (PCNA interaction peptide) (Warbrick, 2000). DNA polymerase δ p66 subunit, the FEN1 RNAse and DNA ligase all contain PIP motifs, linking the major players in lagging strand synthesis (Bruning and Shamoo, 2004). The pre-initiation component Cdt1 is recruited to PCNA where it is ubiquitinated and subsequently degraded, preventing re-initiation of replication (Diffley, 2011). PCNA also plays a vital role in chromatin reformation, binding CAF-1 and a variety of DNA and histone modifying proteins including the DNMT1 methyltransferase, HDAC1 deacetylase and the nucleosome remodeller WSTF (Moldovan *et al.*, 2007). It has been suggested that PCNA may act as a double trimer to allow simultaneous binding of DNA polymerase and CAF-1 so the processes of DNA replication and chromatin reformation are combined (Naryzhny *et al.*, 2005).

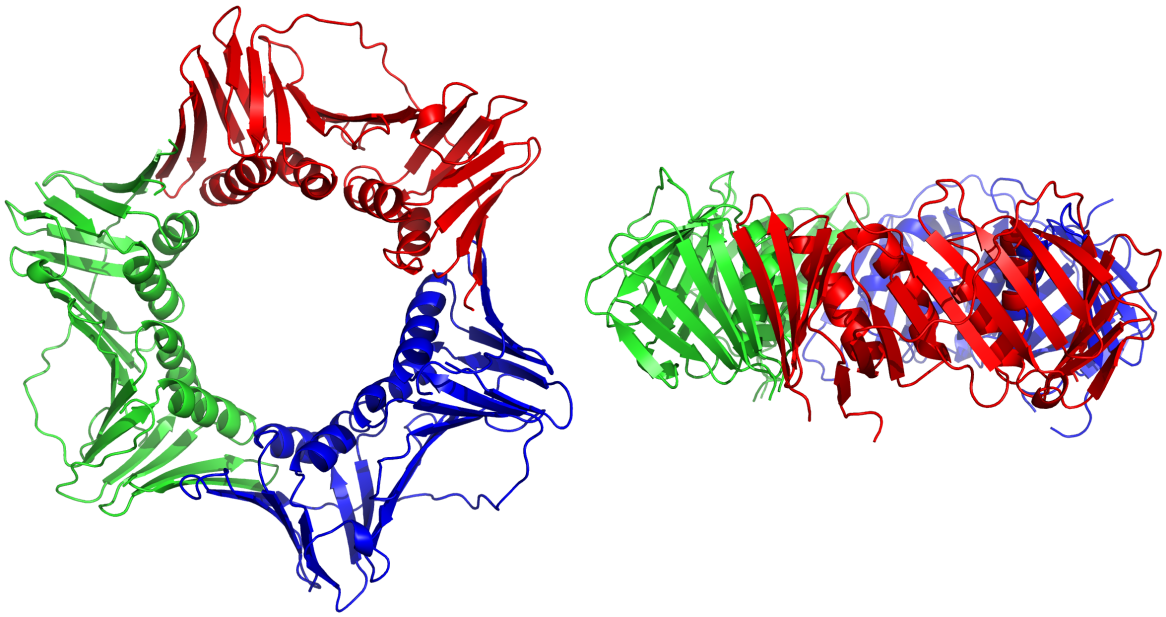


Figure 5: **Structure of the Human PCNA Homotrimer**.
Cartoon representation of the structure of PCNA. Each monomer is coloured red, blue or green. The proteins assembles around DNA which moves through the central gap. Rendered in PyMol (Gulbis *et al.*, 1996).

1.4.2.2 Chromatin Replication

During replication independent assembly, nucleosomes must be totally removed from the DNA as the replication complex passes, deposited onto both daughter strands, the arrangement of nucleosomes re-established and modifications of histone tails repeated. Whilst DNA is replicated in a semi-conservative manner, with a single original strand retained by each new molecule, the same does not appear to be true for nucleosomes. It is known from early studies that both new and old histones are incorporated onto both strands (Sogo *et al.*, 1986; Jackson, 1988) and that the H3/H4 tetramer is not split , resulting in the paradox discussed in section 1.4.2.3.

The disassembly process is thought to be caused by a combination of factors including direct disruption by the replication machinery and chaperone/remodeller dependent disassembly. ASF1 and FACT are thought to be involved as they have been seen to interact with the MCM helicase at the front of the fork (Groth *et al.*, 2007a; Rocha and Verreault, 2008; Tan *et al.*,

2006). They are thought to chaperone histones to behind the replication fork where they are deposited either directly or through association with CAF-1.

The process of deposition after replication is known to be mediated by CAF-1. Localised to the replication fork through interaction with PCNA it deposits the H3.1/H4 tetramer (Shibahara and Stillman, 1999). CAF-1 is known to preferentially (but not exclusively) bind to histone H4 when the N-terminal tails are acetylated at lysine 5, 8 and 12 (Verreault *et al.*, 1996), these marks are added by HAT1, and denote newly synthesised histones (Worcel *et al.*, 1978; Sobel *et al.*, 1995; Parthun *et al.*, 1996). This indicates that the systems by which old and new histones are transferred onto replicated DNA may be distinct. However, recent evidence suggests that parental histones may also be acetylated during the removal/replacement process leaving the question as to how CAF-1 recognises new histones open (Burgess *et al.*, 2010). Addition of the H2A/H2B dimers in a sequential manner to complete the nucleosome is thought to be mediated by a combination of the FACT and Nap1 chaperones (Groth, 2009).

Re-establishment of nucleosome positioning is likely to be initially due to intrinsic nucleosome positioning sequences (see section 1.2) and finalised by the action of remodelling enzymes recruited to the fork. WSTF, a component of the WICH and WINAC remodelling complexes is known to interact with PCNA and WINAC has also been seen to interact with CAF-1 (See section 1.2.5.2) (Groth *et al.*, 2007b). It is possible that WINAC acts with CAF-1 in a similar manner to the interaction of ACF with NAP1, maturing the deposited histone proteins into nucleosomal form.

Further to the deposition and arrangement of histones following the replication fork, the epigenetic marks must be reinstated on new nucleosomes. Both PCNA and CAF-1 remain on new DNA for a significant amount of time which offers a window of recruitment for histone modification activities (Taddei *et al.*, 1999). The re-establishment of methylated DNA is mediated by DNMT1 which uses hemimethylated DNA as a template and is recruited to the replication fork by PCNA (Warbrick, 2000). This methylation can then act as the template for the re-establishment of H3 K9 methylation. MBD1 binds to methylated DNA and recruits the H3 K9 methylase SETDB1. The MBD1-SETDB1 complex also binds to CAF-1 p150 subunits where it could perform the same task during replication dependent assembly without

DNA methylation (Sarraf and Stancheva, 2004). The Sas2 acetyltransferase also interacts with CAF-1 p150 to acetylate H4 K16 after replication (Meijsing and Ehrenhofer-Murray, 2001).

Histone modifications are also implicated in their own duplication: Trimethylation of H3 K27 recruits the PRC2 methylase complex during S-phase, where it is thought to methylate the newly synthesised histone neighbours of nucleosomes already showing the mark (Hansen *et al.*, 2008). Methylation of H3 K9 has the potential to act in a similar manner as it recruits HP1 which in turn recruits the Suv39h1 methylase (Maison and Almouzni, 2004).

1.4.2.3 The Assembly Problem

The composition of the nucleosome, consisting of two copies of each protein, suggests two possible methods of chromatin replication. Firstly, a semi-conservative mechanism may exist, wherein the nucleosome is split in half, with each dimer separated onto a different strand and a new histone dimer added. Secondly, conservative replication moving the whole nucleosome from the front of the replication fork to behind it, with or without disassembly but maintaining the nucleosome as a unit. During conservative replication, the nucleosomes could be deposited onto either just one strand or onto both, in either an equal or stochastic manner, as detailed in Figure 6 (Martin and Zhang, 2007). Early investigations used pulse labelling experiments to show that whilst the H2A/H2B dimer appears to split from the other H2A/H2B dimer and segregate into nucleosomes with newly synthesised H2A/H2B, the H3/H4 tetramer does not appear to be split during the course of replication, leading to nucleosomes containing either new or old H3/H4 tetramers with H2A/H2B dimers from different sources (Jackson and Chalkley, 1981; Annunziato, 2005).

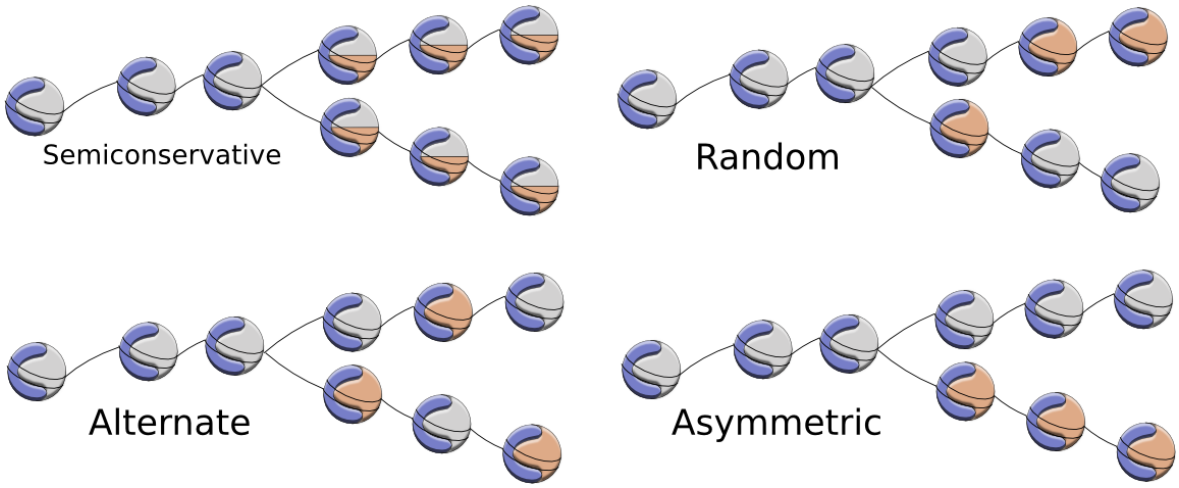


Figure 6: **Possible Chromatin Assembly Strategies**.
Chromatin replication could be either semi-conservative or conservative. There are three possible conservative replication mechanisms, alternate segregation of nucleosomes, asymmetric or random. Current evidence favours the random segregation of nucleosomes. Old histones H3 and H4 are indicated in grey, new histones H3 and H4 in brown. All H2A-H2B dimers are blue.

Recently, this result has been investigated further and a new level of complexity added. Xu and colleagues examined the histone H3 variants H3.1 and H3.3. Investigation of tagged versions of both histones revealed that H3.1/H4 tetramers remain intact on replication, confirming previous results. However, they also discovered that H3.3/H4 tetramers, which are formed in a replication independent manner, are split upon replication, incorporating a newly synthesised H3.3/H4 molecule. This splitting of the tetramer was reduced when replication was inhibited, indicating that it occurs when active chromatin is replicated. Whether this splitting is histone variant specific or rather a result of modifications of the histones is currently unknown (Xu *et al.*, 2010).

Structural evidence of histones H3 and H4 in complex with the chaperone ASF1 has revealed that it binds a dimer of H3/H4 (English *et al.*, 2006), and that ASF1 is capable of disrupting the tetramer to extract the H3/H4 dimer (Natsume *et al.*, 2007). Experiments using epitope tagged H3 to purify H3 containing complexes such as the ASF1/H3/H4 and CAF-1/H3/H4 complexes from the cell did not reveal any untagged H3 protein in tagged H3 containing complexes, despite the tagged protein being expressed at a low level, indicating again that H3/H4 is held in a dimeric form in chaperone complexes in the cell (Nakatani *et al.*, 2004;

Tagami *et al.*, 2004). Recent studies of of the association coefficients of H3 and H4 have also indicated that at the normal concentration seen in the cell they will be present as dimers, with association on DNA due to increased local concentration (Donham *et al.*, 2011).

This information appears to be contradictory; histone tetramers are not split on chromatin replication despite conditions in the cell tending towards dimers of H3/H4 and nearly all histone containing complexes investigated at the molecular level appear to contain only a H3/H4 dimer. How this paradox is resolved is currently under investigation. In this dissertation the CAF-1 complex, known to be involved in deposition of the H3.1/H4 dimer during replication dependent nucleosome assembly is investigated. CAF-1 contains three subunits, all potentially capable of binding histones, at least in a non-specific manner (Kaufman *et al.*, 1995), operating with an unknown oligomerisation state. These experiments aim to dissect the interactions in the CAF-1 complex to provide insight into how it functions. A schematic showing the current understanding of the role of CAF-1 during replication dependent assembly is shown in Figure 7.

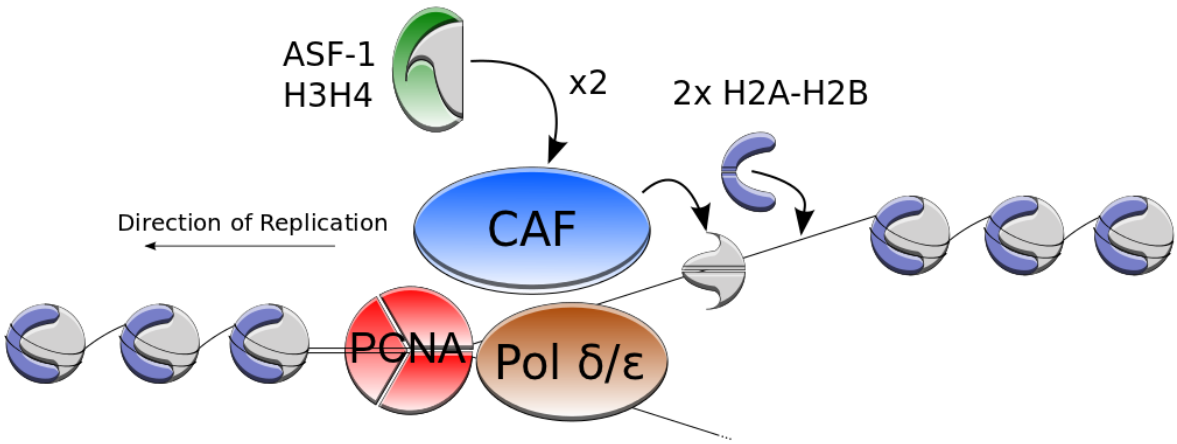


Figure 7: **Schematic of the Role of CAF-1 in Chromatin Formation on Replicating DNA** Following the replication fork and the DNA Polymerase ε *via* interaction with PCNA, CAF-1 deposits the H3/H4 tetramer onto the newly synthesised DNA *via* an unknown mechanism. The nucleosome is completed by the addition of two H2A/H2B dimers. The source for the newly synthesised H3/H4 dimer is ASF1. Old histones are known to be removed from in front of the replication fork with the aid of ASF1, whether CAF-1 is also involved directly in this process in unknown.

1.5 Chromatin Assembly Factor 1

The existence of a replication dependent chromatin assembly factor was first shown in 1986, using the Simian Virus 40 papovavirus (SV40) replication system which packages the 5.2 kb circular viral genome into nucleosomes in a cell free system (Stillman, 1986; Li and Kelly, 1984). This system was later used as an assay to purify this chromatin assembly activity, leading to the discovery of Chromatin Assembly Factor 1 (CAF-1) (Smith and Stillman, 1989).

CAF-1 purified to homogeneity from the 293-T human cell line was shown to consist of three subunits of mass, as estimated by SDS-PAGE, of 150, 60 and 48 kDa and named p150, p60 and p48 respectively (Smith and Stillman, 1989). These subunits, discussed individually later, co-localise to the nucleus at sites of DNA synthesis and are responsible for depositing H3/H4 onto replicated DNA (Verreault *et al.*, 1996; Smith and Stillman, 1989, 1991). The CAF-1 complex binds to either one or two histone H3(.1)/H4 dimers, with a preference for newly synthesised histones (Sobel *et al.*, 1995; Glowczewski *et al.*, 2004). The exact stoichiometry of the complex is currently unknown.

CAF-1 has been identified in all eukaryotic species investigated. *D. melanogaster* (d)CAF-1, consisting of p180, p105 and p55, homologues of p150, p60 and p48 respectively (Kamakaka *et al.*, 1996; Tyler *et al.*, 2001), was purified using a similar method to human (h)CAF-1 and shows functional conservation between species (Song *et al.*, 2007). CAF-1 has also been identified in mouse, *Xenopus laevis* and chicken and has been shown to be essential for development in all these organisms (Quivy *et al.*, 2001; Takami *et al.*, 2007; Houlard *et al.*, 2006).

Both plant and yeast CAF-1 have been identified, although they are not as essential to cell survival as their animal counterparts. *S. cerevisiae* also appears to contain an activity that partially compensates for loss of CAF-1. Whilst mutants appear normal under optimal growth conditions they are sensitive to DNA damaging reagents such as ultraviolet (UV) light and exhibit problems with gene repression (Kaufman *et al.*, 1997; Linger and Tyler, 2005; Enomoto and Berman, 1998).

Mutations of *A. thaliana* FASCIATA 1 (FAS-1), a p150 homologue, cause defects at both stem and root meristems and in other processes requiring cell division, thought to be due to reversion of silenced genes to active states (Kaya *et al.*, 2001). Some FAS-1 mutant lines also show increased numbers of copies of the genome (polyploidy), which is tolerated in plants but not in other organisms, possibly providing an explanation as to why such mutations are lethal to animals (Ono *et al.*, 2006; Exner *et al.*, 2006).

1.5.1 Functions of CAF-1

CAF-1 binds histones H3.1 and H4 to facilitate their deposition onto DNA. Whether the H3/H4 tetramer is constructed within the CAF-1 complex and then deposited onto the DNA or if two dimers are deposited individually is currently unknown. Dimers of H3/H4 have not been observed on DNA (**?**), indicating that if they are assembled after deposition, it is a very fast process. CAF-1 is not unique in binding to the H3/H4 dimer (see section 1.3). However CAF-1's unique feature is that it acts solely on recently replicated DNA (Lassle *et al.*, 1992), making it a marker for actively replicating cells, replication forks and sites of DNA repair requiring synthesis (Polo *et al.*, 2004).

Immunostaining of the p150 and p60 CAF-1 subunits show their localisation to the nucleus (Smith and Stillman, 1991), mainly at replication foci (Krude, 1995). This complex dissociates during mitosis when p60 is hyper-phosphorylated (Marheineke and Krude, 1998). The localisation of p48 is less well defined, it is seen at replication forks but also at other locations, associated with other proteins as part of various chromatin formation and re-modelling complexes (Marheineke and Krude, 1998).

The role of CAF-1 in viral infection has also recently been noted: Human cytomegalovirus (HCMV) recruits CAF-1 to package viral DNA via interaction of p150 with the HCMV IE2 protein which is required for HCMV DNA replication (Lee *et al.*, 2011).

1.5.1.1 Replication

As detailed in section 1.4.2 CAF-1 adds newly synthesised histone H3.1/H4 dimers to replicated DNA. Depletion of CAF-1 causes replication to stall and induces activation of the DNA damage checkpoint pathway, indicating that chromatin formation is critical for a cell undergoing replication of its DNA (Nabatiyan and Krude, 2004; Ye *et al.*, 2003). The specificity of CAF-1 for the replication fork is provided by its interaction with PCNA (Shibahara and Stillman, 1999). CAF-1 can bind PCNA *via* two separate sites on p150. The classical PCNA binding motif (PIP1) between residues 12-34 has the consensus sequence QXXhXXaa where h is a hydrophobic (V/I/L/M) and a is an aromatic (F/Y/W/H) amino acid. The PIP2 site (aa 420-444) contains a substitution of the classical Q to K, which reduces the binding affinity to PCNA but is required and sufficient for assembly of H3/H4 onto DNA (Moggs *et al.*, 2000; Ben-Shahar *et al.*, 2009).

1.5.1.2 DNA Repair

Loss of CAF-1, lethal to animals, results in increased sensitivity to UV and gamma radiation in plants and yeast with increased frequency of somatic homologous recombination events and activation of S-phase DNA damage checkpoints, indicating a possible role in the repair or prevention of DNA damage (Ye *et al.*, 2003; Game and Kaufman, 1999; Endo *et al.*, 2006; Song *et al.*, 2007). Mutation of CAF-1 induces activation of the S-phase and DNA damage checkpoints in *S. cerevisiae*. It is thought that preventing chromatin reassembly leads to replication fork stalling, which when resolved incorrectly creates double strand breaks (Ye *et al.*, 2003).

In quiescent cells, where CAF-1 levels are normally low, the complex is up-regulated in response to damage such as double strand breaks (Nabatiyan *et al.*, 2006) and complementation assays have indicated a role for CAF-1 in the Rad6 mediated repair pathways of *S. cerevisiae* (Martini *et al.*, 1998; Ye *et al.*, 2003; Game and Kaufman, 1999). Subsequent work has shown that CAF-1 is involved in repackaging chromatin during the three DNA

repair mechanisms requiring synthesis: Nucleotide excision repair (NER), homologous recombination (HR) and non-homologous end joining (NHEJ) (Linger and Tyler, 2005; Green and Almouzni, 2003), outlined in Figure 8.

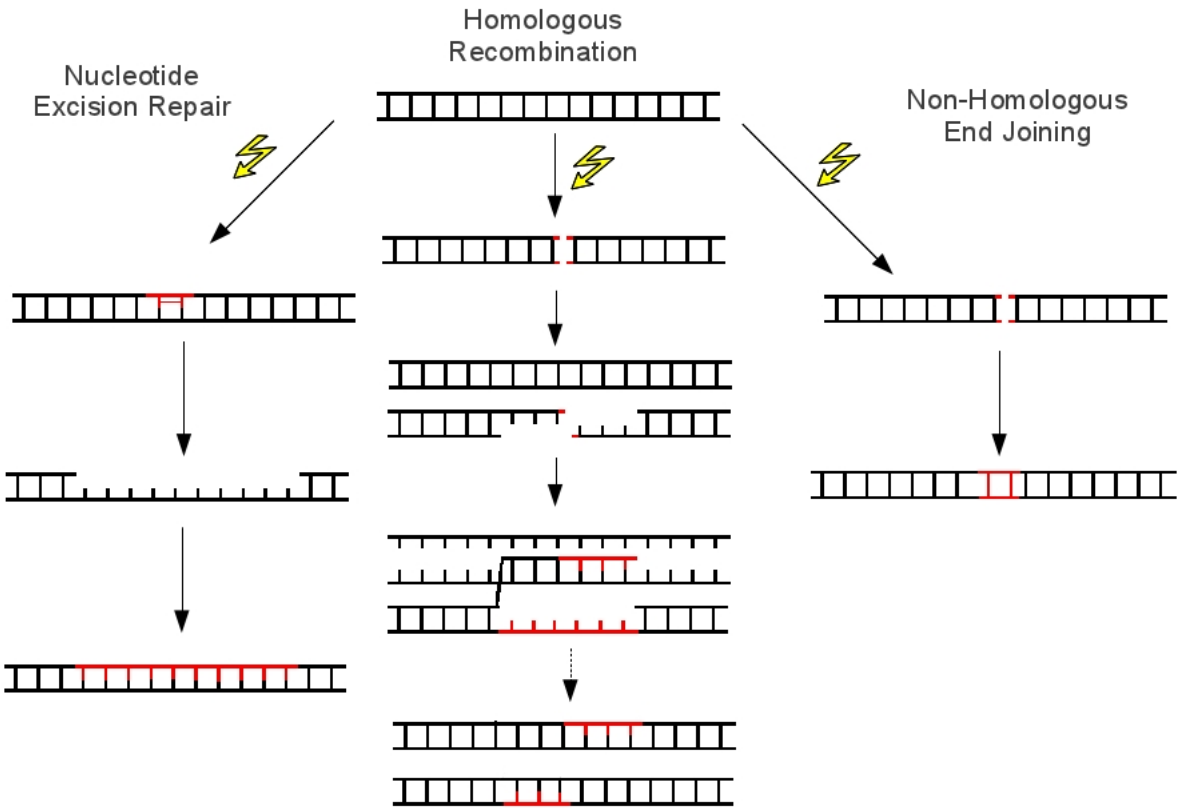


Figure 8: **Schematic Representation of Replication Requiring DNA Repair Mechanisms**.
Newly synthesised DNA is indicated in red. The DNA polymerase machinery allows recruitment of CAF-1 for replication dependent chromatin formation.

NER is used to correct small, helix distorting lesions, such as thymine adducts caused by UV light, cyclobutane pyrimidine dimers and 6-4 photoproducts (Escargueil *et al.*, 2008). Damage recognition of distortions of the DNA backbone is performed by the proteins RPA, XPA and XPC-TFIIH which form a recruitment complex for the nucleases XPG and XPF. TFIIH is a helicase which unwinds the DNA to allow access to the nucleases, which create single strand breaks either side of the lesion (Sancar *et al.*, 2004). The resulting single strand break is then filled by DNA polymerase δ or ε using a recruitment pathway similar to that of genomic DNA replication. This in turn recruits CAF-1 through its interaction with PCNA

where it re-assembles the new DNA into chromatin (Green and Almouzni, 2003).

Double strand breaks are often caused by ionising radiation or as a result of genomic crossover events. Breaks are marked by the phosphorylation of H2A.X which is required for the recruitment of the repair machinery. These include the histone acetyltransferase NuA4 which relaxes chromatin, allowing access to the DNA (Tsukuda *et al.*, 2005; Escargueil *et al.*, 2008). Chromatin is remodelled in response to double strand breaks in an INO80 dependent manner, which results in the loss of nucleosomes from the damaged site, which must then be replaced after damage repair (Tsukuda *et al.*, 2005).

There are two main mechanisms of double strand break repair: HR and NHEJ. During HR a single strand of DNA 'invades' complementary DNA to allow recombination and retrieval of the genetic information from the other strand. This results in a crossover event forming a Holliday junction which is resolved as seen during meiosis. The resolution of the Holliday junction requires DNA synthesis, which is likely to recruit CAF-1 to repair sites (Sancar *et al.*, 2004). If a template strand is not available then the damage can be crudely repaired by ligating the ends together by insertion of a few bases followed by non-specific ligation by XRCC4 ligase4 (NHEJ) (Dudasova *et al.*, 2004; Linger and Tyler, 2005).

The role of CAF-1 in these events is not completely clear. Complexes containing extensively phosphorylated p60 are recruited to sites of UV damage (Martini *et al.*, 1998) in a PCNA dependent manner (Moggs *et al.*, 2000; Okano *et al.*, 2003) and incorporate newly synthesised histones into sites of repair. However, CAF-1 is not essential for activation of the checkpoint or the repair itself and in NER it has been shown that chromatin assembly propagates from the site of repair even when DNA synthesis is inhibited, (Gaillard *et al.*, 1996). CAF-1 is recruited late in the repair process, therefore it is thought to be required for the re-formation of chromatin after damage repair and is likely to be recruited and function in a similar manner to during genomic replication (Polo *et al.*, 2006; Green and Almouzni, 2003).

1.5.1.3 Heterochromatin

In *S. cerevisiae* loss of CAF-1 p150 results in a loss of maintenance of transcriptional silencing, at both genes and telomeres indicating that in yeast CAF-1 is not directly involved in the Sir-dependent establishment of heterochromatin (Enomoto and Berman, 1998; Monson *et al.*, 1997). This silencing can also be mediated by Hir proteins, indicating a mechanism by which *S. cerevisiae* compensates for the loss of CAF-1, reversing the lethality observed in animals (Kaufman *et al.*, 1998). This loss of transcriptional silencing is also seen in *A. thaliana*(Ono *et al.*, 2006), with methylation of DNA acting as a partially redundant pathway (Schönrock *et al.*, 2006).

CAF-1 is also thought to contribute to kinetochore formation at centromeric chromatin during mitosis; loss of CAF-1 p150 and Hir1 in *S. cerevisiae* causes S-phase arrest and a delay in passing from G2 to M phase, coinciding with an increased number of chromosome mis-segregation events and an alteration in the structure of centromeric chromatin (Sharp *et al.*, 2002).

In higher animals CAF-1 is seen to localise to heterochromatic replication foci following the replication of euchromatin (Krude, 1995) and is required for replication of heterochromatin structure (Quivy *et al.*, 2004). Loss of p150 causes arrest in mouse embryos at the 16 cell stage due to alterations in heterochromatin organisation (Houlard *et al.*, 2006), results in defects in nuclear organisation and cell death in *Xenopus* (Quivy *et al.*, 2001) and is hemizygous lethal during *D. melanogaster* development with tissue-specific disruption causing severe developmental defects (Song *et al.*, 2007).

CAF-1 p150 can localise HP1 to heterochromatin via the interaction of the p150 MIR (MOD interacting region) with HP1, an area of p150 dispensable for chromatin assembly function (Murzina *et al.*, 1999). Given the independent activity of the N and C terminal domains of p150 one would expect that H3/H4 and HP1 can be bind concurrently. However when examined *in vivo* whilst CAF-1 was seen to bind HP1 and histones it did not appear to bind both simultaneously (Quivy *et al.*, 2004).

As well as having a role in chromatin formation, it has been suggested that CAF-1 may aid

in the destabilisation of heterochromatin. Depletion of CAF-1 in human cell lines results in cell arrest prior to heterochromatin replication (Hoek and Stillman, 2003) and depletion in *Xenopus* prevents pericentromeric heterochromatin replication, causing cell cycle arrest (Quivy *et al.*, 2008). In *Xenopus* this was shown to be a result of the loss of the interaction between p150 and HP1, indicating a potential role for CAF-1 in the removal of HP1 before chromatin deposition (Quivy *et al.*, 2008).

1.5.1.4 Aberrant Expression

Correct function of CAF-1 is required for gene silencing. Cancer tissue is characterised by aberrant gene expression. It therefore follows that there may be a link between the two. Several studies have associated over expression of p60 with increased severity and poorer outcomes of cancer (Staibano *et al.*, 2007, 2009; Yang *et al.*, 2002; Mascolo *et al.*, 2010). Conversely, depletion of p60 is tolerated by quiescent cells (Nabatiyan and Krude, 2004). p150 is down regulated in a minority of cancers with low concentration associated with poor outcomes (Kong *et al.*, 2007; Pacifico *et al.*, 2007; Staibano *et al.*, 2007). As loss of p150 leads to massive chromosomal defects, a hallmark of late stage cancer, this is not unexpected (Ye *et al.*, 2003; Endo *et al.*, 2006; Song *et al.*, 2007). RbAp48, the p48 subunit, was originally discovered interacting with the tumour suppressor Rb and has been seen to be down-regulated during cancer. The down regulation of p150 and p48 expression during cancer indicates roles for these proteins as tumour suppressors. However, the seemingly contradictory role of p60 ever-expression is not well understood as yet. Currently it is unknown whether the change in levels of CAF-1 proteins is a cause or effect of disease, given the complex nature of cancer both are possible as part of a larger network of expression changes.

As well as its possible role in sporadic cancers, CAF-1 is associated with other proteins implicated in diseases causing susceptibility to cancer or uncontrolled cell growth. The BLM (Bloom) and WRN (Werner) RecQ family helicases are known to cause Bloom's and Werner syndromes respectively, when mutated. Both disorders are characterised by genomic instability, Bloom's patients have a high incidence of cancer whilst Werner patients exhibit

premature ageing (Amor-Gueret, 2006; Jiao *et al.*, 2004). The loss of either protein has been shown to impair CAF-1 mediated chromatin assembly during DNA repair by decreased mobilisation to sites of damage without affecting protein activity (Jiao *et al.*, 2004, 2007).

PCNA and the p48 subunit of CAF-1 are known to be regulated by the genes TCS1 and 2, mutations in which cause the disease Tuberous Sclerosis, characterised by the development of tuberous, cancer-like growths throughout the body (Hengstschläger *et al.*, 2003). These diseases indicate how important the chromatin assembly pathway is for correct development and control of cells.

1.5.2 CAF-1 Subunits

CAF-1 consists of three subunits, known in humans as p150, p60 and p48. Their basic properties are shown in Table 1. All three proteins are found in the CAF-1 complex, however p150 and p48 are seen to perform other roles in the cell. The subunits are discussed individually below.

CAF-1 Subunit	**p150**	**p60**	**p48**
Number of Amino Acids	956	559	425
Theoretical Molecular Mass (kDa)	106.9	61.5	47.7
Theoretical pI	5.7	7.2	4.75
Major Structural Motif	Small coiled coil	WD repeat protein	WD repeat protein
Location in Genome	19p13.3	21q22.13	1p35.1
Gene Length (kb)	42.8	31.4	29.4
Number of Introns	14	14	12
Splice Variants	2	None known	None known

Table 1: **Summary of the Properties of the Human CAF-1 Subunits.**

1.5.2.1 p150

The largest subunit of CAF-1 has little similarity to proteins other than its functional homologues, which exhibit a great deal of sequence variation between species. Various features have been identified by bioinformatic and experimental means, summarised in Figure 9. Originally human p150 was thought to be 938 aa in length, however more recent bioinformatic information revealed a further 18 aa at the N-terminus of the protein. This has required sequence information in older papers to be updated for Figure 9. p150 has been shown to bind a variety of proteins associated with chromatin, as discussed below, but has very little predicted structure itself, the only motif seen by bioinformatic investigation of the sequence is a coiled-coil region, a common motif involved in protein-protein interactions over a variety of different cellular functions (Burkhard *et al.*, 2001). This coiled-coil spans the KER region of the protein thought to be involved with the interaction of p150 with histones (Kaufman *et al.*, 1995).

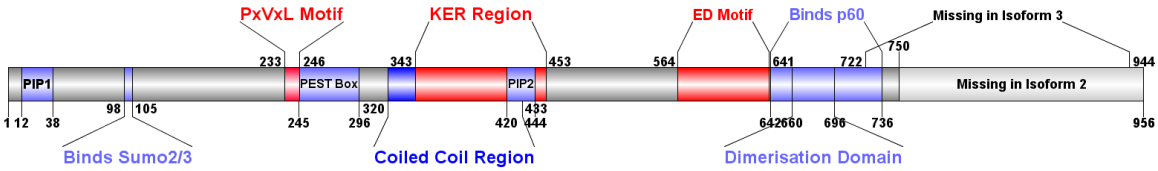


Figure 9: **Map of the p150 Peptide**.
PIPs 1 and 2 are PCNA Interaction Peptides, the PxVxL binds HP1, the coiled-coil region is predicted from sequence identity. KER and ED regions are specified by the amino acid sequence. Mutant isoforms have been shown to exist in certain cell lines but biological importance remains unknown. Constructed using information from (Kaufman *et al.*, 1995; Dong *et al.*, 2001; Quivy *et al.*, 2001; Murzina *et al.*, 1999; Moggs *et al.*, 2000; Reese *et al.*, 2003; Ben-Shahar *et al.*, 2009) with DOG 1.0 (Ren *et al.*, 2009).

p150 has been shown to form a homodimer through a conserved 36 aa sequence motif which is an absolute requirement for chromatin formation *in vitro* (Quivy *et al.*, 2001). The dimerisation event is regulated by phosphorylation of p150 by Cdc7-dbf4, a cyclin dependent kinase required for the initiation of DNA synthesis, linking chromatin assembly to the overall regulation of S-phase in the cell (Gerard *et al.*, 2006). Unphosphorylated p150 exists in a monomer/dimer equilibrium which is driven to the monomer by phosphorylation.

Monomeric p150 has an increased affinity for PCNA, facilitating binding to the replication complex. However, only dephosphorylated p150 is active in chromatin assembly. Dephosphorylation is thought to be mediated by protein phosphatase 1, known to be necessary for CAF-1 activity (Gerard *et al.*, 2006; Keller and Krude, 2000). This two-sided regulation of p150 allows fine-tuning of initiation of chromatin assembly within the cell.

Two additional splice variants of p150 have been identified by bioinformatic analyses and RT-PCR. These C-terminal truncations have a tissue specific localisation, whereas full length CAF-1 is ubiquitous; splice variant 2 is seen in the thymus, colon and small intestine whilst all three variants are seen in the testes. If these variants have a function it is not yet understood (Dong *et al.*, 2001).

Mutations and deletions of p150 resulting in loss of nucleosome assembly activity results in uncontrolled gene re-activation, increased transcription at repression sites, increases in homologous recombination as well as more general structural defects in chromatin (Tchénio *et al.*, 2001; Kirik *et al.*, 2006; Poleshko *et al.*, 2010).

The major role of p150 is thought to be a scaffold, initially for the CAF-1 complex but recent data has identified several roles in chromatin complexes. In the context of CAF-1 p150 binds p60 and p48 to PCNA, primarily using the non-classical PIP2 motif for binding PCNA during chromatin assembly through a mechanism slightly different from the classical motif (Tyler *et al.*, 2001; Ben-Shahar *et al.*, 2009). Experimental dissection of p150 has revealed that the first 296 amino acids are not required *in vitro* for nucleosome assembly or binding to p60. Despite the loss of this section being lethal *in vivo* it is thought that in terms of function, p150 can be divided into two regions: the N-terminal third, important for localising p150 and interactions with other, usually heterochromatic complexes and the C-terminal two-thirds, containing the KER and ED regions as well as the C-terminus being required for chromatin assembly (Kaufman *et al.*, 1995).

The N-terminal portion of p150 has been shown to interact with a variety of proteins associated with chromatin. The HP1 family bind to the PxVxL motif which may target p150 and subsequently CAF-1 to heterochromatin replication foci during late S-phase (Murzina *et al.*, 1999). When bound to HP1, p150 is also seen to bind MBD1, possibly forming a complex

involved in the formation and maintenance of repressive complexes and remains bound until late G2 (Reese *et al.*, 2003). The presence of p150 in heterochromatin may explain why the loss of CAF-1 results in destabilisation of silencing, as discussed earlier.

Other proteins seen to interact with p150 include the SUMO2/3 modifying protein which is recruited to DNA replication foci, linking p150 the re-establishment of histone modifications (Uwada *et al.*, 2009), the Lim15/Dcm1 recombinase which may link p150 to homologous recombination (Ishii *et al.*, 2008) and Rtt106, a yeast histone chaperone also involved in silencing heterochromatin (Huang *et al.*, 2005). In contrast to its general theme of repression, p150 has been seen to be involved in gene activation of the human cytomegalovirus major immediate early promoter, binding to DNA *via* the KER domain and resulting in recruitment of p300 histone acetyltransferase, although this function of p150 remains to be investigated more thoroughly (Lee *et al.*, 2009).

1.5.2.2 p60

CAF-1 p60 is the least well investigated of the components of CAF-1. Bioinformatic analysis reveals seven WD repeats and a C-terminal PEST sequence with multiple phosphorylation sites (Figure 10). p60 is known to bind the C-terminus of p150 but the p150 interacting site on p60 is unknown. The unstructured C-terminal tail of p60 binds to ASF1 (Takami *et al.*, 2007; Tang *et al.*, 2006). p60 is well conserved between species, exhibiting 47 % sequence homology between human and yeast in the area of the WD repeats. Outside its strict homologues p60 displays some sequence similarity to the HIRA histone chaperone (Tang *et al.*, 2006).

p60 is known to be a target for phosphorylation by Protein Kinase C, Casein Kinase and several tyrosine kinases (Smith and Stillman, 1991). Hyperphosphorylation of p60 results in its recruitment to p150, with dephosphorylated p60 less active when assembling chromatin after DNA repair (Martini *et al.*, 1998). The function of p60 within the CAF-1 complex remains largely unknown, although the sequence similarity to HIRA indicates a direct involvement in chromatin assembly.

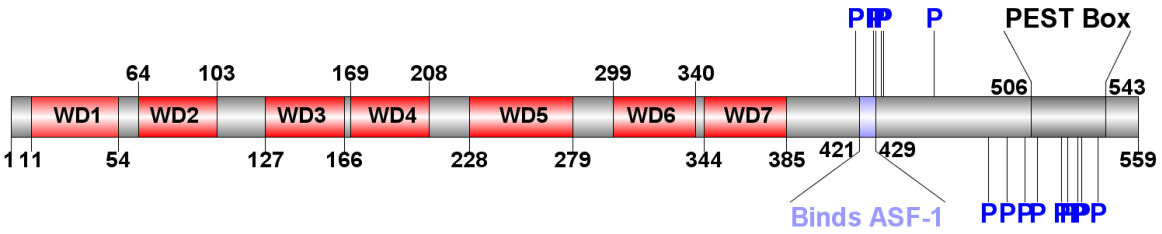


Figure 10: **Map of the p60 Protein**.
WD regions form the predicted beta-propeller structure. P indicates a phosphorylation site. There is no current structure of p60.

1.5.2.3 p48

p48, also known as RbAp48, NURF55 (drosophila) or RBBP4 is a highly conserved protein that binds H3/H4, identified in 1993 as a component of the Rb complex (Qian *et al.*, 1993). p48 is the best understood of the components of the CAF-1 complex, with the structure solved by the SGC (Structural Genomics Consortium, PDB ID: 3GFC) and independently by S. Lejon in this lab (Lejon *et al.*, 2011). The structure of p48, shown in Figure 11, is that of a seven repeat beta-propeller proteins, a common protein-protein interaction motif, also predicted to be seen in p60. The structure of p48 is very similar to that of RbAp46, a paralogous gene with 89 % identity. The structure of p46 was also solved in this lab, in complex with a histone H4 peptide (Murzina *et al.*, 2008). Both p46 and p48 appear to have the same function - that of a H3/H4 chaperone but operate in different complexes. It is thought that the distinction between the two proteins lies more in their physical location than biochemical properties as they appear to function redundantly (Verreault *et al.*, 1998).

The binding of p48 to H3/H4 is currently a topic under investigation. The structure of p46 with a H4 peptide indicates an atypical binding mode of the H4 peptide, as it is seen to bind to the side of the protein rather than the top or bottom face, as is more usual for WD repeat proteins (Murzina *et al.*, 2008). More intriguingly, the binding pocket seen for the H4 peptide is unable to accommodate a fully-folded histone H3/H4 molecule, indicating that there is likely a conformational change in the histone structure on binding to p48 or 46 (Murzina *et al.*, 2008). This feature is also seen in the structure of p55, the *D.melanogaster* homologue of both p48 and p46 (Song *et al.*, 2008).

The histone binding function of p48 is utilised by the cell in many different contexts. As well as assembling chromatin with HIRA and CAF-1, as previously mentioned it is involved in histone assembly of CENP-A, with which it makes the minimal deposition complex (Furuyama *et al.*, 2006). The PRC2 (Polycomb repressive complex) demethylation complex utilises p48 via its interaction with Suz12 to localise the complex to histones (Margueron and Reinberg, 2011). p48 is also involved in acetylation of histones, binding to HAT1 interchangeably with p46 to allow HAT1 to acetylate newly synthesised histones (Verreault *et al.*, 1998). The interaction of p48 with the deacetylase HDAC1 localises it to the Rb gene repressor complex, which represses genes regulated by the E2F transcription factor (Nicolas *et al.*, 2000) and it is also part of the core deacetylase complex of NuRD and Sin3 with HDAC1 and 2 (Zhang *et al.*, 1999).

The NuRD complex is multifunctional and contains remodelling, deacetylase and demethylase activities. Mi-2α/β (CHD3/4) in conjugation with p48, HDAC 1/2, and LSD 1 is responsible for each of these respective activities. Other subunits of the complex include a methyl-CpG binding domain (MBD2/3) subunit which binds to methylated DNA and the MTA1/2 deacetylase recruitment proteins (Ramírez and Hagman, 2009). Within various cell types a number of transcription factors are known to interact with and target NuRD. The regulator FOG1 (friend of GATA1, a nuclear transcription factor) is known to bind NuRD through the p48 subunit *via* an interaction between FOG1's N-terminal tail and p48's top surface, a canonical binding mode to WD proteins. This site is separate to the known H4 binding site of the similar p55 and p46, indicating that p48 can bind to both proteins simultaneously, recruiting proteins to histone H3/H4 (Lejon *et al.*, 2011). This ability to bind both histones and other proteins simultaneously is likely to be central to p48's role in these diverse complexes, allowing a diverse set of functions to be performed by the same protein.

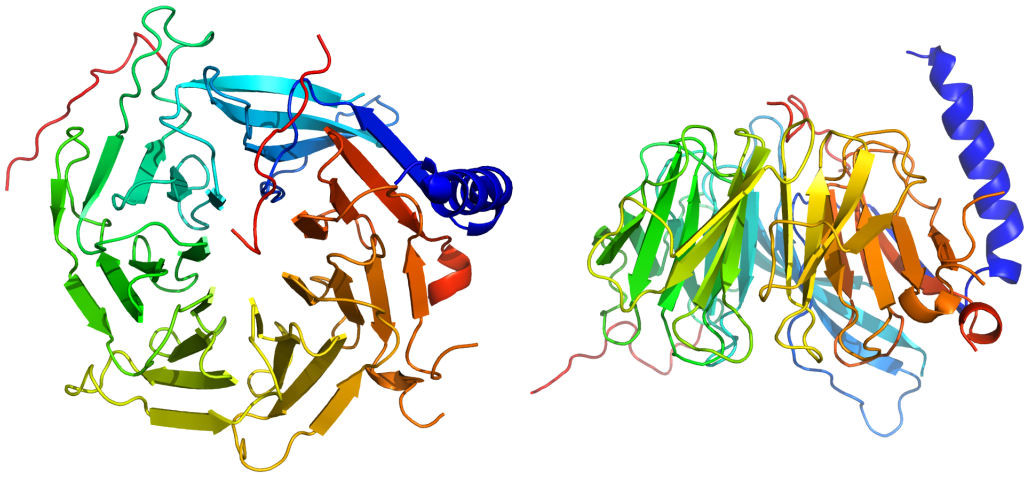


Figure 11: **Structure of p48**.
Crystallised and solved by S. Lejon in the EDL lab. The protein is rainbow coloured from the N to C terminus. Figure compiled in PyMol

Chromatin assembly factor 1 is clearly an important complex in the cell, required for accurate replication of the epigenetic state. Understanding how CAF1 functions will help our understanding of this important part of chromatin dynamics.

1.6 Aims of this Project

Whilst the function of CAF-1 has been widely investigated, the molecular mechanism by which it deposits histones onto DNA and the molecular details of the interactions both within the CAF-1 complex and with other proteins remain unknown.

The primary aim of this study was to produce and crystallise recombinant CAF-1 as well as biochemically characterise the complex in terms of interactions and stoichiometry. To do this CAF-1 had to be produced recombinantly using a new system that allowed reliable co-expression of proteins in baculovirus and purified to homogeneity before crystallisation was attempted.

Chapter 3 describes methods intended to co-express the CAF-1 complex, as a ternary co-expression self-cleaving fusion protein and as a binary complex. The purification of both complexes is also described but did not produce protein suitable for crystallisation. Alternative methods were used to produce individual subunits for further investigation as described in chapter 4. The protein produced was still not sufficiently homogeneous for crystallisation and so further investigation of the interactions between subunits and stoichiometry of the complex was undertaken, as detailed in chapter 5.

Chapter 2

Materials and Methods

2.1 Materials

Yeast Extract and Tryptone were from Duchefa Biochime, Protease Inhibitors were "Complete Mini EDTA Free Protease Inhibitor Cocktail" tablets from Roche. All other chemicals were from Sigma-Aldrich unless otherwise specified. Phusion Polymerase and Taq were from Finnzymes and Roche, respectively. Thrombin from Merck, In-Fusion kit from Clontech and all other enzymes from New England Biolabs. Plasmid purification, PCR clean-up and gel extraction kits were from Macherey-Nagel, MidiPrep kits from Qiagen and oligos from MWG (now Eurofins). Protein Gels were run using the Nu-PAGE system from Invitrogen. Ni-NTA resin was from Qiagen. All protein purification columns were from GE Healthcare.

The MultiBac and iACEMBL systems were kindly provided by the Berger Laboratory (Grenoble), His-p48-pFBDM, His-p150-pFBDM, His-p150p-60-pFBDM and pFBDM-GST-p48 plasmids were previously made by Dr Murzina in this lab and pNIC-Bsa4 vector was a gift from O. Gileadi (SGC, Oxford). The following *E. coli* strains were used: BW23474 and BW23473 cells were from the Berger Lab; Top10 cells were from Invitrogen and Arctic-Express from Stratagene. Anti-p60, anti-p48 and anti-Myc antibodies were from AbCam (ab8133, ab38135 and ab9106 respectively). Anti-his and secondary antibodies (anti-mouse

and anti-rabbit both conjugated to HRP) were from Sigma. FOG1 1-15 peptide was synthesised by Auspep.

2.2 Methods

Buffer recipes are in Appendix B, *Escherichia. coli* genotypes in Appendix C, Plasmids in Appendix D Primers in Appendix E and protein sequences in Appendix F. A benchtop Eppendorf 5417R centrifuge was used for centrifugation of volumes up to 2 ml and a benchtop Eppendorf 5810R centrifuge for volumes up to 50 ml. Larger volumes were spun in a Beckman J-26 XP centrifuge using a JLA10.500 rotor. High speed spins were performed in the same centrifuge using a JA25.50 rotor.

2.2.1 DNA Manipulation

2.2.1.1 Growth and Purification of Plasmids

A single colony of *E. coli* containing the plasmid of interest was grown in 10 ml LB supplemented with relevant antibiotic overnight. Antibiotic concentrations are shown in Table 2. The cells were spun down at 1,500 g and the supernatant discarded. To purify plasmids for cloning experiments a miniprep was performed according to the manufacturer's instructions. Where this did not produce sufficient DNA a midiprep was tried, in the case of pUCDM the alkaline lysis method was also used.

Abbreviation	**Full Name**	**1000x Stock Concentration (mg/ml)**
Amp	Ampicillin	100
Cm	Chloramphenicol	34 (EtOH)
Gent	Gentamicin	10
Kan	Kanamycin	50
Tet	Tetracycline	10 (EtOH)

Table 2: **Concentrations of Antibiotics Used in this Study**

2.2.1.1.1 Alkaline Lysis Method 500 ml of LB enriched with appropriate antibiotic was inoculated and grown overnight at 37 ^{o}C, 250 rpm. The cells were pelleted at 6700 g for 5 min at 4 ^{o}C. The pellet was washed with 100 ml ice cold TES buffer and re-suspended in 28 ml alkaline lysis buffer with a few crystals of lysozyme. This was incubated at room temperature for 10 min then 56 ml NaOH/SDS solution was added, the mixture shaken until clear and incubated on ice for 10 min. 42 ml of ice-cold 5M potassium acetate was then added, the mixture gently inverted and incubated on ice for 10 min. The solution was then spun at 48,000 g for 30 min at 4 ^{o}C and the supernatant filtered. 72 ml isopropanol was added and the mixture left at room temperature for 10 min. This was spun as before at 20 ^{o}C and the supernatant removed. The remaining pellet was rinsed with 70 % EtOH and left to air dry for 10 min. The pellet was re-suspended in 2.5 ml TE buffer with 2.5 µl of RNAse A at 10 mg/ml and incubated at 37 ^{o}C for 30 min, vortexing occasionally. The solution was then phenol extracted by adding 2.5 ml Phenol:Chloroform solution, spun at 8,600 g for 15 min at 20 ^{o}C and removing the aqueous (top) phase to a fresh tube. This was repeated 3-5 times, making the aqueous phase up to 2.5 ml each time with TE buffer. The remaining DNA was then ethanol precipitated by adding 0.1 volumes 3 M NaAc and 2.5 volumes ice-cold 100% EtOH, then spun at 8,600 g for 15 min at 4 ^{o}C. The supernatant was removed and the pellet washed in ice-cold 70 % EtOH. This was then air dried for about 10 min and re-suspended in 0.5 ml water. The concentration and quality of DNA was checked as described in section 2.2.1.2 and the DNA stored at -20 ^{o}C.

2.2.1.2 Quantification of DNA

To calculate the amount of DNA in a sample a ND-1000 spectrophotometer (Nanodrop) was used. The quantity of DNA is calculated based on its absorbance at 260 nm using the Beer-Lambert law: $A=\varepsilon cl$, where A is absorbance at 260 nm, ε is the extinction coefficient, c is the concentration and l is the path-length. The extinction coefficient for double stranded DNA is 0.02 $(ng/\mu l)^{-1}$ cm^{-1}. The purity of the DNA (protein contamination) was assessed based on the ratio of absorbance at 260 nm/280 nm. Pure DNA has a ratio of around 1.8. Protein contaminated samples have a ratio of 1.6 and were discarded or re-purified.

2.2.1.3 Chemically Competent *E. coli*: $CaCl_2$/RbCl Method

An isolated cell colony was inoculated into 5 ml of LB and incubated at 37 ^{o}C, with shaking at 250 rpm overnight. 1 ml of this culture was inoculated into 100 ml LB supplemented with 20 mM $MgSO_4$. The media was then incubated at 37 ^{o}C, with shaking at 250 rpm until the optical density at 600 nm ($O.D._{600}$) reached 0.4-0.6. The culture was then spun at 600 g,for 10 min at 4 ^{o}C. The pellet was re-suspended in 40 ml TFB1 and incubated on ice for 5 min. This was spun down again, the supernatant discarded and the cells re-suspended in 4 ml of TFB2. This was incubated on ice for 15-60 min then flash frozen in 100 µl aliquots in pre-cooled 1.5 ml Eppendorf tubes and stored at -80 ^{o}C.

2.2.1.3.1 Transformation into Chemically Competent *E. coli* An aliquot of competent cells was thawed on ice, 50-200 ng of plasmid solution added and incubated on ice for 30 min. The cells were then heat-shocked at 42 ^{o}C for 45 s and returned to ice for 2 min. 500 µl LB was then added and the tubes incubated at 37 ^{o}C with shaking at 250 rpm for 60 min. 20-500 µl of cell solution was then plated out onto LB-agar plates containing the appropriate antibiotic and left at 37 ^{o}C overnight.

2.2.1.4 Electrocompetent *E. coli*

An isolated colony of the cells to be made competent was inoculated into 100 ml LB- + antibiotic(s) and grown overnight at 37 ^{o}C, with shaking at 250 rpm. This was used to inoculate 1 L LB-antibiotic which was grown to an $O.D._{600}$ of 0.5. Cells were then cooled and spun down at 3000 g for 15 min at 4 ^{o}C. The cells were then re-suspended in 500 ml ice cold 10% (v/v) glycerol solution (sterile) and spun as before. This was repeated with 250 and 50 ml 10% (v/v) glycerol solution and finally the cells were re-suspended in 2 ml 10% (v/v) glycerol and 100 µl aliquots were flash-frozen in liquid nitrogen and stored at -80 ^{o}C.

2.2.1.4.1 Transformation into Electrocompetent *E. coli* An aliquot of competent cells was thawed on ice, 0.5-2 µl of transformation mixture was added and the cells transferred

into an ice-cold electroporation cuvette. The cuvette was placed in an EC100 electroporator (E-C Apparatus Corporation) and a pulse of 1800V applied. 900 µl LB was then added and the cells removed into a sterile 1.5 ml microtube which was incubated at 37 °C, with shaking at 250 rpm for 60 min. 20-200 µl was then plated out onto LB-agar plates containing appropriate antibiotic and left at 37 °C overnight.

2.2.1.5 Digestion of Plasmids

Restriction digestion of plasmids for cloning was performed in 50 µl of solution containing: 1 x NEB Enzyme buffer; 2 mM DTT; 100 µg/µl BSA; 1-3 µg of plasmid DNA and 10 U (1 µl or to a maximum 0.5% total volume) enzyme(s). Reactions were incubated at 37 °C for 3 h, and analysed by agarose gel. If further digestion was required a further 10 U of enzyme was added and the reaction incubated at 37 °C for 3-16 h. Digestions were inactivated by incubation at 70 °C for 20 min. Short-term storage was at 4 °C. The product was checked by agarose gel as above and purified by PCR clean-up or gel extraction as appropriate. Samples were then stored long-term at -20 °C. Analytical digests were performed in 1/5 volume with the same ratio of ingredients.

2.2.1.6 Agarose DNA Gels

Agarose gels were made up to 1 % (w/v) Agarose in 1x TAE containing 0.4 µg/ml ethidium bromide. Samples were prepared by mixing sample with DNA running buffer to a final concentration of 1x and 5-20 µl loaded per well (small comb) or up to 50 µl (large comb). Gels were run in a tank containing 1x TAE for 30 min at 80 V. Bands were visualised by UV using a SYNGENE Gene Genius gel-imaging system. Where bands were to be excised gels were viewed using a Vilber Normad transilluminator with appropriate protective equipment and the bands excised.

2.2.1.7 Primers Design

In-Fusion cloning, Ligation Independent Cloning (LIC) and Sequence and Ligation Independent Cloning (SLIC) primers were designed with a sequence overlap with both vector and insert of 15 or 20 bp respectively, as illustrated in Figure 12. If a tag was to be inserted this was incorporated into the primer between the vector sequence and restriction site. Specific primers are described in Appendix E. All these cloning methods work using a variant of T7 DNA polymerase which in the absence of nucleotides exhibits an exonuclease activity. This is used to provide long overhangs that are then annealed. Any missing amino acids and strand breaks are removed by *E. coli* after transformation.

Figure 12: **Design of In-fusion Primers.**
Oligo design for an In-Fusion reaction at a *Bam*HI restriction site. 15 bp of vector sequence is followed by the restriction site then 15 bp of overlap with the insert sequence. 15 bp of overlap each side is required for In-Fusion and LIC cloning, 20 bp for SLIC

2.2.1.8 Annealing of Oligos

All oligo stocks were made to 100 μM with 10 mM Tris-HCl, pH 8. To anneal two primers, 25 μl of 2 μM solution of each primer were combined in a PCR tube and plunged into 2 L boiling water. This was left to cool to room temperature overnight and the oligos run on an 8% (w/v) acrylamide-TBE gel (Invitrogen) with a 10 bp ladder (Invitrogen) to detect fragments. The gel was run in the Nu-PAGE apparatus in 1x TBE buffer and soaked in TBE supplemented with 0.4 μg/ml ethidium bromide for 30 min and visualised as for agarose gels.

2.2.1.9 High Fidelity PCR

To generate gene inserts using Phusion Polymerase a 50 µl reaction containing: 1 x Reaction buffer; 200 µM each dNTP; 0.5 µM forward and reverse primers; 0.5 µl miniprep of template and 4U enzyme was prepared. The reaction was incubated as shown in Table 3. Finished reactions were stored at 4 ^{o}C until checked on an agarose gel, purified using a PCR purification kit and used in SLIC reactions. Where it was important to remove template DNA 10 U of *DpnI* was added to the PCR mixture before purification and incubated at 37 ^{o}C for 30 min, then purified as before.

Step	**Temperature**	**Time**	**Step**
1	98 ^{o}C	30 s	Initial Denaturation
2	98 ^{o}C	10 s	Denaturation
3	62 ^{o}C	20 s	Annealing
4	72 ^{o}C	90 s	Extension
5	Repeat 25-30 times		
6	72 ^{o}C	5 min	Final Extension
7	4 ^{o}C	hold	

Table 3: **Program for High Fidelity PCR**

2.2.1.10 Ligation

Fragments to be ligated were either double digested to have non compatible ends or treated with 0.5 µl Antarctic phosphatase (NEB) and 1 µl 10 x Antarctic phosphatase buffer. Restriction digested fragments were ligated together in 10 µl solution containing: 1 x NEB T4 ligase buffer; 0.5 mM ATP (Amersham); 5 mM DTT; 20,000 U T4 ligase; 100 ng linear vector and insert at 2 x, 4 x or 10 x molar vector concentration. The reaction was incubated at room temperature for 3 h followed by transformation of 5 µl into competent *E. coli*. Antibiotic resistant colonies were isolated and checked by colony PCR or analytical digestion.

2.2.1.11 In-Fusion Cloning

In-Fusion dry-down cloning kits were used according to manufacturer's instructions.

2.2.1.12 Ligation Independent Cloning (LIC)

The *E. coli* expression vector pNIC28-Bsa4 is suitable for LIC cloning. The vector contains the suicide gene SacB, removed on digestion with *BsaI,* which is lethal in the presence of sucrose. This is used as a selection marker in place of blue/white screening. Linearised vector was combined with: 1 x T4 polymerase buffer; 5 mM DTT; 2.5 mM dGTP; 0.1 mg/ml BSA and 8 U T4 polymerase. This was incubated at 22 ^{o}C for 30 min then inactivated at 75 ^{o}C for 20 min. Insert was treated as above but with dCTP in the place of dGTP. Vector and insert were combined at a molar ratio of 2:1 and incubated at room temperature for 10 min before transforming 3 μl into competent cells as above. Cells were plated onto LB-Kan,5 % (w/v) sucrose.

2.2.1.13 Sequence and Ligation Independent Cloning (SLIC)

Vector and insert to be combined by SLIC were independently treated in a 20 μl reaction containing: 1x T4 polymerase buffer; 5 mM DTT; 0.2 M Urea; 1-3 μg DNA (usually the total purified from the previous step) and 3 U T4 DNA polymerase for 20 min at room temperature. This reaction was arrested by addition of 25 mM EDTA pH 8.0 and inactivated by incubation at 75 ^{o}C for 20 min. Half of each reaction was then combined and incubated at 65 ^{o}C for 10 min and cooled very slowly to room temperature. 5 μl mix was then transformed into competent *E. coli*. Antibiotic resistant colonies were then checked by colony PCR or analytical digestion.

2.2.1.14 Colony PCR

Colony PCR was used to check for insertion of the correct gene following ligation or SLIC. One 20 μl reaction containing: 1 x Taq buffer; 0.2 mM each dNTP and 100 nM forward and

reverse primers was inoculated with a colony from a transformation plate, which was also patch plated. This was run according to the program shown in Table 4 The reaction was then analysed by electroporation on an agarose gel.

Step	Temperature	Time
1	95 ^{o}C	5 min
2	Add 1 U Taq Polymerase	
3	95 ^{o}C	1 min
4	55 ^{o}C	1 min
5	68 ^{o}C	1 min
6	Repeat 3-5 30 times	
7	4 ^{o}C	hold

Table 4: **Program for Colony PCR**

2.2.1.15 Sequencing

Sequencing was done in-house by the Department of Biochemistry DNA sequencing facility. 10 µl of 100 ng/µl DNA was sufficient for one reaction. The oligo used for sequencing was dependent on the vector. All *E. coli* expression vectors were sequenced using the T7 oligo and insect expression vectors were sequenced using the polH oligo (see Table 5).

Name	Sequence
T7	TAATACGACTCACTATAGGG
PolH	CTATAATATATTGTGTTGGG

Table 5: **Oligos used for Sequencing**.

2.2.1.16 Mutagenesis

Mutagenesis was done using the QuikChange II Site-Directed Mutagenesis Kit according to manufacturer's instructions.

Primers were designed using PrimerX (http://www.bioinformatics.org/primerx/).

2.2.1.17 *In Vitro* Cre-Lox Recombination

To combine pUCDM and pFBDM-Lox a total of 0.25 μg DNA at 1:1 molar ratio was combined in a 50 μl reaction with 1x Cre buffer and 1 U Cre recombinase. This was incubated for 30 min at 37 ^{o}C, the reaction was stopped by incubation at 70 ^{o}C for 10 min and stored at 4 ^{o}C. 5 μl was then transformed into Top10 competent cells. Cells were allowed to grow for 3 hours after transformation before being plated onto LB-agar containing Cm and Amp to allow selection for both antibiotic markers. Recombinant plasmids were then purified and checked for characteristic digestion patterns calculated using the CreRecEMBL program (Written for and used with permission of the Berger Lab). Correct plasmids were then used for expression trials.

2.2.2 Protein Production

2.2.2.1 *E. coli* Expression

Production in *E. coli* is IPTG dependent for all vectors used in this study. For initial trial expressions 10 ml of cell culture were grown in LB with appropriate antibiotic at 37 ^{o}C, with shaking at 250 rpm until the O.D.$_{600}$ reached 0.8, at which point they were cooled and moved to a shaker at an appropriate temperature and IPTG added to 1 mM. To check expression, cells were spun down, lysed using 50 μl BugBuster solution (Novagen) and spun to remove insoluble debris. The insoluble pellet was re-suspended in an equal volume of bugbuster. Total, soluble and insoluble fractions were taken to assess protein production.

For large scale expression 10 ml cells were grown overnight at 37 ^{o}C, with shaking at 250 rpm and used to inoculate up to 2 L LB-Antibiotic. Cells were grown to an OD$_{600}$ of 0.8 and IPTG added to a concentration of 1 mM. Once induced constructs of p150 were grown for 5 hours at 20 ^{o}C, with shaking at 250 rpm in Rosetta pLysS and p60 for various

times, as detailed in the results section. Once grown cells are spun down at 6,700 g for 20 min at 4 ^{o}C and frozen at -80 ^{o}C until needed.

2.2.2.2 *Sf21* Insect Cells Expression *via* Baculovirus Infection

Protocols in this section are based on protocols from Dr I Berger (Fitzgerald *et al.*, 2006, 2007; Trowitzsch *et al.*, 2010). Basic cell culture technique was used as described in "Guide to Baculovirus Expression Vector Systems (BEVS) and Insect Cell Culture Techniques" (Invitrogen).

To express protein in insect cells it is necessary to generate a viral stock that can be used to infect growing cells. To make the initial virus, a plasmid containing the gene of interest and the Tn7L transposon element was transformed into competent *E. coli* containing the MultiBac-YFP bacmid (BacY from the iACEMBL system). After electroporation as described above 900 µl of LB was added to the tube and the cells incubated overnight at 37 ^{o}C, with shaking at 250 rpm before being plated out onto LB-Agar containing Kan, Tet, Gent, Xgal (40 µg/ml) and IPTG (1 mM). Incubation at 37 ^{o}C for 24-48 h gives blue and white colonies. White colonies were re-streaked onto a fresh plate to check colour selection for gene incorporation. Colonies that were still white after re-streaking were inoculated into 3 ml LB-Kan/Tet/Gent and incubated at 37 ^{o}C, with shaking at 250 rpm for 10-18 h then spun down at 4000 rpm.

To extract the recombinant bacmid the cell pellet was re-suspended in 300 µl Alkaline Lysis Buffer, 300 µl NaOH/SDS buffer was added, followed by 300 µl 3.0 M potassium acetate, pH 5.5. The solution was then spun at full speed for 3 min in a bench-top centrifuge, and the supernatant removed to a fresh tube. 0.7 volumes of isopropanol was added and the tube centrifuged again at full speed for 10 min. The supernatant was then removed and the pellet washed with 20 % (v/v) EtOH. The DNA was then air-dried under sterile conditions and re-suspended in 20 µl sterile water.

Transfection of the recombinant bacmid was done in 6-well plates (Nunc) containing $0.5^{x}10^{6}$ cells per well. Cells were allowed to attach to the bottom of the plate for 15 min at 27 ^{o}C

in 2 ml Sf900 II Serum Free Media (Invitrogen). Bacmid DNA was diluted 10x in media and combined with a 20x dilution of Fugene (Roche). 150 µl was then added to the cell plate in duplicate. The plate was incubated at 27 °C for 48-60 h. The supernatant from the cells containing virus V_0 was removed to sterile tubes. This virus was used to infect flasks containing 25 ml of cells at 0.5^x10^6 cells/ml and incubated at 27 °C, with shaking at 70 rpm. The cells were then counted every 24 h until they stopped proliferating. The cells were spun down gently at 800 rpm in a bench-top centrifuge and the supernatant removed to form V_1 virus. The cells were re-suspended at 0.5^x10^6 cells/ml and a sample of 1^x10^6 cells was taken every 12 h and processed to measure YFP. For future growths the V_1 virus was used to infect 500 ml flasks at 0.5^x10^6 cells/ml, 10 ml virus was usually sufficient.

To measure the YFP content the sample was spun down at full speed in a bench-top centrifuge for 3 min and the media removed. This was then re-suspended in 500 µl PBS, sonicated for 10 s and a sample taken. The solution was then spun down as before and a sample of the soluble protein taken. The remaining solution was then examined for fluorescence emission at 515-550 nm, using a Perkin Elmer Instruments LS 55 Luminescence Spectrometer with excitation at 513 nm, slits set at 2.5 nm and the photomultiplier voltage at 750 V. Peak fluorescence was measured at 530 nm. Where CFP was measured the machine was set to scan from 450-500 nm with the peak measured at 475 nm.

Once the YFP fluorescence reached a plateau the cells were harvested by centrifugation at 4,000 rpm in a Beckman JLA-10.500 rotor. Pellets were flash frozen in liquid nitrogen and stored at -80 °C until required.

2.2.3 Protein Purification

All proteins were purified from *Sf21* insect cells unless otherwise stated.

Denaturing purifications were carried out identically to the native purification except at room temperature and with the addition of urea to 8M after the initial centrifugation step. Samples were bound to Ni-NTA resin (Qiagen) as described in the native purification and then washed

from 8M to 6M to 2M and then into 0M urea in wash buffer before continuing as detailed in the protocol.

2.2.3.1 CAF-1: HIs-p150ΔN/p60/CBP-p48

2.2.3.1.1 Small Scale A cell pellet from 25 ml of CAF-1 expression was lysed by re-suspension in 14 ml CAF Lysis Buffer and spun at 48 000 g for 20 min at 4 ^{o}C. The supernatant was added to 0.5 ml of Ni-NTA resin (Qiagen), equilibrated in CAF Lysis Buffer and incubated on a rotating wheel for 2 h at 4 ^{o}C. This was then spun down at 360 g for 10 min at 4 ^{o}C, washed with 30 ml CAF High Salt Buffer followed by 30 ml CAF Wash Buffer. The resin was then washed with 5 ml CAF E20 and E40 followed by E100, and CAF E500 as step elutions. Samples were checked for presence of the three subunits on an acrylamide gel.

2.2.3.1.2 Large Scale A cell pellet from an 1 L expression was re-suspended in 50 ml CAF Lysis buffer and mixed gently until homogeneously re-suspended. This was passed twice through an EmulsiFlex cell breaker (Avestin), spun at 48 000 g for 20 min at 4 ^{o}C. The supernatant was then added to 2 ml Ni-NTA resin (Qiagen) equilibrated with CAF Lysis Buffer and incubated on a rotating wheel for 2 h at 4 ^{o}C. The resin was then washed with 50 ml CAF High Salt Buffer and poured into an empty K 9/15 FPLC column and packed by gravity flow. The column connected to an Äkta explorer was washed with 50 ml CAF Wash Buffer and a linear gradient of 0-1 M imidazole in CAF lysis buffer lacking NP-40 was applied at 1 ml/min over 30 CV (column volumes). SDS-PAGE was run to analyse the samples. During later experiments CAF-1 was eluted using sequential elutions of 0.5 ml CAF E500. The eluate from the Ni-NTA column was collected and incubated with Chelix resin (Bio-Rad) to remove any excess Ni^{2+}ions from the eluate.

2.2.3.1.3 CBP based purification 500 µl of calmodulin beads were washed with CAF lysis buffer enriched with 0.2 mM $CaCl_2$. The protein was combined with beads and incubated on a rotating wheel for 2 h at 4 ^{o}C. The solution was then spun at 500 g and the

supernatant removed. Beads were washed three times with 1 ml CAF wash buffer then incubated with 1 ml CAF calmodulin elution buffer on a rotating wheel for 20 min at 4 ^{o}C. Beads were spun as before and the supernatant removed and analysed by SDS-PAGE.

2.2.3.1.4 Ion Exchange Ion exchange chromatography was performed on an Äkta explorer (GE) using a MonoQ HR 5/5 column washed with 5 CV CAF IX buffer B and pre-equilibrated with 5 CV CAF IX buffer A. The protein was diluted by half in 50 mM Tris-HCl pH 8 to give a final salt concentration of 200 mM KCl, bound to the column in 20 % CAF IX buffer B and eluted with a linear gradient between 20 and 100 % CAF IX buffer B over 40 CV.

2.2.3.2 His-p150ΔN/p60

2.2.3.2.1 Small and Large Scale Purifications The tow subunits of CAF-1 were purified following the same protocol as the CAF-1 complex containing three subunits. In addition gel filtration was performed using a Superdex S200 16/60 column using an Äkta Explorer (GE) in CAF GF buffer.

2.2.3.2.2 Low Salt Purification To separate p60 from p150ΔN cells were lysed in CAF lysis buffer containing 50 mM KCl and the supernatant removed from the pellet. The supernatant was then used for ion exchange experiments using an 1 ml Q HP column (GE) with CAF IX buffers A and B as before.

2.2.3.2.3 Dephosphorylation 2 ml of p60 supplemented to 1 mM manganese chloride from the low salt purification was treated with 10 µl (4000 units) λ protein phosphatase (NEB) for 30 min at room temperature. This fraction was then run on a Q HP column as before.

2.2.3.3 p150ΔN

2.2.3.3.1 p150ΔN Native Large Scale Purification The p150ΔN subunit of CAF-1 was purified following the same protocol as the 2 subunit CAF-1 complex, with the addition of an ultracentrifugation step after initial clarification. After centrifugation at 48 000 g the lysate was spun at 100 000 g for 1 h at 4^{o} C in a Beckman Optima LE 80 K ultracentrifuge using a SW 28 rotor.

2.2.3.3.2 Gel Filtration A 24 ml S200 10/300 gel filtration column (GE) was run at 1 ml/min in CAF GF Buffer for 1.5 CV. Standards for estimation of molecular mass were run by M. Chaillet on the same column used for gel filtration of p150ΔN. Dextran blue was run to find the dead volumn of the colmn and beta amylase (200 kDa), alcohol deshydrogenase (150 kDa), bovine serum albumin (66 kDa), carbonic anhydrase (28 kDa) and cytochrome C (12.4 kDa) run to generate a calabration curve.

2.2.3.3.3 Truncated p150 Constructs Initial Purification *E. coli* cell pellets from 100 ml expressions containing p150 constructs were re-suspended in 25 ml of 50 mM Tris-HCl pH 8, 50 mM KCl, 5 mM imidazole and lysed by four passes through an EmulsiFlex cell breaker (Avestin). The lysate was clarified by centrifugation at 48,000 g for 30 min at 4 ^{o}C and passed through 0.5 ml Ni-NTA resin pre-equilibrated with lysis buffer. This was washed with CAF high salt buffer and eluted with a step gradient of CAF E40, E100 and E500 containing 50 mM KCl.

2.2.3.3.4 p150 Construct C Large Scale Purification An *E. coli* cell pellet from 1 L expression was re-suspended in 50 ml CAF Lysis Buffer + 10 mM imidazole and lysed by four passes through an EmulsiFlex cell breaker (Avestin) and the lysate cleared by centrifugation at 48,000 g for 30 min at 4 ^{o}C. Urea was then added to 8 M and the solution passed through 2 ml of pre-equilibrated Ni-NTA (Qiagen) resin. This was washed sequentially with 40 ml CAF Lysis buffer + 6, 4, 2 and 0 M urea. Bound beads were removed here for pull-

downs. When protein was required in solution it was eluted with five washes of 0.5 ml CAF E500. Gel filtration was performed in CAF GF buffer using a S200 10/300 column.

2.2.3.4 p60

2.2.3.4.1 Small Scale Myc- Based Purification A pellet from 200 ml of cell culture was lysed by re-suspension in 20 ml CAF lysis buffer, 200 mM KCl. The lysate was cleared by centrifugation at 48000 g for 30 min at 4 ^{o}C and incubated with 0.25 ml EZView Red Anti-c-Myc Affinity Gel (Sigma) pre-equilibrated with CAF lysis buffer, 200 mM KCl on a rotating wheel for 2 h at 4 ^{o}C. The beads were then washed three times with 1.5 ml CAF lysis buffer, 200 mM KCl, 0.1 % (v/v) NP-40. To release the protein the fusion was cleaved by addition of 2.5 mM $CaCl_2$, 1-2 U of thrombin (sigma) and the mixture incubated on a rotating wheel overnight at 4 ^{o}C. The supernatant was recovered and three 0.25 ml washes of the beads were taken and treated with anti-thrombin beads (Sigma) by incubation a rotating wheel for 2 h at 4 ^{o}C. The supernatant from this reaction was used in pull-down experiments.

2.2.3.4.2 Larger Scale Tagless Purification Cells from 500 ml culture were lysed by re-suspension in 30 ml CAF lysis buffer, 50 mM KCl and clarified by centrifugation at 48,000 g for 30 min at 4 ^{o}C. The clarified lysate was then run on a 1 ml HiTrap Q HP column (GE) over 30 CV from CAF IX Buffer A to Buffer B over 40 CV. The eluate from the column was then run through a 2 ml hydroxyapatite column (Bio-Rad) was packed and run manually with step elutions of 5, 10, 25, 40, 70 and 100% HA Buffer B in Buffer A. Gel filtration was then conducted on the hydroxyapatite eluate using a S200 10/300 and CAF GF buffer. Further ion exchange was performed on the gel filtration eluate after dilution to 100 mM KCl before running it on a MiniQ column (GE) over 30 CV from CAF IX Buffer A to Buffer B over 40 CV.

2.2.3.5 p48

Expression and purification of p48 was performed by either W. Zhang, A. Murthy or S. Lejon as described in (Lejon *et al.*, 2011).

2.2.4 Acrylamide Gel Electrophoresis

Protein content of samples was checked by acrylamide gel electrophoresis. Precast 15 well 4-12% (w/v) Bis-Tris NuPAGE Novex gels (Invitrogen) were run according to manufacturer's instructions with samples supplemented to 1x with SDS-PAGE sample buffer. Gels were stained using Instant Blue stain (Novexin) until bands were clearly visible. Silver staining was performed according to manufacturer's instructions (Fermentas Silver Staining kit).

2.2.5 Western Blotting

A SDS-NuPAGE gel was run and proteins transferred to a PVDF membrane (Millipore) using the Invitrogen NuPAGE blot module according to manufacturer's instructions. Once blotted the membrane was removed and soaked in 5 % (w/v) milk (Marvel) in PBS, and the remaining gel was stained with Instant Blue (Novexin) to check transfer efficiency. After an hour of blocking in milk with gentle agitation at room temperature (or overnight at 4 ^{o}C) the milk was removed by three 10 min washes with PBS. The primary antibody was then applied in 10 ml 5 % (w/v) milk in PBS to the concentration recommended by the manufacturer. This was incubated with gentle agitation for 1 hour at room temperature, the blot washed three times in PBS and the secondary antibody applied as for the first. After the blot had been washed again it was incubated for 1 min at room temperature in 10 ml luminol solution, 100 µl enhancer solution and 3 µl 100 % (v/v) H_2O_2 (Leong *et al.*, 1986). The blot was then quickly drained and placed in an autoradiography cassette. The membrane was used to expose photographic film for 30 s in a darkroom and developed using a AFP Imaging Mini Medical 90 automatic developer. Further exposures were done as required.

2.2.6 Quantification and Concentration of Protein

Protein concentration was estimated using the absorbance at 280 nm and the Beer-Lambet relationship $A = \varepsilon cl$, where A is the absorbance at a given wavelength, ε is the extinction coefficient, c is the concentration and l is the path-length used to measure the absorbance. An approximate extinction coefficient was calculated using the Expasy ProtParam program (http://www.expasy.ch/tools/protparam.html) and this was refined using absolute quantification when proteins were of sufficient quality to sent for amino acid analysis. For p48 $\varepsilon_{280\ nm}$= 83670 M^{-1} cm^{-1}. Values for the other proteins in the CAF-1 complex were not calculated as it was not possible to get the protein clean enough for amino acid analysis.

Protein was concentrated using Amicon Ultra-centrifugal filter with a 10 kDa molecular weight cut-off regenerated cellulose membrane unless otherwise stated. These were used according to manufacturer's instructions.

2.2.6.1 Bradford Assay

The Bradford assay was performed in 1 ml of solution containing 1x Bradford solution diluted from 5x stock (Bio-Rad) and 20 μl protein solution. The absorbance of the solution was read at 595 nm using a Shidmatzu UV-1601 spectrophotometer. A standard curve was constructed using BSA diluted to concentrations between 0 and 800 μg/ml and the slope used to calculate protein concentration of samples. Where the protein sample analysed gave readings outside this range it was diluted until it fell within the standard curve.

2.2.7 Pull-down Assays

Proteins for pull-down experiments were run out on a gel to gauge relative concentration. Approximately equal concentrations of soluble and bead-bound proteins were combined in a tube and 10 μl of well mixed solution was taken for an input sample. 0.5 ml His or Myc pulldown buffer was then added and the mixture incubated with rotation at 4 ^{o}C for 3 hours

or overnight. The beads were then spun down at 200 g in a benchtop centrifuge, washed three times with 0.5 ml of pull-down buffer and re-suspended in the original volume. Input and pull-down samples were then run on a gel and analysed by Western blot or Coomassie staining.

2.2.8 Competition Experiments

Competition experiments were performed as for pull-downs with the additional step of mixing the protein to be pulled down with a ten times excess of the competing peptide before adding to beads.

2.2.9 Limited Proteolysis

Approximately 5 mg protein was digested with subtilisin (Sigma, stored as 1 mg/ml stock at -20 ^{o}C) in 50 mM Tris-HCl, pH 8.0 400 mM KCl at 27 ^{o}C as indicated in the results section. Samples were taken at 0, 15, 30, 60 and 120 min and the reaction stopped imminently by boiling for 2 min in 1x SDS-PAGE sample buffer. Samples were stored at -20 ^{o}C prior to analysis by SDS-PAGE.

2.2.10 Non-Specific Crosslinking

Non-specific crosslinking was performed using formaldehyde. Formaldehyde was added to samples to a final concentration of 1 % (v/v). Reactions were incubated at 27 ^{o}C and samples were taken at 0, 5, 10, 20 and 30 min. The reaction was quenched with 1/10 volumes of 1 M glycine (pH 8), mixed with SDS-PAGE sample buffer to a concentration of 1x. Samples were heated gently at 27 ^{o} instead of boiling and frozen at -20 ^{o}C .

2.2.11 Specific Crosslinking

Specific amine crosslinking with BS^3crosslinker (Thermo Scientific) was performed at 27 ^oC for 30 min. A 2 mg vial of crosslinker was re-suspended in water to give a 3 % (w/v) stock which was diluted 1 in 5, 20, 50 and 100 times. 1 µl of each dilution was added to 10 µl of protein in buffer containing 50 mM HEPES, pH 8, 400 mM KCl. To stop the reaction 1 µl of Tris-HCl, pH 8 was added to each reaction and incubated at 27^oC for 15 min. SDS-PAGE sample buffer was then added to 1x concentration and the samples frozen at -20 ^oC until required. Where accuracy was required for superimposition of Western blots samples were run on the same gel with Precision Plus Protein WesternC Standards (Bio-Rad), visualised separately with appropriate antibody and a 1:20000 dilution of StrepTactin-HRP (Bio-Rad) added to visualise the markers.

2.2.12 Crystallisation Trial Experiments

Initial screens for crystals were performed in 96-well format. Three screens were conducted: Index (Hampton), Pro Complex and JCSG+ (Qiagen). 40 µl of each well solution was dispensed into a 96-well SwissCi 3 well SBS format crystallisation plate. An Oryx 6 robot (Douglas Instruments) was used to pipette sitting drops of volume 0.4 µl containing 50 % protein 50 % reservoir solution. Plates were sealed with crystal sealing film (Hampton) and stored at 4 ^oC. Wells were examined directly after pipetting and then after 1 day, 3 days, 1 week and thereafter every 2 weeks. Crystals and other features were noted. Where crystals were seen they repeated in 24-well hanging drop format of drop size 1 µl. Crystals that could be repeated were subjected to standard tests for salt including poking with a fine needle, examining birefringence for patterns, Izit crystal dye (Hampton) and exposure to home source X-rays.

Chapter 3

Co-Expression of the CAF-1 Complex

3.1 Introduction

Very little biochemical data exists concerning the CAF-1 complex and so I wanted to make recombinant CAF-1 to allow investigation of the complex using methods that require more protein than can be purified from natural sources.

It was known that p150 is prone to degradation when expressed recombinantly in insect cells (N. Murzina, personal communication). To try and prevent this degradation p150 was co-expressed with the more structured and stable members of the complex (p60 and p48) in the hope that they would protect p150 from degradation where they were bound. Full length versions of both p60 and p48 were used as they had both been shown to be expressed without excessive degradation (N. Murzina and S. Patel, personal communications; Lejon *et al.* (2011)). The p150 construct (p150ΔN) used for these co-expression experiments consists of the C-terminal two thirds of the protein (314-956), which have been shown to be sufficient for chromatin assembly both *in vitro* and *in vivo* (Kaufman *et al.*, 1995; Hoek *et al.*, 2011). This is more stable than full length p150 as it does not contain the PEST region.

P150 produced in *E.coli* is mostly insoluble and both p60 and p48 are totally insoluble when produced in *E.coli*. Expression in insect cells, however, has produced both p46 (a p48 homo-

logue) and p48 that are fully folded and soluble (Murzina *et al.*, 2008; Lejon *et al.*, 2011). Co-expression of the three subunits of CAF-1 was therefore attempted in insect cells.

Although the three subunits of CAF1 were expressed previously in insect cells using an old Baculovirus expression system (Kaufman *et al.*, 1995), the expression level was very low and totally unsuitable for biochemical and structural studies. The complex was expressed using a separate virus for each subunit, co-infecting cells with the three viruses in an attempt to co-express the complex. This is an inefficient method of complex expression as many cells are not infected by all three viruses, or receive multiple copies of one or two of the viruses, resulting in differential expression patterns across a cell population. This is problematic when one complex member is required to stabilise another as there is no guarantee that they will both be expressed at similar levels in the same cell, resulting in degraded or aggregated protein. Therefore different systems were investigated for co-expression.

To co-express the three subunits of CAF-1 in one virus the MultiBac/iACEMBL system designed by I. Berger (Fitzgerald *et al.*, 2007) was used. The MultiBac system consists of two plasmids (pFBDM and pUCDM, see Appendix D). Each plasmid contains two promoters for protein expression (*p10* and *Polh*) within a "multiplication module" which can be used to increase the number of proteins expressed from a single plasmid, see Fitzgerald *et al.* (2007) for further details. The two vectors can be combined by *cre-lox* recombination to form a multi-protein expression vector which is then used to generate a bacmid by site-specific recombination. The bacmid is then purified and used to produce a single virus expressing multiple genes. The advantage of this system is that different combinations of vectors can be used to produce a single virus. This means that multiple proteins can be reliably expressed in one go and that different combinations of proteins can be expressed without further gene cloning being required.

The later developed iACEMBL is a continuation and improvement on the same system incorporating more plasmids (Berger *et al.*, 2004; Trowitzsch *et al.*, 2010). Further to this, the MultiBac viral DNA has been modified to remove a protease and a chitinase, reducing protein degradation and cell lysis. The bacmid also includes a YFP gene under the same promoter (Polh) as the protein cloned into the major expression sites of the plasmids. This allows pro-

tein expression to be monitored without the need for Western blotting as the expression of YFP peaks at the same time as expression of the protein of interest without affecting protein production (Trowitzsch *et al.*, 2010).

Originally our plan was to to clone and express CAF-1 using the MultiBac system as a binary complex of p150 and p60 with p48 being added later, either by cloning in a further gene or by addition of purified protein to reconstitute the complex. We hoped that p60 would protect p150 from degradation and that we would be able to make deletion mutants of both p150 and p60 to identify suitable constructs for crystallisation work. p48 did not need to be investigated as it had already proven amenable to crystallisation.

His-tagged p150ΔN had previously been cloned into pFBDM. Myc-tagged deletion constructs of p60 were to be cloned into pUCDM and used for co-expression with HIs-p150ΔN, as detailed in Figure 13. When a p60 construct that bound p150ΔN had been identified, deletion mutants of p150 would be designed, expressed and their interaction with p60 examined. However, problems were encountered cloning into pUCDM (as detailed in chapter 4) and therefore alternative methods of expression of the intact ternary CAF-1 complex were also investigated, see below.

A self-cleaving fusion protein containing all three members of the CAF-1 complex was therefore also expressed in collaboration with the I. Berger laboratory, alongside work in Cambridge, to co-express a binary complex consisting of p150 and p60 from a single plasmid. For the latter, our intention was to reconstitute the complex with p48, expressed and purified separately.

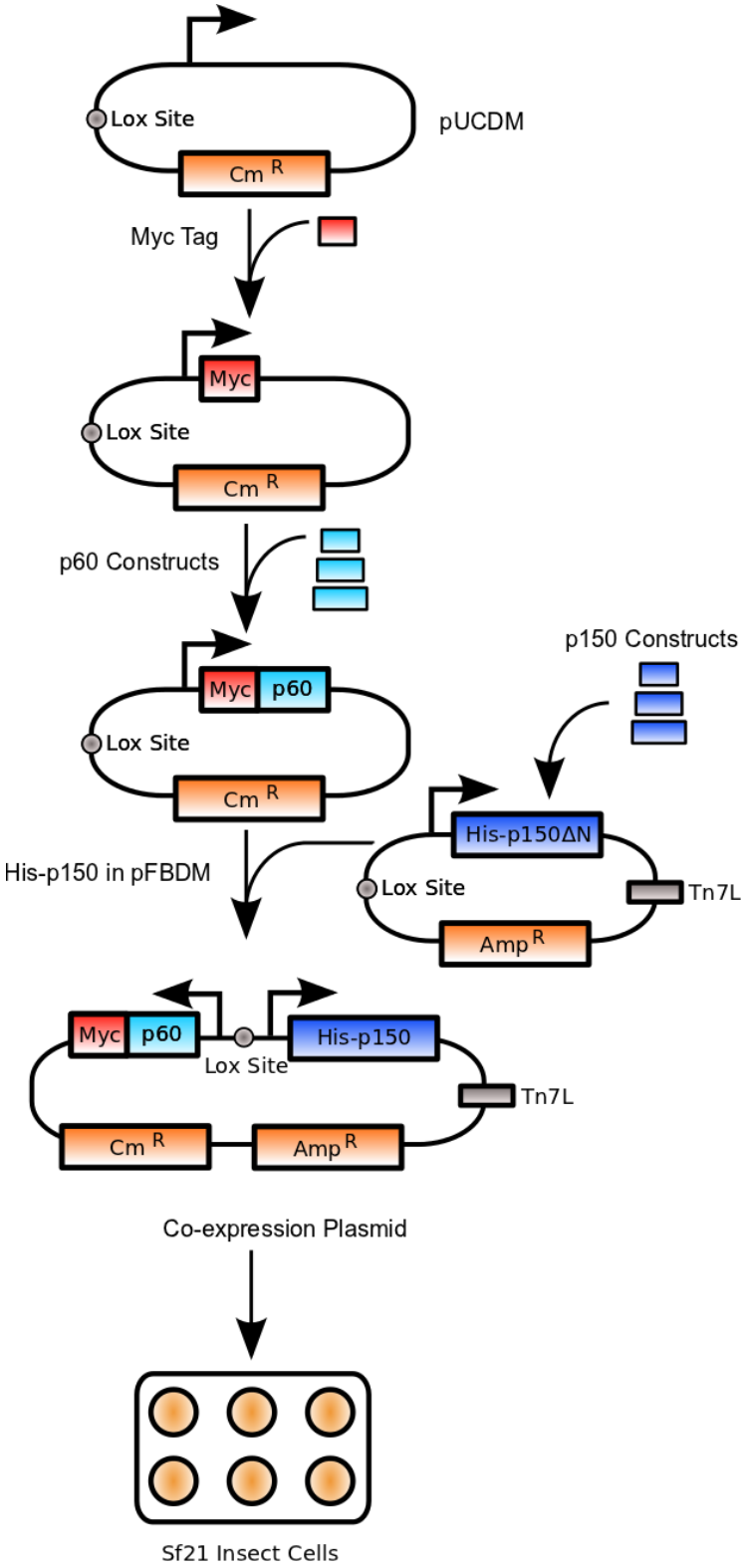


Figure 13: **Cloning Plan for Co-expression of p60 and p150.**
Deletion constructs of p60 were designed for cloning into pUCDM containing a Myc-tag. The p60 constructs would then be combined with p150 for co-expression and identification of interacting constructs. This method would then be used in further experiments to test various p150 deletion constructs.

3.2 CAF-1 Fusion Protein

3.2.1 Cloning

A construct expressing the ternary CAF-1 complex as a single, self-cleaving polypeptide was designed to ensure that all proteins in the complex were expressed at the same level. The protein contains the TEV protease, the three CAF-1 subunits, and a C-terminal cyan fluorescent protein (CFP) molecule (to provide a check that the whole fusion protein was being expressed), as illustrated in Figure 14, was constructed by Y. Ni in the Berger lab. All polypeptide components are separated by TEV cleavage sites. This was then used to express the ternary CAF-1 complex. To allow for purification of the complex p150 was tagged with a hexa-his tag and p48 was tagged with a calmodulin binding peptide (CBP).

Figure 14: **Structure of the CAF-1 Co-Expression Cassette.**
A single polypeptide containing all three subunits of the CAF-1 complex, a CFP marker and TEV protease were used to express the ternary CAF-1 complex. All proteins were separated by a TEV cleavage site allowing the fusion protein to self-cleave into the proteins of interest. Figure constructed using DOG 1.0 (Ren *et al.*, 2009).

3.2.2 Expression and Small Scale Purification

After initial expression trials performed by F. Garzioni in the Berger lab as described in section 2.2.2.2, the virus expressing the ternary CAF-1 fusion protein was used to infect a 25 ml culture of *Sf21* cells. Cells were counted daily and once cell growth had arrested (denoted dpa, day post arrest) YFP and CFP expression was followed by fluorescence spectroscopy until expression peaked and the cells were harvested (Figure 15). Both YFP and CFP expression follow the same pattern and peak at the same time, indicating that YFP is a good indicator of protein expression and that the whole complex is being translated and expressed as CFP is the C-terminal peptide (Figure 15 A). Total and soluble protein production was assessed by SDS-PAGE (Figure 15 B). A distinct band at the expected size p150ΔN (indicated

by arrow) was observed to increase in intensity over the expression time-course (compare B lane 3 to lane 13, Figure 15), however the other proteins could not be seen as they run at the same size as cellular proteins. Western blotting (Figure 15 C) confirmed this with a discrete band for each protein and a double band for TEV as it self-cleaves.

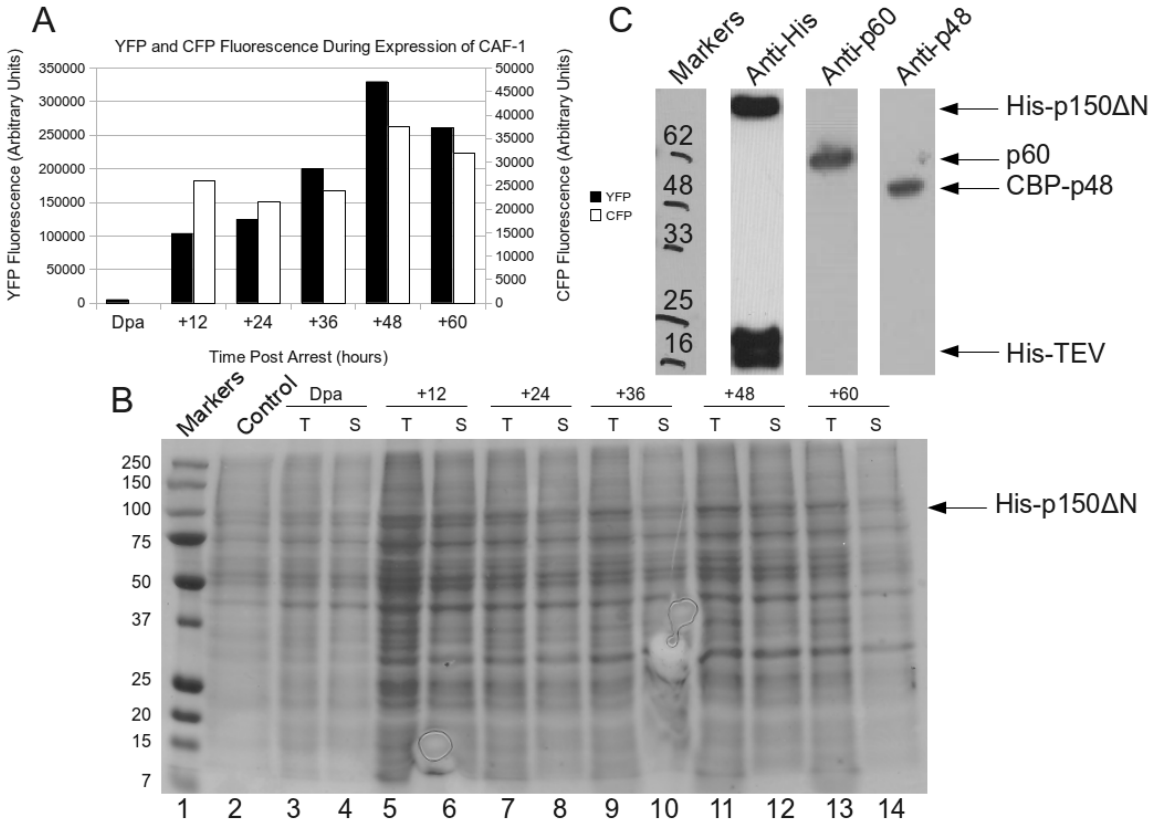


Figure 15: **Expression of CAF-1 Fusion Protein.**
A: Fluorescence of YFP and CFP over time. 1 x 10^6 cells were lysed in 0.5 ml PBS and the fluorescence maxima measured at 530 nm for YFP and 475 nm for CFP. Dpa: day of growth arrest. Fluorescence was measured in arbitrary units. B: 10 % acrylamide SDS-PAGE of total (T) and soluble (S) cell extract at each time point. The control was uninfected cells. Molecular Weight markers are in kDa. C: Western blot of total lysate from Dpa+60. Anti-his antibodies were used to detect p150 as no specific antibody is available against p150ΔN.

After the cells were harvested the complex was purified using a standard protocol for nuclear proteins as described in section 2.2.3.1. This contains 400 mM KCl to enhance removal of chromatin bound proteins. Figure 16 shows a typical SDS-PAGE analysis of the resulting purification. The cells were lysed using NP-40 and spun to remove cell debris. His-tagged p150ΔN with it's associated proteins was purified by affinity chromatography with Ni-NTA

resin (Figure 16, lanes 3 and 5). The resin was washed to remove non-specifically bound proteins and the bound complex eluted with an imidazole step gradient (Figure 16, lanes 9-11). SDS-PAGE analysis showed that three bands were eluted from the beads at sizes corresponding to those expected for p150ΔN, p60 and p48.

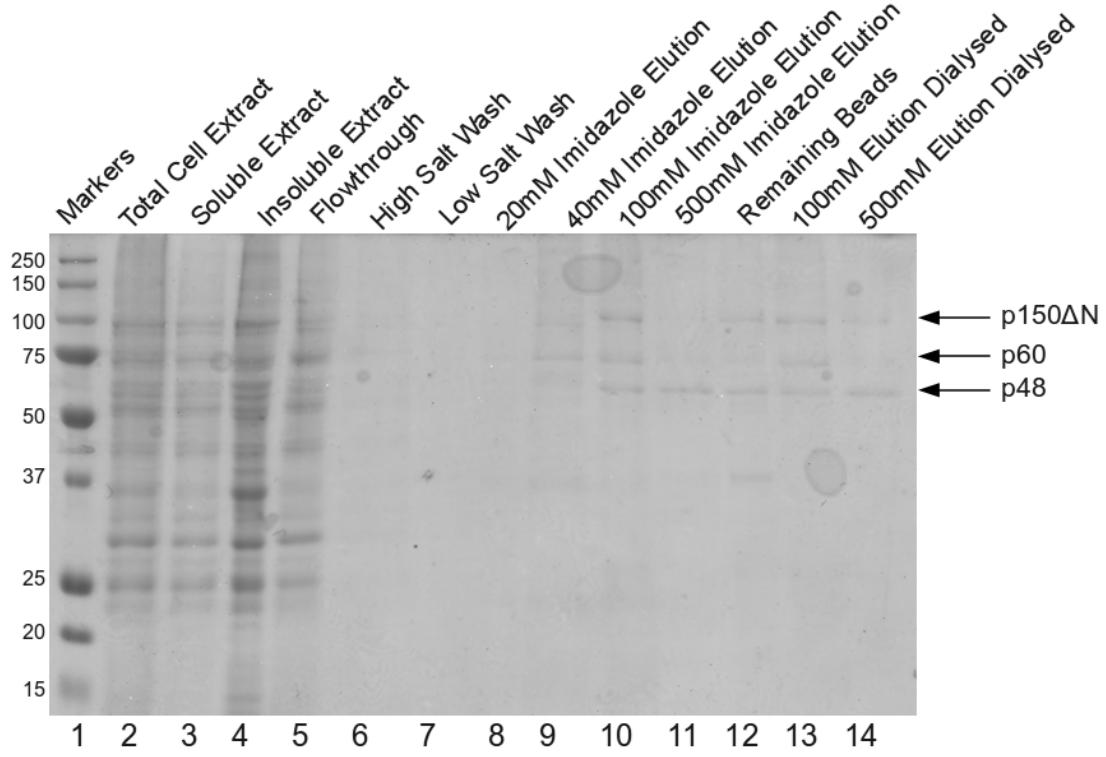


Figure 16: **Small Scale Purification of Co-expressed CAF-1 Complex.**
Cells expressing the CAF-1 fusion protein were lysed and the proteins were bound to Ni-NTA beads in batch, washed and eluted. The eluted proteins were then dialysed to remove imidazole. 10% acrylamide SDS-PAGE analysis shows three bands can be seen p150ΔN, p60 and p48 and were confirmed by Western Blotting (data not shown). Markers are in kDa.

To assess the concentration of imidazole required to elute CAF-1 fusion protein from beads a further purification was performed as before, but the washed beads were packed into a column and eluted using a continuous imidazole gradient between 0 and 1 M (Figure 17). This showed that the complex elutes over several fractions (Figure 17, lanes 4-10) but is eluted by the time the gradient has reached 400 mM imidazole. P60 appears to elute earlier and in smaller amounts than the other proteins, assuming each protein binds coomassie dye equally. Trials of methods to concentrate the eluted protein revealed that the addition of 5 % (v/v) glycerol and use of regenerated cellulose membranes reduced losses to around 20 %.

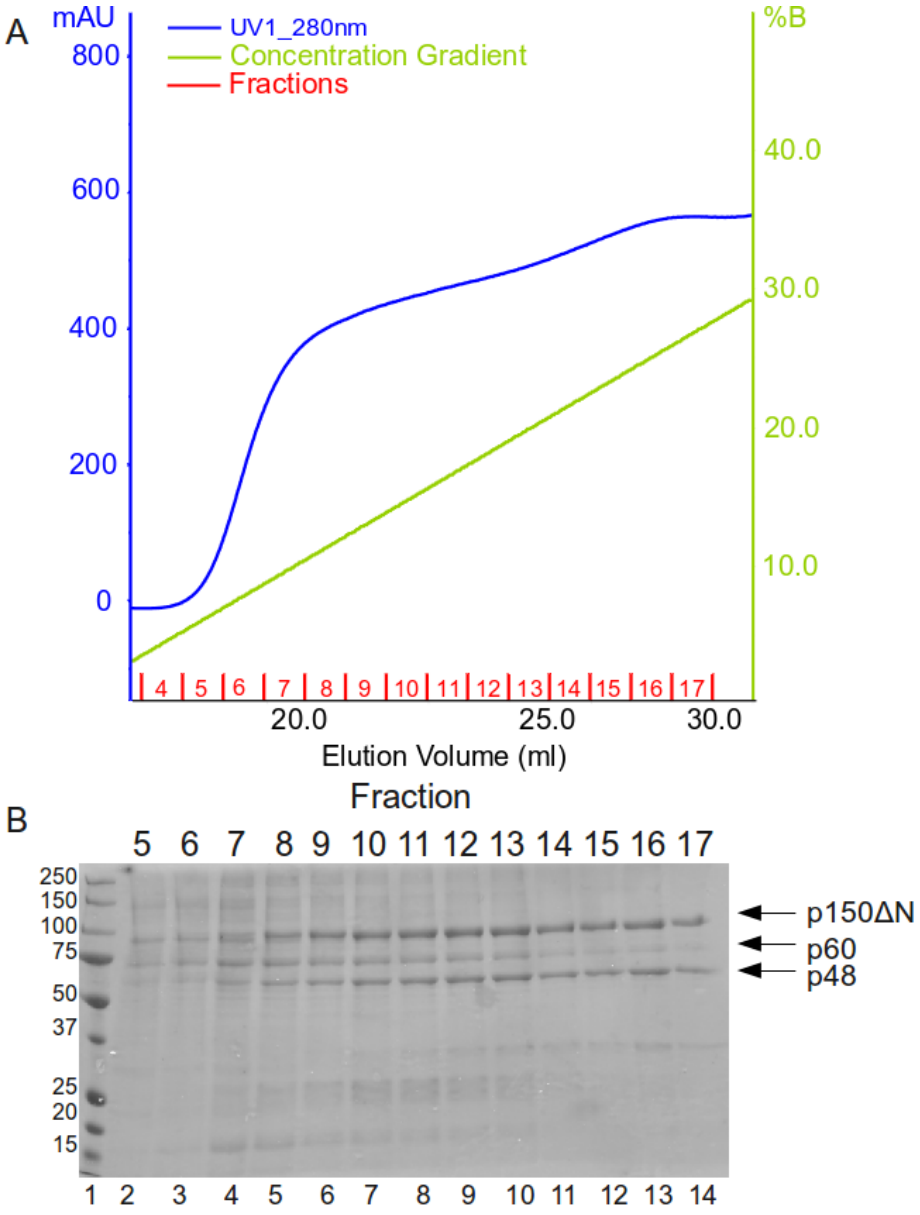


Figure 17: **Elution of the CAF-1 Complex from a Ni-NTA Column.**
Clarified lysate of cells expressing the CAF-1 fusion protein were bound to Ni-NTA beads and the complex eluted in buffer containing a 0-1 M imidazole gradient over 30 CV and collected in 1 ml fractions. A: Elution Profile, Imidazole absorbs UV at 280 nm, causing an increasing baseline. B: 100 µl of each fraction was analysed by 10 % acrylamide SDS-PAGE after being concentrated by evaporation. Markers are in kDa.

3.2.3 Large Scale Purification

Initial attempts at larger scale expression using cell pellets from a 1 L expression and purification resulted in a less pure sample compared to the small scale purifications. One reason for this may be a lower protein expression and therefore a reduced complex:cellular protein ratio. This was not improved by adding more virus when infecting cells to increase expression. A variety of methods were trialled to clean the sample further after affinity purification. As a first approach, the second tag in the complex, CBP, which was fused to p48 was used (Figure 18). It proved possible to bind the complex to calmodulin beads although the subsequent elution was less clean than expected with some of the His-tagged TEV clearly binding to the beads (Figure 18, lanes 5 and 6). This was not removed by washing with high salt buffer, unlike p60 which is lost even before the stringent wash (compare lane 2 with lane 5). Further to this the protein elution from the beads was very inefficient and subsequent experimentation yielded no better results. Consequently, this was not used as a further purification technique. Similar experiments using p48 specific antibodies bound to protein A beads also did not provide better results (data not shown).

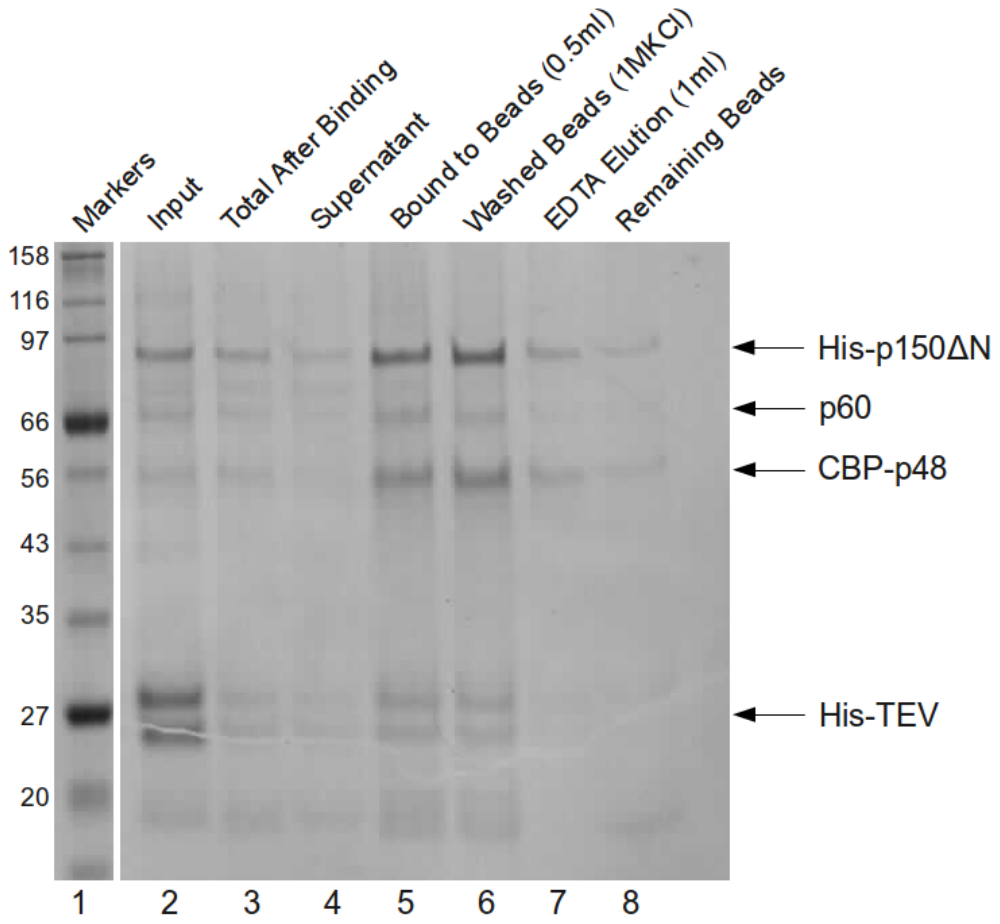


Figure 18: **Purification of the CAF-1 Complex Using Calmodulin Beads.**
Protein eluted from a Ni-NTA column was supplemented with $CaCl_2$ and bound to calmodulin beads. TEV self-cleaves to give a double band. Markers are in kDa.

An alternative purification based on protein net charge was attempted using a MonoQ column at pH 8 (section 2.2.3.1). However it was not possible to lower the salt concentration below 200 mM KCl, (which may have improved binding of the protein complex to the column) because the protein precipitates out of solution if the salt concentration is less than 200 mM KCl.

Purification on the MonoQ column revealed p60 eluting over a wide range of salt concentrations (Figure 19 C, lanes 3-9) whereas p48, along with much of the p60 was eluted in one peak (Figure 19 C, lane 9). Two bands were visualised by anti-p60, one at the expected size, around 60 kDa and a larger band at around 85 kDa. The larger band is of unknown

composition and not seen in other purifications (e.g. Figure 23 and Figure 25) so may be an artefact of this expression method.

The Western blot using anti-his to visualise p150ΔN did not visualise any protein, this is because p150ΔN transfers poorly onto the membrane and at the concentrations here is insufficient to react with the antibody. Ensuring that the gel was in the transfer tank with the higher molecular weight bands at the bottom increased transfer efficiency but was still insufficient for this particular blot.

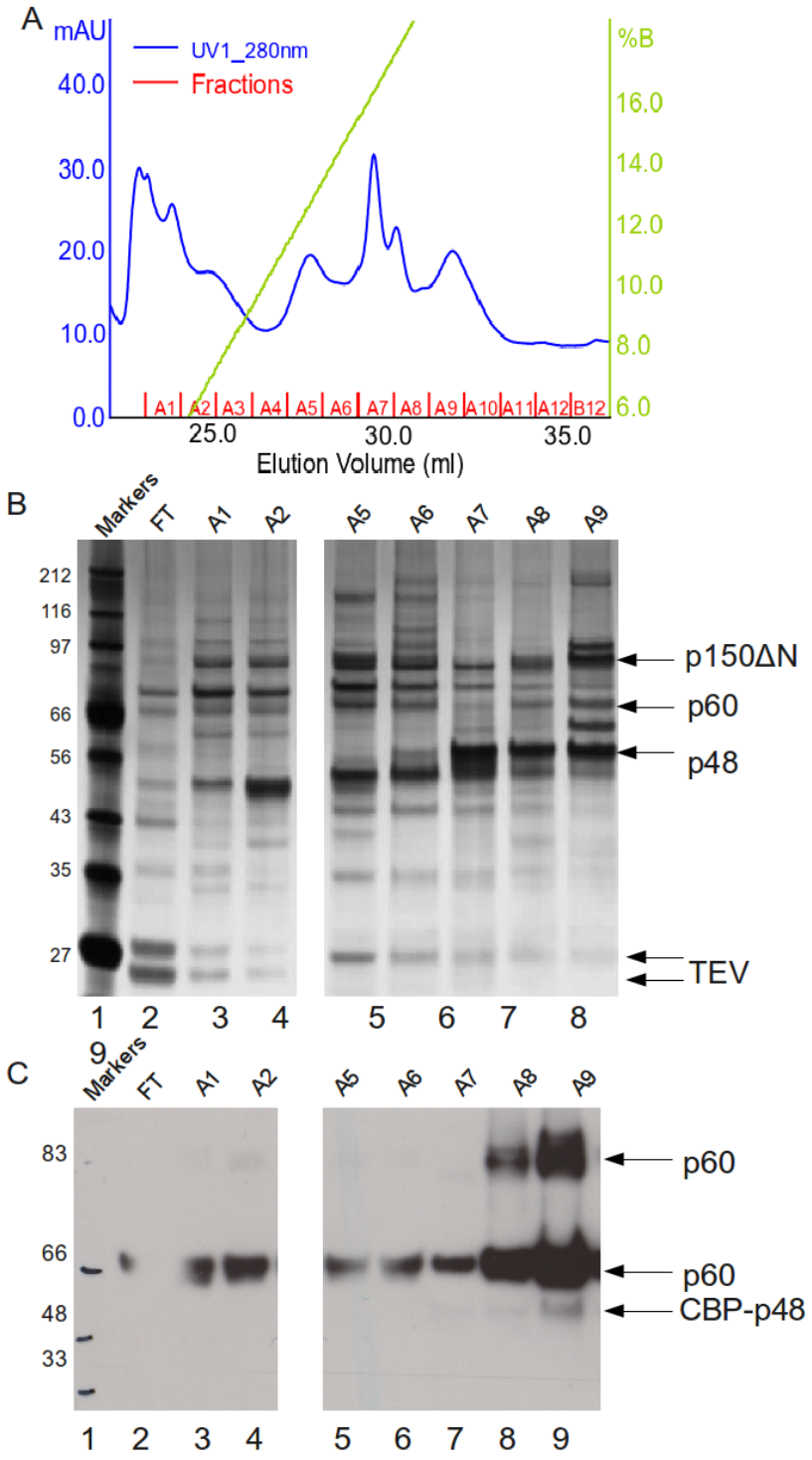


Figure 19: **Anion Exchange Chromatography of the CAF-1 Complex.**
A: Elution profile of Ni-NTA purified CAF-1 complex eluted from a MonoQ column with a salt gradient from 0.2 to 1 M KCl over 40 CV. B: Silver stained SDS-PAGE gel of unbound sample (FT) and relevant fractions from the column with fraction number indicated at top of the gel. Arrows indicate expected sizes of proteins. C: Western blot of gel fractions. The Western blot was probed with anti-p60 (reacting bands labelled p60) followed by anti-p48 (reacted band labelled CFP-p48). Anti-His did not visualise p150ΔN as there was insufficient protein concentration (not shown). Markers are in kDa.

Attempts to lower the amount of salt used in the CAF-1 fusion protein purification in an effort to preserve the complex revealed that extraction of p150ΔN requires 400 mM KCl , which is characteristic for a chromatin bound nuclear protein (Dignam, 1990); extraction using less salt (200 mM or 50 mM) results in insoluble p150ΔN (Figure 20), compared to 400 mM KCl where sufficient protein is extracted to be visible on a Coomassie stained SDS-PAGE (See Figure 17). It was however observed that a fraction of p60 and p48 was soluble in the lower salt conditions, indicating that some of these proteins are not bound to chromatin (Figure 20, compare lanes 2 and 3 for p60 and p48). It was not however possible to purify the soluble protein further as p60 is not tagged and as seen before CBP-p48 cannot be efficiently eluted from calmodulin columns (Figure 18).

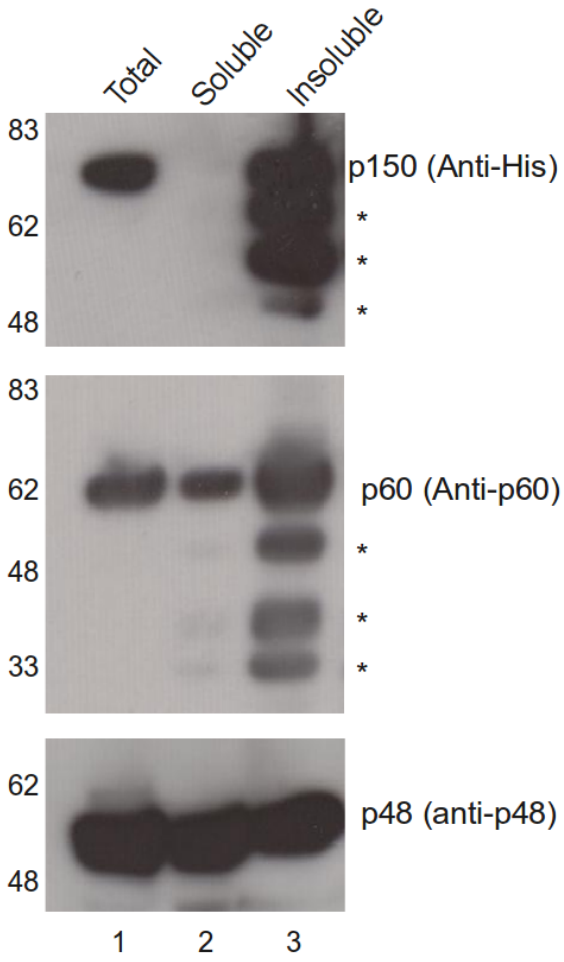


Figure 20: **Low Salt Solubility of CAF-1 Complex Subunits**.
Cells expressing CAF-1 fusion protein were lysed in buffer containing 50 mM KCl, separated into soluble and insoluble fractions and analysed by Western blot with antibodies as indicated. Asterisks (*) indicate degraded proteins. Markers are indicated on the left in kDa.

These results show that whilst we have produced CAF-1 as a self-cleaving fusion protein in insect cells it is not possible to purify the complex to homogeneity, mainly due to low yield making it difficult to remove impurities, which seem to interact with the complex as a anion exchange purification does not remove them. If the yield of the virus was increased this may help, however the tendency of the complex to form non-specific interactions with other proteins which would have to be overcome.

The separation of the parts of the CAF-1 fusion protein into components that appear to be stable separately suggested that it might be possible to use the MultiBac system to express the proteins individually. Expression of the binary p150ΔN/p60 complex is discussed below and individual protein expression is discussed in Chapter 4.

3.3 p150ΔN/p60 Co-expression

Concurrent with expression of the CAF-1 self-cleaving fusion protein a binary complex containing p150 and p60 was expressed and examined. The p150ΔN/p60 co-expression vector was made by N. Murzina and consisted of His-tagged p150ΔN expressed from the *Polh* promoter and untagged p60 expressed from the weaker promoter (see Appendix D). The His-tag on p150 should selectively purify p150 that has not been N-terminally degraded along with p60 which it was hoped would stabilise the C-terminal end of the protein.

3.3.1 Expression and Purification

The complex was expressed and an initial purification performed as for the CAF-1 fusion protein (see section 2.2.3.1). The initial purification is shown in Figure 21. The yield of protein from this virus was higher than that of the CAF-1 fusion protein virus (compare Figure 21 lanes 7 and 8 with Figure 16, lanes 10 and 11), allowing a gel filtration column to be run after purification on Ni-NTA resin (Figure 22). There appears to be some aggregated p150ΔN (lanes 2-5) in the initial peak but p150ΔN/p60 complex can be purified (lanes 6-12).

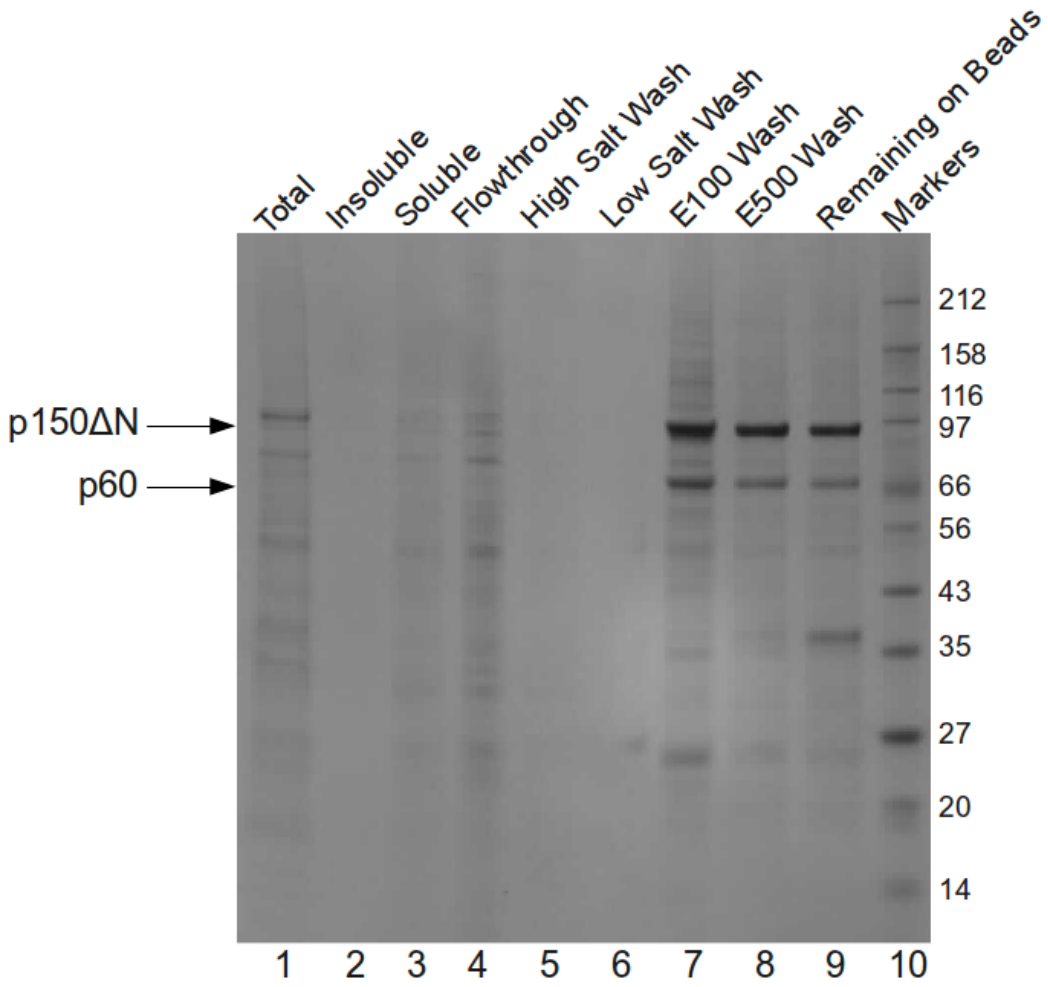


Figure 21: **Small Scale Purification of p150ΔN/p60 Binary Complex**
SDS-PAGE of purification from 25 ml of *Sf21* cell culture infected with p150ΔN/p60 virus using Ni-NTA resin. Markers are in kDa.

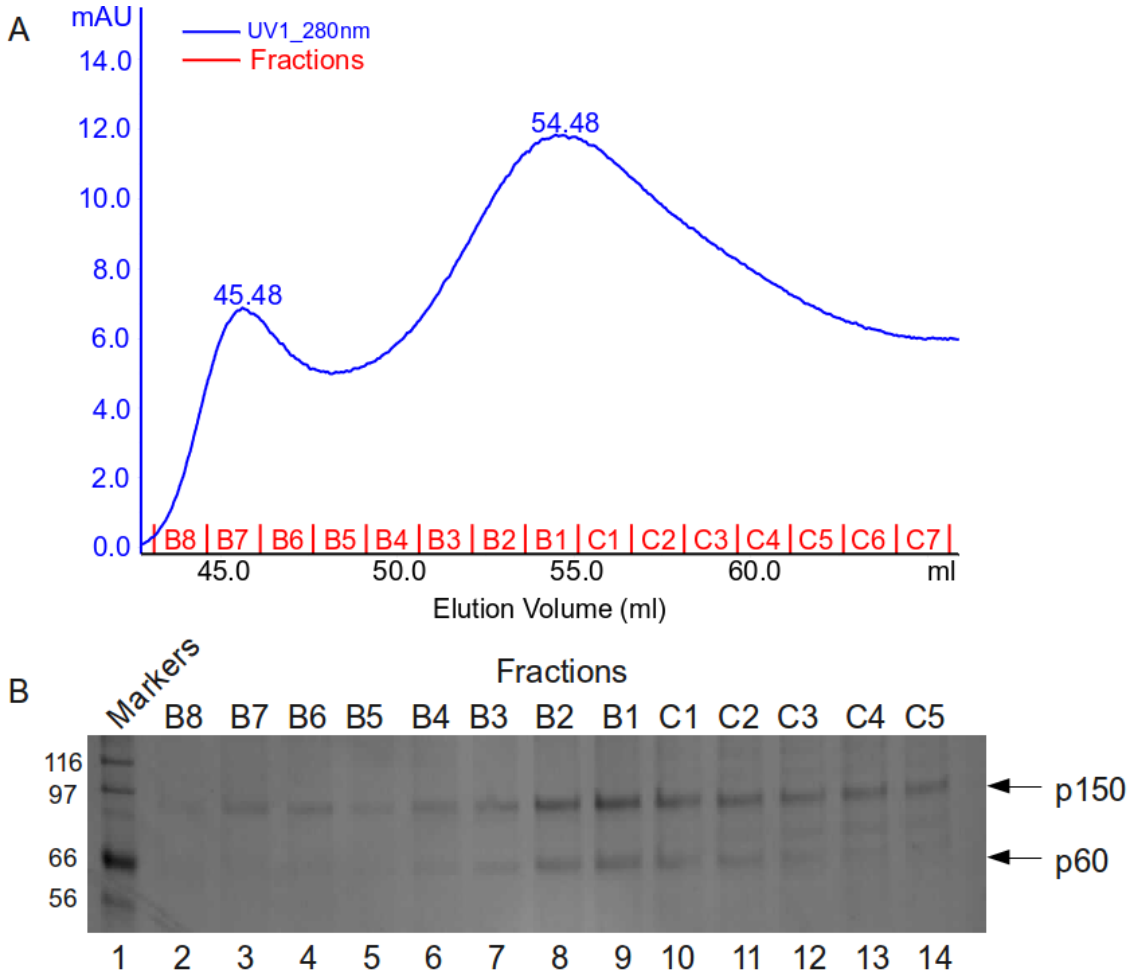


Figure 22: **Gel Filtration of p150ΔN/p60**.
Eluate from the Ni-NTA beads was concentrated and loaded onto a S200 16/60 size exclusion column, Elution fractions of 1.5 ml were collected. A: Elution Profile B: SDS-PAGE analysis of 50 µl of each fraction. The fraction numbers are indicated at the top of the gel. Markers are in kDa.

3.3.2 Ion Exchange Chromatography

While small scale expression and purification resulted in a relatively pure complex, scaling up to a purification from 1 L of cell culture revealed many containments. In an attempt to remove contaminants ion exchange chromatography was performed: the protein eluted from Ni-NTA beads was diluted to 200 mM KCl and loaded onto a MonoQ HR 5/5 column at pH 8.0. The eluate was analysed for the presence of p60 and His-p150ΔN by Western blotting (Figure 23). This showed that a significant fraction of protein did not bind to the column (lane 3 of gel and 2 of Western blot) and the protein that did was spread across the whole elution profile (lanes 4-13 of gel, lanes 3-10 of Western blot). Changing the salt concentration of the starting buffer did not make any difference as the complex precipitates out of solution irreversibly at lower salt concentrations. Therefore ion exchange could not be used to purify the p150/p60 complex.

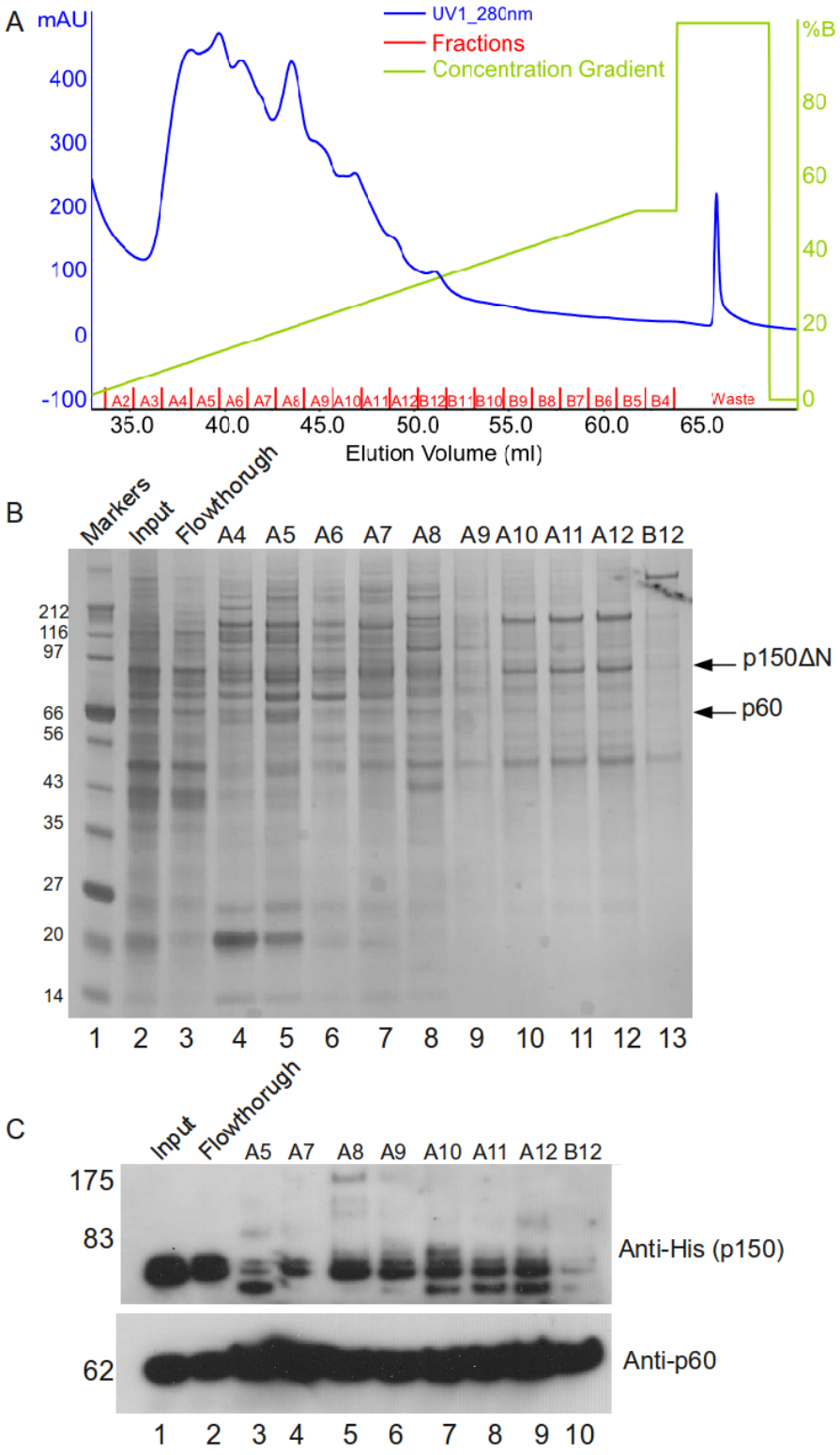


Figure 23: **Anion Exchange Chromatography of Ni-NTA Purified p150ΔN/p60**.
A: Elution profile of p150ΔN/p60 eluted from MonoQ column with 0-1 M KCl gradient over 40 CV B: SDS-PAGE of fractions as indicated at the top of the gel, (15 µl of each fraction was loaded). C: Western blot of the indicated fractions using anti-His or anti-p60 antibodies to identify His-p150ΔN and p60, respectively. Markers are indicated on the right in kDa.

The proteins eluted from the ion exchange column over wide range of salt concentration. This indicates that there may be different charge species present. It was obvious from the Western blot that there is some degradation of p150ΔN, which could account for it eluting in multiple peaks but p60 is spread more widely in both this experiment and also when isolated separately from p150ΔN in low salt, as described below.

p60 is known to be hyperphosphorylated under certain circumstances (Martini *et al.*, 1998). Multiple phosphorylation sites would cause different charge species, therefore a fraction of p60 isolated in low a low salt condition was treated with λ phosphatase as described in section 2.2.3.2. This treated fraction was run through an anion exchange column under the same conditions as an untreated sample and the two elution profiles were compared. The results are shown in Figure 24. If multiple phosphorylation states were the main cause of the wide distribution of p60 in the elution profile then the λ phosphatase treatment should have resulted in a narrower peak. However, whilst phosphatase treatment does alter the profile of p60 elution from an ion exchange column it spreads it out even further, indicating that there are likely other reasons behind this wide distribution of charges, such as clipping of charged C-terminal residues from the unstructured tail of p60 or binding to other proteins such as various chaperones.

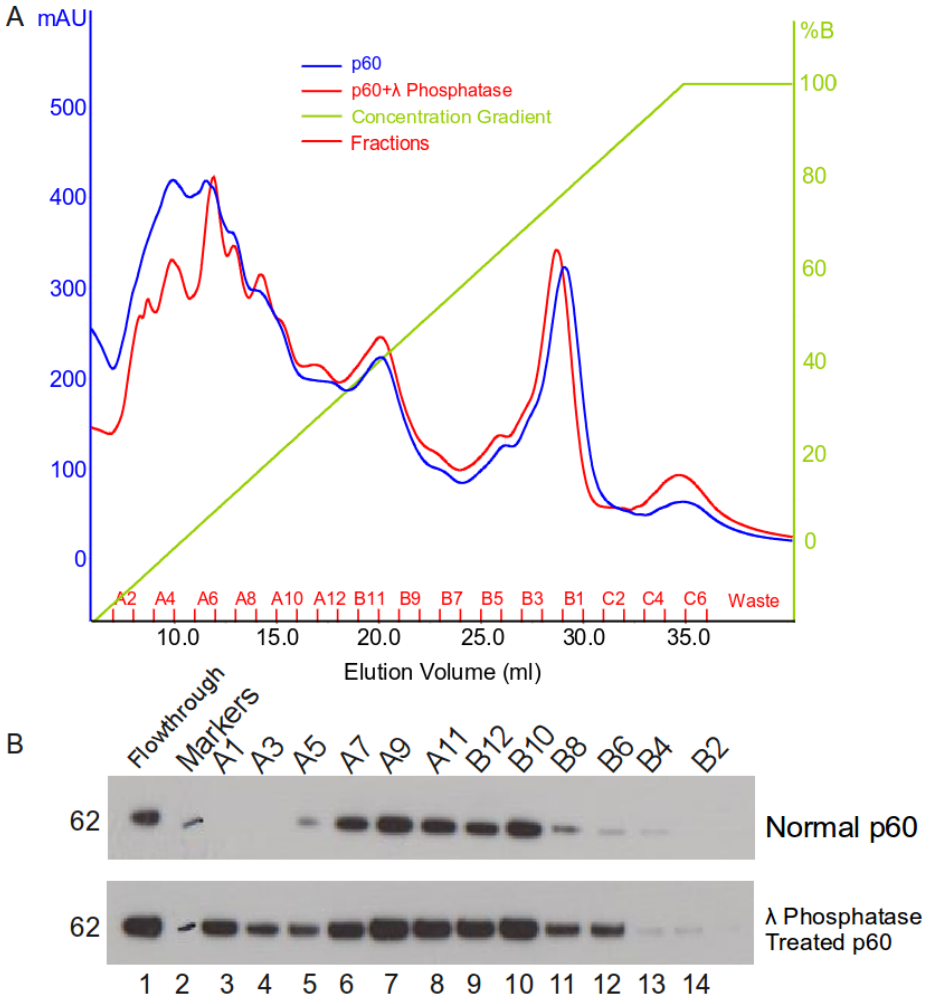


Figure 24: **Comparison of p60 Treated and Untreated with λ Phosphatase.**
An aliquot of p60 was treated with λ phosphatase and the elution profile after running through a Q HP column was compared to untreated sample. A: profiles of both p60 elutions from the anion exchange columns. B: Western blot of fractions from elutions visualised using an anti-p60 antibody. Fraction numbers are indicated on the top of the blot, markers are indicated on the right in kDa.

As it was not possible to further purify the p150ΔN/p60 complex by ion exchange, the Ni-NTA elution was optimised to remove as much contamination from the complex as possible (Figure 25). After using a spin column (Generon) for the Ni-NTA beads and 50 CV of lysis buffer without NP-40 as a wash step, the protein produced was more concentrated and cleaner giving a yield of approximately 4 mg total protein per litre of cells (estimated by Bradford protocol, see section 2.2.6.1). At least 75% (3 mg) of this is p150ΔN and p60 (estimated from gel, Figure 25). This protein exhibits no degradation of p150ΔN (see Western blot,

Figure 25). Concentration of this protein complex was still not efficient, hence gel filtration was not used to purify the complex further. This protein was used in the experiments detailed in Chapter 5. Crystal trials were set up as described in section 2.2.12 using E500 fraction 3 (well 8) but did not provide protein crystals. The protein likely needs to be cleaner and more homogeneous before crystallisation can be undertaken successfully.

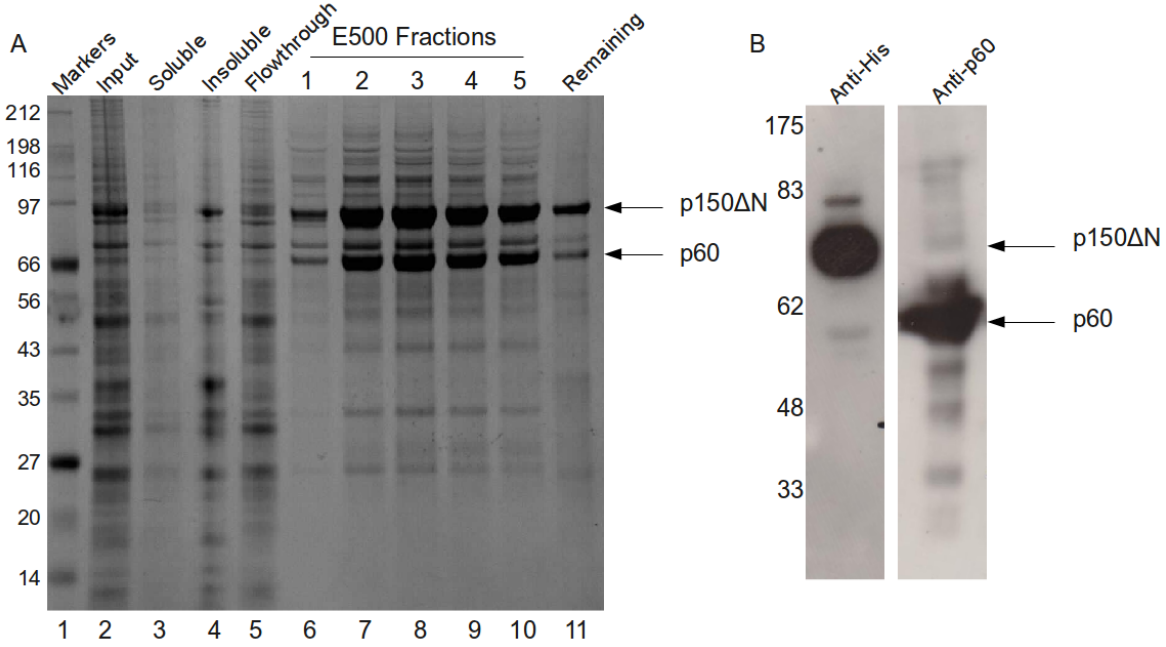


Figure 25: **Concentrated Preparation of p150ΔN/p60 from Ni-NTA Beads.** p150ΔN/p60 expressing cells from one litre of cell culture were lysed by incubation with NP-40 and loaded onto Ni-NTA resin by passing through a spin column. After extensive washing the bound proteins were eluted with buffer containing 500 mM imidazole A: SDS-PAGE analysis of the purification. B: Western blot of E500 fraction 3. Markers are on the right of each gel and are in kDa. Markers used for Western blotting are pre-stained and hence run at a slightly higher molecular weight.

3.4 Conclusions

In this chapter attempts to produce the CAF-1 complex by co-expression are discussed. It proved possible to express both the full complex and a binary complex containing p150 and p60 in insect cells using a baculovirus based co-expression system. The optimised expression was relatively low for the whole complex and slightly better for the binary complex. Purification strategies were examined and it was discovered that high salt (400 mM KCl or higher) was required to extract p150 from the cells and that it was possible to extract and partially purify p150ΔN in the presence of p60 with or without p48. This problem is also described in Hoek *et al.* (2011) and is believed to be due to the tight binding of p150 to chromatin.

Whilst the protein complex produced from these methods is not totally homogeneous it is a large improvement in quantity over previous work. 16.3 μg of CAF-1 was purified from 64 litres of human 293 cells (Smith and Stillman, 1989) and previous work in insect cells previously produced 0.75 μg of p150 (Kaufman *et al.*, 1995) from a volume of cells approximately equivalent to a 0.5 L cell pellet as grown in this study, compared to an yield of at least 3 mg of p150ΔN/p60 purified from 1 L of insect cells here. Whilst the Bradford method is not an accurate measurement of protein concentration it illustrates a significant increase in protein yield compared to previous work.

Work of other groups has focused mainly on protein produced either in cells at low levels (Takami *et al.*, 2007), tagged endogenous protein (Smith and Stillman, 1991; Nabatiyan and Krude, 2004), *in vitro* transcription translation (Quivy *et al.*, 2001) or small, synthetic peptides (Tang *et al.*, 2006). This higher yield purification if the purity of the final sample can be improved has the potential to allow studies of CAF-1 that were previously not possible with low-yield proteins, such as kinetic measurements and crystallographic studies.

Ion exchange of the complex and individual proteins revealed that p60 consists of a variety of charged species. These species may include both degradation products, post-translationally modified protein and complexes with other proteins. Full-length p60 contains a PEST box which targets proteins for degradation and the elution profile of p60 from an ion exchange

column does not improve on treatment with phosphatase, suggesting that whilst some protein may be phosphorylated it is not the main problem encountered during purification. As the PEST box is in the long unstructured C-terminal tail of p60 (Marheineke and Krude 1998; NCBI protein record Q13112) a method to improve the behaviour of the protein would be to make constructs lacking all or part of this tail. This is investigated in the next chapter.

Unexpectedly, there was very little C-terminal degradation of p150ΔN seen as Western blots, for example Figure 25 shows no ladder of degraded protein. The deletion of two known proteases from the baculoviral genome in the MultiBac system may account for this. As the major reason for expressing p150ΔN in a complex was to prevent this degradation it may be possible to express p150ΔN alone in this system. This offered the prospect of being able to examine the interactions of p150ΔN with other members of the CAF-1 complex and to be able to express further truncated constructs. Production of p150ΔN alone and of shorter constructs of p150 are described in Chapter 4.

Chapter 4

Individual CAF-1 Subunits: Constructs Design, Protein Expression and Purification

4.1 Introduction

After the investigation of co-expression of the CAF-1 complex it was thought that it would be possible to express the subunits of the complex individually. Previous work in this laboratory has shown that it is possible to produce pure p48 from insect cells (Lejon *et al.*, 2011). Work in this thesis has suggested that p150ΔN can be expressed without degradation and therefore may be stable alone and that p60 can also be separated from the complex but is subject to degradation.

Deletion mutants of p150 and p60 were also designed and constructs were produced for expression in *E. coli*. Constructs of p150 were designed to locate minimal binding regions to either p60 or p48 and constructs of p60 were designed to remove the C-terminal unstructured tail that we thought might have caused problems initially. As p48 has been stably expressed the established protocol could be used and intact protein produced at high quality and quantity.

For *E. coli* based expression of constructs the Ligation Independent Cloning (LIC) vector pNIC-Bsa4 was used as it allows for fast generation of constructs (Savitsky *et al.*, 2010). Insect cell expression was performed as detailed in chapter 3.

4.2 p150ΔN Production and Purification

P150ΔN was produced with the aid of the Berger laboratory at EMBL Grenoble using methods as described for the CAF-1 fusion protein. An initial purification using nickel resin was performed using the same protocol as for the fusion protein (see section 2.2.3.1), with and without the presence of urea to remove as much background as possible and to show that the protein was being purified *via* the His-tag. This is shown in Figure 26. The total protein yield appeared to be slightly higher than previously seen for the whole complex (compare lane 7 Figure 26with lane 10 Figure 16, section 3.2.2). As seen previously, p150ΔN was not soluble in low salt solutions and required 400 mM KCl to extract it from chromatin. The purified protein appears to be of a larger size than predicted when analysed by SDS-PAGE, running at around 100 kDa where expected to 77 kDa. This is not unusual as full length p150, named for its approximate mass by SDS-PAGE analysis, has an estimated mass of 107 kDa and unstructured proteins tend to run at larger molecular weights. Western blotting using anti-His antibodies was used to confirm that the band was His-tagged p150ΔN.

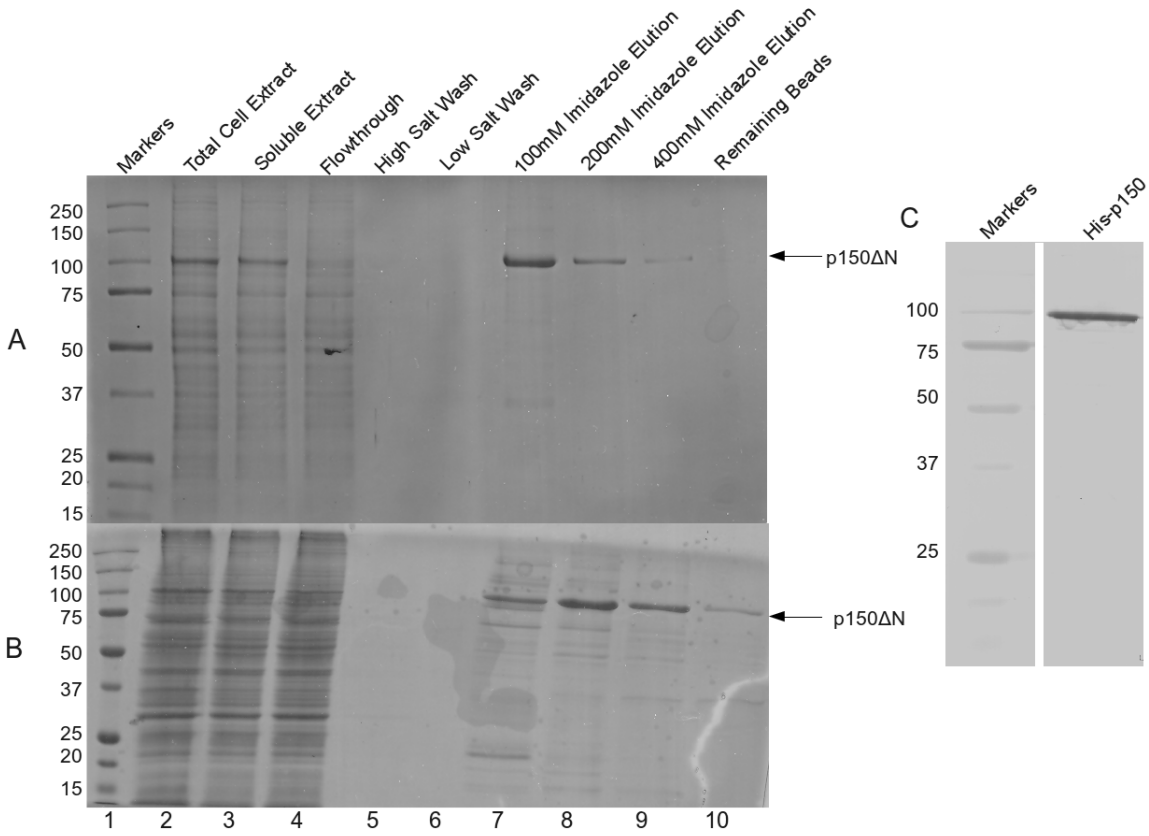


Figure 26: **Denaturing and Native Purifications of p150ΔN**.
A: SDS-PAGE analysis of the purification p150ΔN in the presence of 6M urea B: SDS-PAGE analysis of the purification of p150ΔN under native conditions. C: Western blot of the 100 mM Imidazole elution (lane 8), visualised using anti-His antibodies. Molecular weight markers are in kDa.

To examine the stability and quality of the purified protein gel filtration was performed, as shown in Figure 27. This shows some of the protein is bound to DNA as indicated by the 260:280nm ratio (first peak) but the second, smaller peak appears to contain protein of 70-80 kDa. As p150ΔN is around 70 kDa this is correct for a monomer.

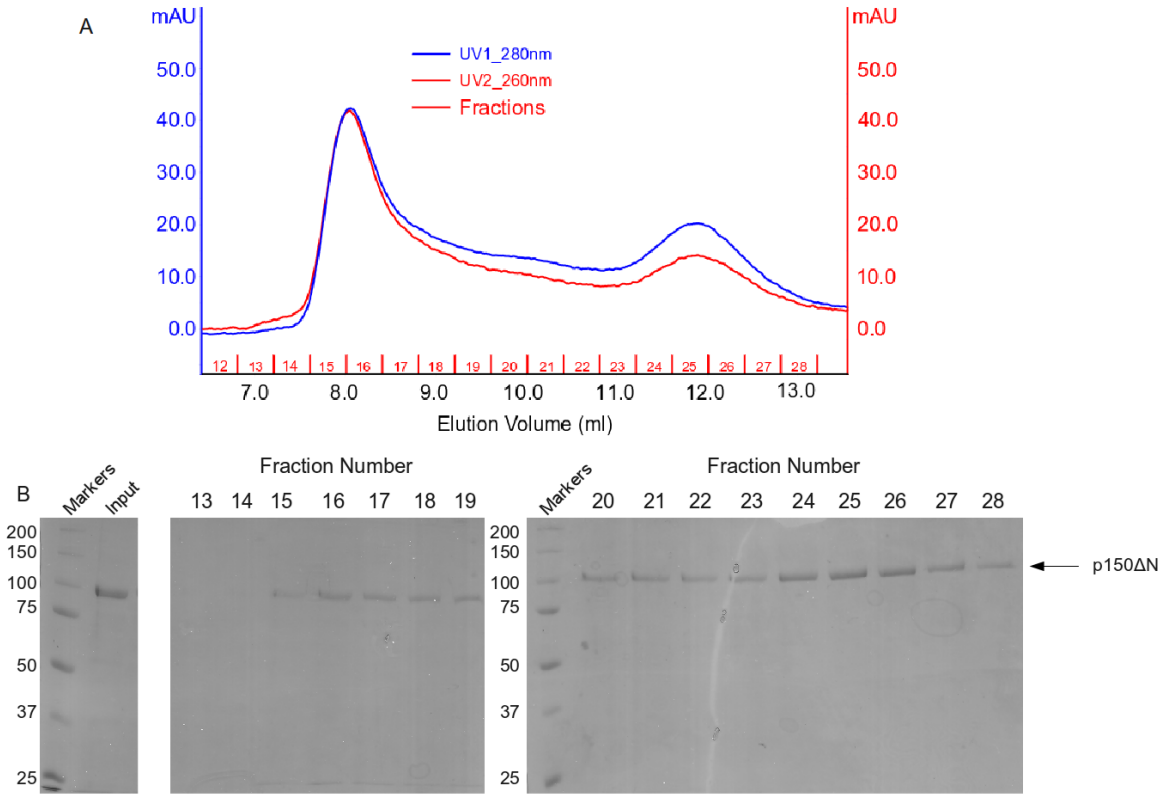


Figure 27: **Gel Filtration of Native p150ΔN**.
p150ΔN eluted from Ni-NTA resin was run on a S200 10/300 column, 0.5 ml fractions were collected. A: Elution Profile B: SDS-PAGE analysis100 µl of each fraction. Fraction numbers are indicated above the SDS-PAGE analysis. Markers are in kDa.

To scale up the production of p150ΔN a pellet from 0.5 L culture of cells was treated as described in section 2.2.3.3, however the protein eluted from the beads was dilute and contained many contaminants when concentrated (Figure28 lane 10). Attempts were made to purify p150ΔN further by ion exchange but this chromatography method proved impossible as p150ΔN is unstable in low salt and precipitated immediately. Gel filtration was possible, as seen above, however when concentrated by centrifugation p150ΔN precipitates onto the membrane. Alternative membranes and concentration methods were tried as for CAF-1 (section 3.2.3). However, neither changing the membrane, removing imidazole, addition of glycerol nor sucrose based dehydration allowed concentration of protein. In most cases the protein totally precipitated on the membrane and the final concentration of p150ΔN as seen by SDS-PAGE was lower than initially, e.g. Figure 27, lane 16 .

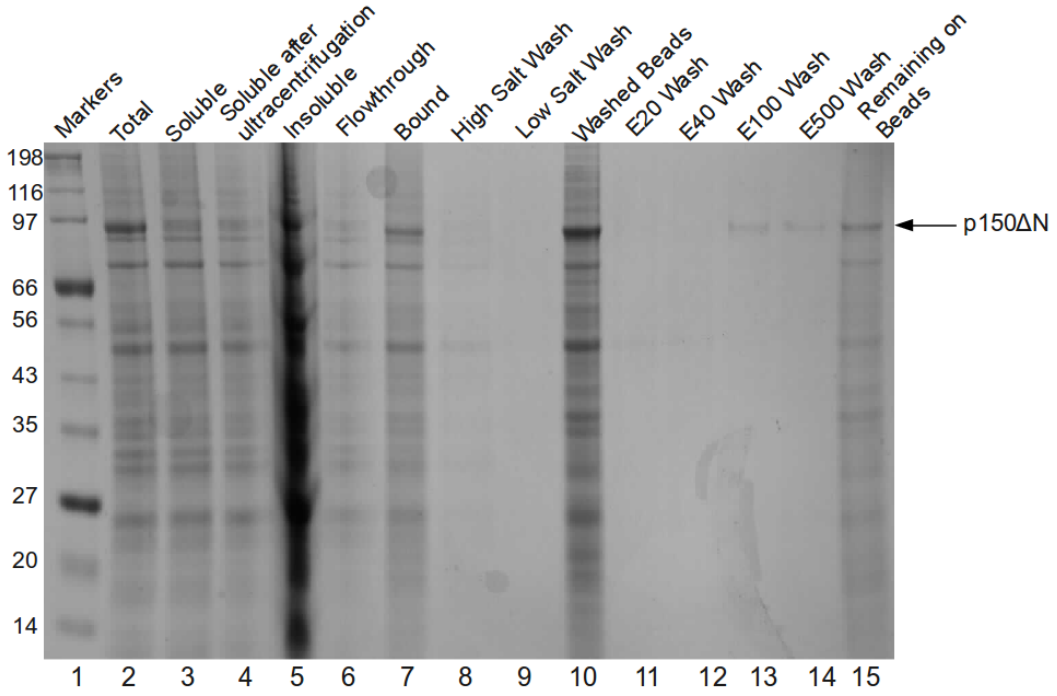


Figure 28: **Initial Purification of 0.5 L p150ΔN Pellet**.
SDS-PAGE analysis of purification. A 0.5 L culture of cells expressing p150ΔN was lysed, clarified by centrifugation at two speeds and purified by interaction with Ni-NTA beads. Markers are in kDa.

These results indicate that whilst it is possible to produce p150ΔN alone it is not a stable protein and has a tendency to aggregate. This was also seen when attempting to store p150ΔN at -80 ^{o}C after flash-freezing in liquid nitrogen as it aggregates on defrosting, meaning that it had to be freshly purified each time before use. Despite these problems sufficient protein was purified via nickel beads to use in pull-down and crosslinking experiments, as detailed in Chapter 5.

4.3 p150 Deletion Mutant Design, Production and Purification

4.3.1 Bioinformatic Analysis of p150

Bioinformatic analysis predicts p150 to be a mostly unstructured protein with no predicted motifs other than a coiled-coil region covering the KER rich region in the middle of the protein predicted by SMART analysis (http://smart.embl-heidelberg.de/). Constructs to isolate the p60 and p48 binding region(s) were designed based on sequence alignments of p150 from various species produced using Phyre (http://www.sbg.bio.ic.ac.uk/~phyre/), predicted structural motifs using PSIPRED (http://bioinf.cs.ucl.ac.uk/psipred/) and the information on the interactions between p150/p60 and p150/p48 published by Krude (1995) and Takami *et al.* (2007) (Figure 29). The resulting constructs A-D take into account the highly conserved regions seen between species with the removal of some repetitive regions. Construct E is based on the minimal region for interaction with p60 from Takami *et al.* (2007).

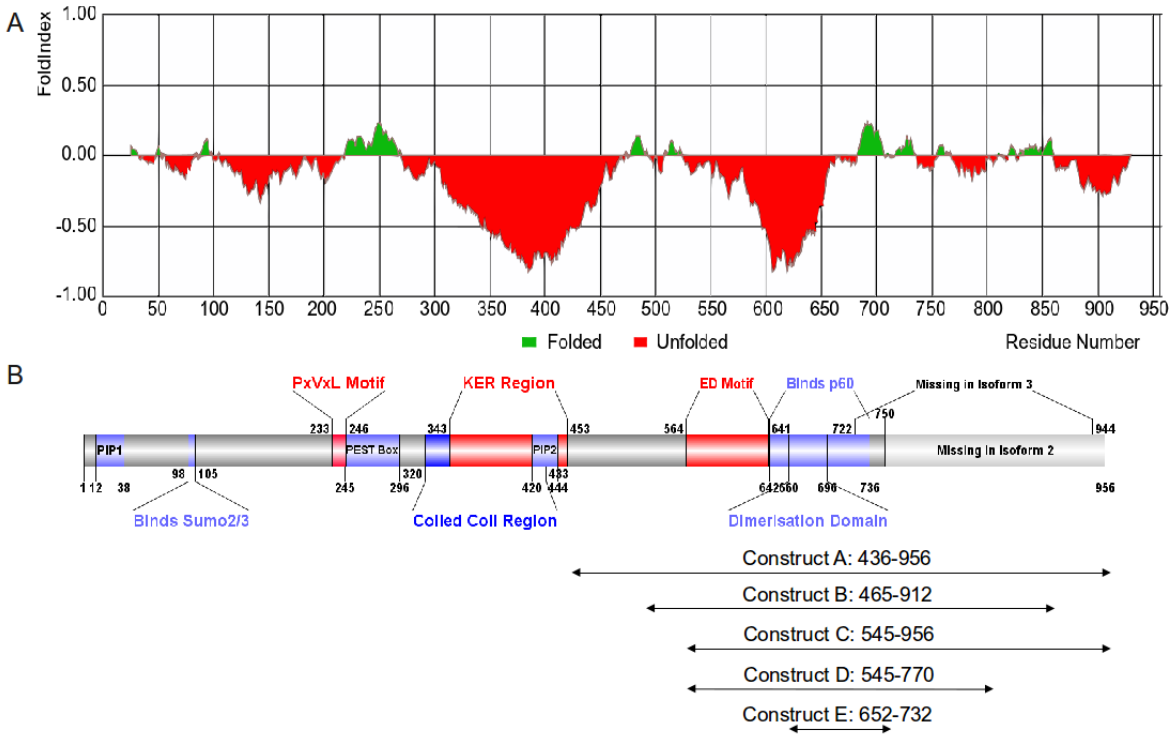


Figure 29: **Bioinformatic Analysis of p150 Sequence Showing the Constructs**.
A: Fold Index (http://bip.weizmann.ac.il/fldbin/findex, Prilusky *et al.*, 2005) prediction for p150, showing it is predicted to be mostly unfolded (red). B: Summary schematic of the p150 peptide where the constructs made are indicated.

4.3.2 PCR Primer Design

Longer p150 constructs were His-tagged to allow for large scale purification as required. Construct E is a short peptide and therefore was produced fused to a GST tag to increase solubility and stability in the cell. The LIC vector pNIC-Bsa4 was used for production of His-tagged proteins containing a TEV cleavage site. The GST-tagged protein was incorporated into pGEX-2T using the SLIC cloning method (Li and Elledge, 2007). Primers were designed as described in Chapter 2, and are shown in Appendix E.

4.3.3 Cloning

The p150 fragments were generated by PCR using the standard method described in section 2.2.1.9. Plasmids were digested with *Bsa*I (Constructs A-D) or *Bam*HI and *Eco*RI (Construct

E) as described in section 2.2.1.5 until less than 10 colonies were observed on transformation of the digested plasmid into Top10 cells (section 2.2.1.3). These were then treated according to the (S)LIC protocols (sections 2.2.1.12 and 2.2.1.13) and the resulting mixtures transformed into Top10 cells and plated on appropriate antibiotic. Colonies were grown overnight and checked by restriction digest and sequencing (section 2.2.1.15). Plasmids containing the correct genes were used in expression trials.

4.3.4 *E. coli* Expression Trials and Purification

His-tagged p150 constructs A-D were transformed into the *E. coli* expression strain Rosetta 2 pLysS (Invitrogen). 10 ml LB was inoculated with a single colony and grown to an O.D. (600 nm) of 0.4-0.6. The cells were then induced with 1 mM IPTG and grown at 20^{o}C for 3 hours. Samples were taken and lysed using Bugbuster, the soluble and insoluble fractions were analysed by SDS-PAGE and Western blot. Construct E could theoretically be expressed directly from Top10 cells as the GST vector contains a Tac promoter recognised by bacterial RNA polymerase. However, cells containing construct E would not readily grow in suspension culture and when induced were seen to lyse, indicating the construct E is lethal to *E. coli*, therefore only the His-tagged constructs were taken forward for further purification.

The His-tagged proteins were expressed in larger culture volumes and used for purification. A protocol similar to that used for the purification of p150 from insect cells (section 2.2.3.3) was applied but with a lower salt concentration in the lysis and purification buffers as the proteins were soluble when lysed with Bugbuster. Each cell pellet was lysed, spun and the soluble supernatant run over 1 ml Ni-NTA beads. A SDS-PAGE gel of the purification results is shown in Figure 30. Constructs A-C appear to have been expressed and purified at a low but usable level (lanes 6, 12 and 18). Construct D did not appear to be present above background levels (lane 25) and was therefore not used in further experiments. When analysed by SDS-PAGE the proteins appear to be larger than predicted, as has been seen previously for p150ΔN. The proteins were shown to contain a His-tag by Western blot (data not shown). These proteins were used for experiments in chapter 5.

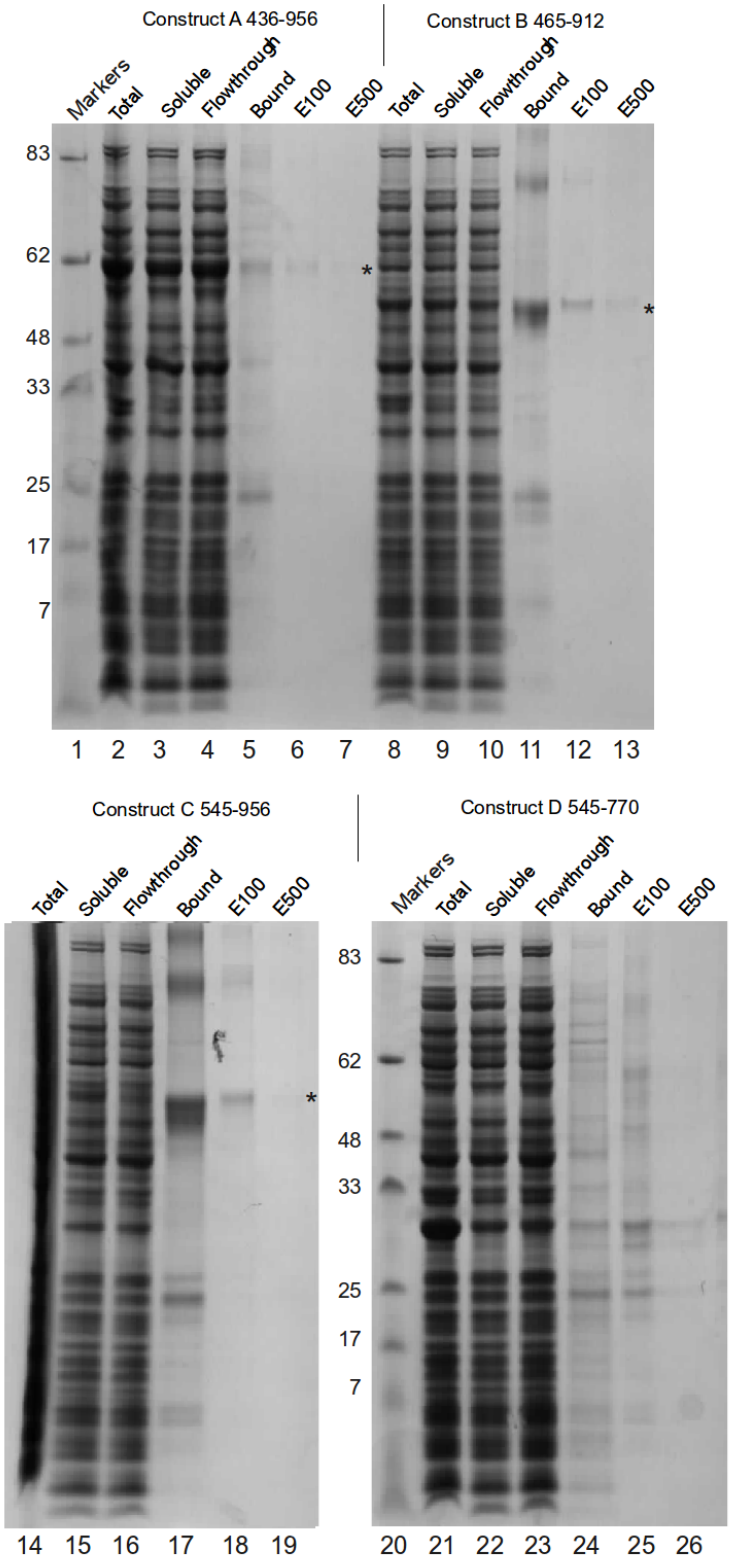


Figure 30: **Initial Purification of p150 Deletion Mutants.**
Each construct was purified by binding to Ni-NTA beads, washed with 1 M KCl and eluted with increasing concentrations of imidazole. Protein samples were analysed by SDS-PAGE. * indicates the protein of interest, purified proteins were shown to contain a His-tag by Western blot (results not shown). Construct D has not been purified. Markers are in kDa.

4.4 p60 Constructs

4.4.1 Bioinformatic Analysis of p60

To assess the conserved domain composition of p60 a variety of bioinformatic analyses were performed as for p150. P60 is predicted to be a WD repeat containing protein which would form a beta propeller structure similar to that of p48 with an unstructured tail which is known to interact with ASF-1 (Malay *et al.*, 2008). Visual representations of some of this data and the chosen constructs are shown in Figure 31. Using this data five constructs were designed. Three constructs containing the WD repeat region, a construct containing the C-terminal tail and full length p60. The whole of p60 was sampled to make constructs as it is currently unknown which section interacts with p150 and the various end sites for the WD containing constructs reflect the uncertainty as to where the final WD repeat needs to be cut to maintain a stable structure.

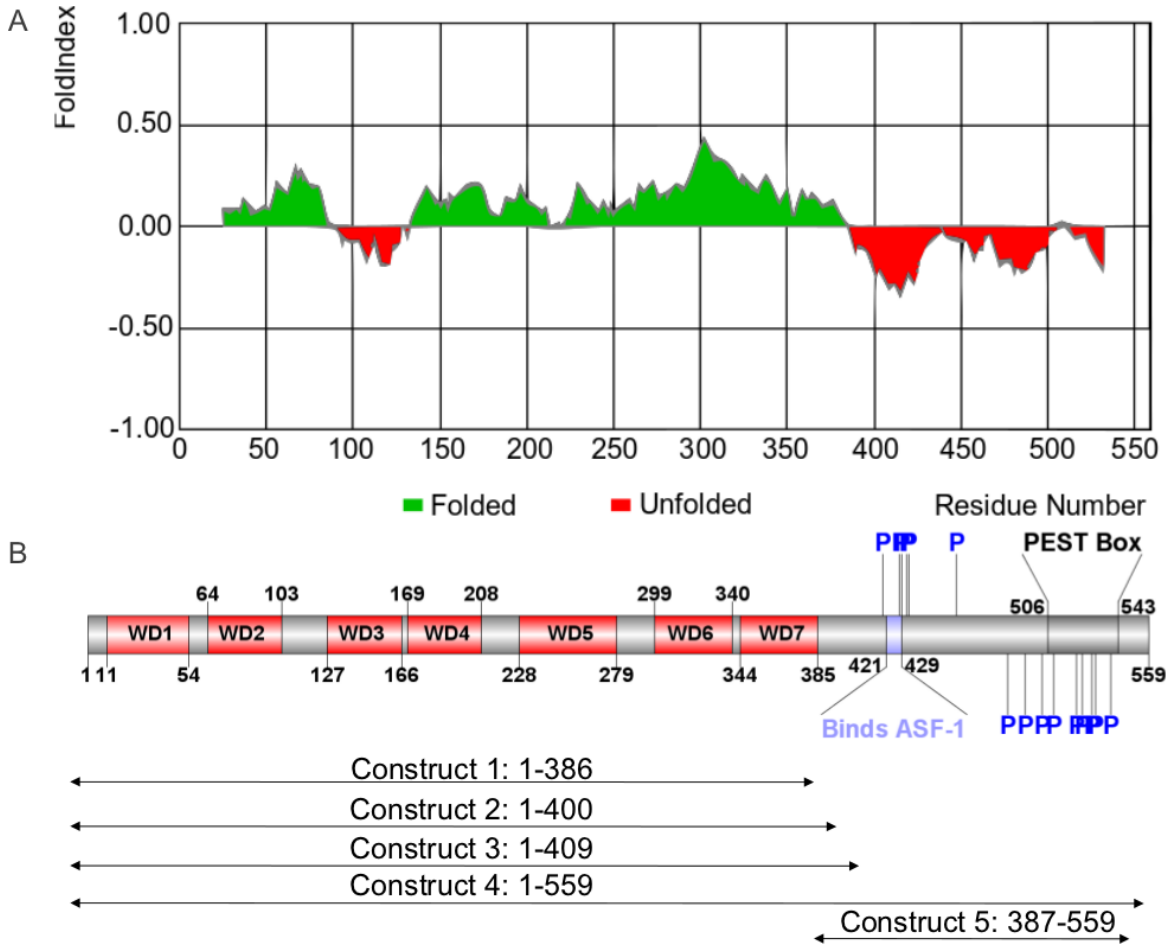


Figure 31: **Bioinformatic Analysis of the p60 Sequence Showing the Location of Constructs.**
A: FoldIndex prediction of structured regions of p60 (Prilusky *et al.*, 2005). B: peptide diagram summarising key features of p60. WD denotes predicted locations of beta propeller blades. P, predicted phosphorylation sites. Below the protein the constructs chosen for cloning are indicated, the numbers denote the amino acid residues.

4.4.2 *E. coli* Expression

4.4.2.1 PCR Primer Design

Constructs were to be cloned into both *E. coli* and insect cell expression vectors. Previous work showed that full length p60 did not express solubly in bacteria (N. Murzina, personal communication). For expression in *E. coli* the plasmid pNIC-Bsa4 was used. Primers were

designed to contain the LIC sequence followed by a section of p60 and are described in Appendix E.

4.4.2.2 Cloning

The p60 fragments were synthesised by PCR using the basic protocol described in section 2.2.1.9. The reactions were purified with a PCR clean up kit and used for LIC reactions. The vector pNIC-Bsa4 was linearised by treatment with *BsaI*. Constructs and linearised vector were treated with T4 DNA polymerase and combined. The DNA was transformed into Top10 cells and colonies growing on LB-agar supplemented with kanamycin and sucrose were isolated, grown overnight for minipreps and confirmed by sequencing to contain the correct gene.

4.4.2.3 p60 *E. coli* Expression Trials

Several p60 expression trials were undertaken using Rosetta and Arctic Express *E. coli* strains. Cells were grown at 37 ^{o}C for 3 hours, 27 C for 8 hours and 15 ^{o}C overnight with IPTG concentrations of 1, 0.4 and 0.1 mM IPTG. None of these conditions produced soluble protein and as this had previously been the case with p48 in this laboratory it was decided to concentrate on insect cell expression. Figure 32 shows an example of an expression profile from the expression in Rosetta cells at 27 ^{o}C for 8 hours using 0.4 mM IPTG.

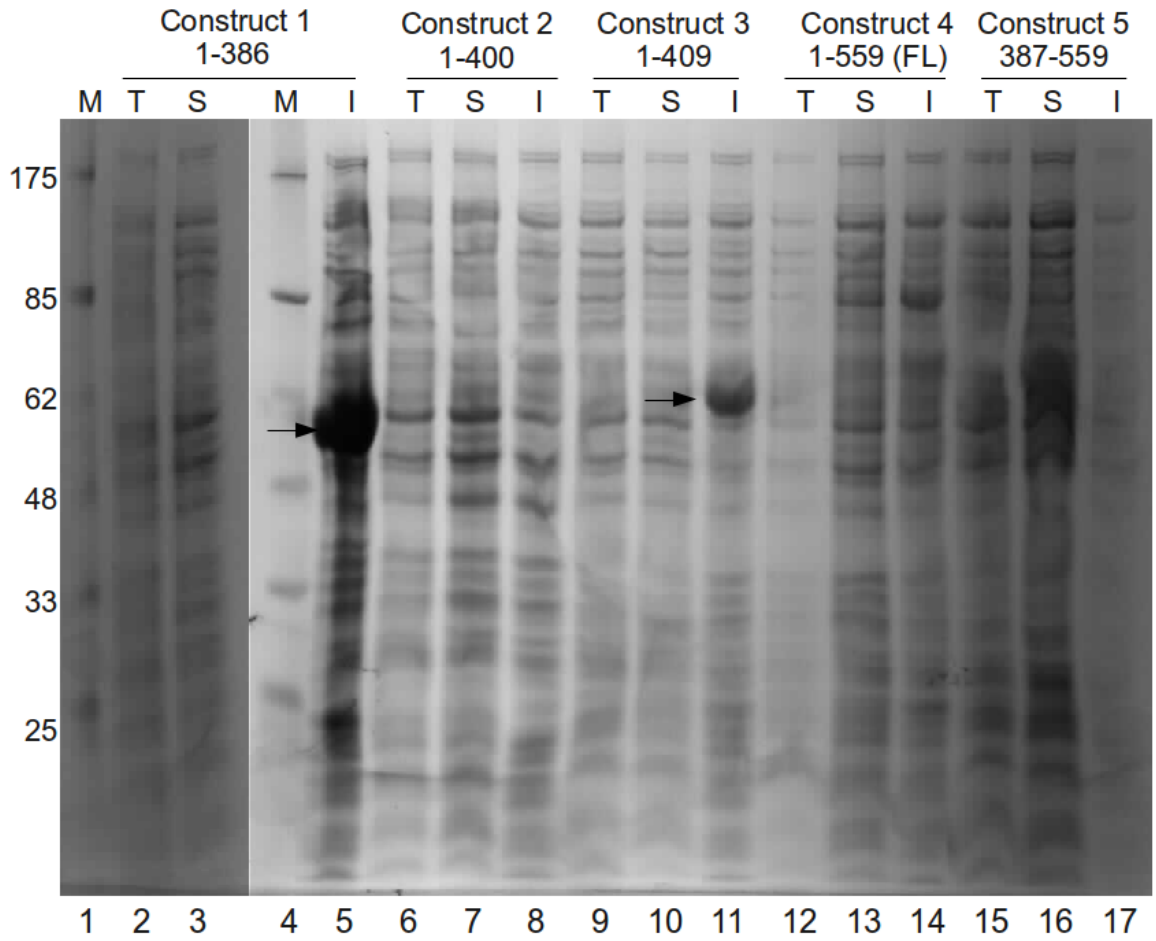


Figure 32: **Trial Expression of p60 in *E. coli*.**
Rosetta cells containing p60 expression plasmids were grown at 27 ^{o}C for 8 hours and induced with 0.4 mM IPTG, lysed with Bugbuster and spun down to separate soluble and insoluble proten. Protein samples were analysed using SDS-PAGE. T = Total cell extract, S = Soluble and I = Insoluble. Expected sizes are 47kDa for construct 1 (1-386), 49 kDa for construct 2 (1-400), 50 kDa for construct 3 (1-409), 66 kDa for construct 4 (1-559, full length) and 22 kDa for construct 5 (387-559). Black arrows indicate insoluble constructs 1 and 3. Constructs 2, 4 and 5 are less obvious.

4.4.3 Insect Cell Based Expression of p60

As *E. coli* based expression did not produce soluble protein the constructs were cloned into the MultiBac system for expression in insect cells (Berger *et al.*, 2004). To allow for future work co-expressing p60 with other proteins, primers were designed to clone constructs 1-5 into pUCDM containing a Myc-tag. For this we used the In-Fusion cloning system which

allows for sequence and ligation independent cloning (SLIC). Primers were designed to insert a Myc-tag into the *Bam*HI site of pUCDM and the gene of interest between the *Bam*HI and *Xba*I restriction sites of the vector and are detailed in Appendix E.

4.4.3.1 pUCDM Plasmid Preparation

The MultiBac system consists of two plasmids: pFBDM which has a standard *E. coli* origin of replication and pUCDM, which contains the conditional origin of replication R6Kγ that requires cells to express the *pir*+ gene product for plasmid propagation. The *E. coli* strain BW23474 was used for this propagation but initially provided very low yields of plasmid (<20 ng/μl on elution in 30 μl water). A variety of different miniprep kits were tried (Macherny-Nagel, Anachem and Qiagen) but did not improve the yield. Midipreps (Qiagen) did produce larger quantities of DNA, but at similar concentrations (up to 20 ng/μl), and were unreliable with some preps failing to produce any DNA. Alkaline lysis (section 2.2.1.1) produced large amounts of DNA but did not remove low molecular weight nucleic acids even when large amounts of RNAse was added to the reaction. To increase the concentration of DNA produced from a midiprep the lysed and centrifuged solution (stage 7 in the Qiagen manual) was passed through a miniprep spin column (rather than a midiprep column) and eluted in 50 μl water, this produced reasonable quality DNA with a 260/280 nm ratio greater than 1.8. The results from both methods is shown in Figure 33.

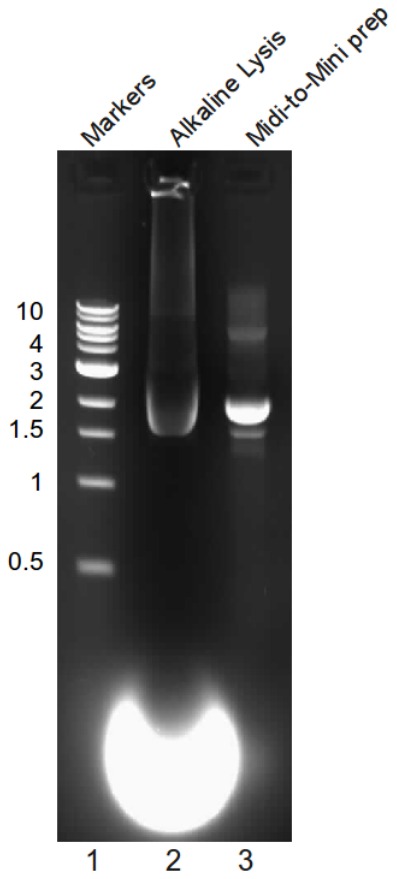


Figure 33: **Alternate Preparation Methods of pUCDM.**
Agarose gel analysis of pUCDM samples. DNA was prepared from BW23474 cells containing pUCDM using the alkaline lysis method and the midi-to-mini prep method where the contents of a midiprep was transferred to a miniprep spin column and treated as a miniprep.

When sufficient DNA had been produced, digestions were set-up (section 2.2.1.5) but it was discovered that the plasmids did not digest to completion. The majority of the pUCDM plasmid digested except for a supercoiled sub-population that is not removed even on addition of a second (or third) portion of enzyme. Increasing the enzyme concentration and digestion time did not make any difference. Different sources of pUCDM were examined but the lower yield BW23473 cells, pir+ cells from Epicentre Biotechnologies and a fresh batch of BW23474 cells from the Berger laboratory both produced very low yield plasmids which also were not susceptible to digestion.

After consultation with I. Berger it was also suggested that pUCDM containing cells should be grown for 8 hours and harvested whilst in exponential phase rather than stationary phase.

A comparison of these two methods are show in Figure 34. Whilst the yield and purity of the plasmid is affected, this DNA did not digest readily either. However the DNA from the concentrated midiprep could be digested to a certain extent and was therefore used for the Myc cloning experiment as it was sufficient to make colonies containing the correct inserts.

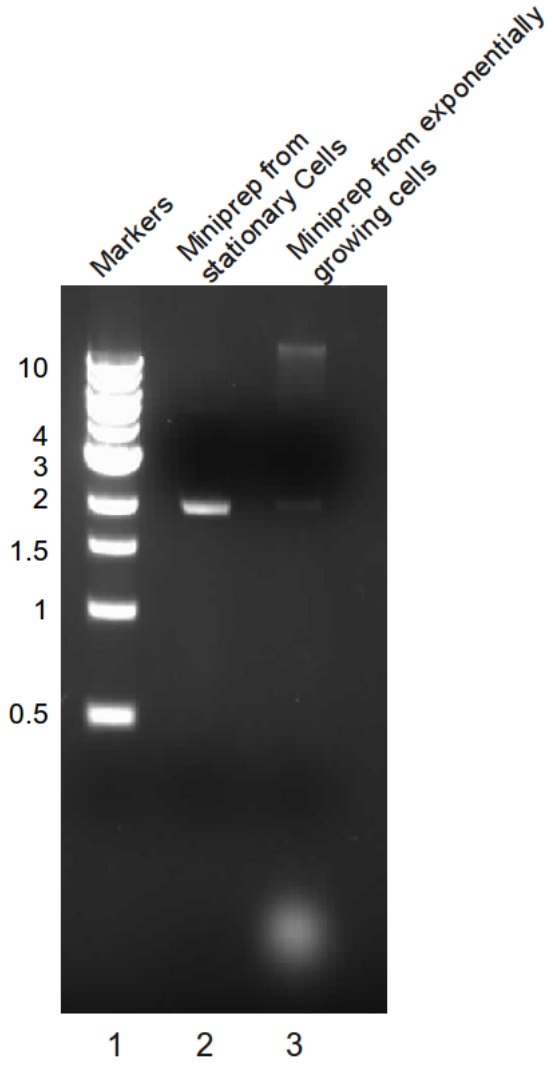


Figure 34: **Comparison of Exponentially Growing and Stationary Phase Cells for Minipreps.**
DNA was prepared from cultures grown overnight or for 8 hours and compared by agarose gel electrophoresis. Markers are kbp.

4.4.3.2 Insertion of Myc Tag into pUCDM

The sequence encoding a Myc-tag following a thrombin cleavage site was inserted at the *Bam*H1 cleavage site of pUCDM, introducing an *Nde*I cleavage site which is absent in pUCDM. Complimentary oligos containing the Myc-tag were designed (Appendix E) and annealed by plunging a mixture of oligos into a beaker of boiling water and leaving them to cool slowly and anneal. Annealing was checked by 8 % (w/v) acrylamide gel and the DNA was ligated into pUCDM prepared as described above, cut with *Bam*HI and cleaned using a PCR clean-up kit. The ligation reaction was transformed into *pir*+ cells and left to grow overnight. There were many colonies seen on control plates due to poor digestion of pUCDM, but as ligation plates contained around ten times as many colonies as the controls, colonies were picked, grown overnight and the plasmid DNA prepared. Plasmids were then digested with *Nde*I and *Pme*I, giving a 600 bp fragment in correct integrants, shown in Figure 35. Colonies appearing correct by digestion analysis were confirmed by sequencing with a PolH primer. The final pUCDM-Myc plasmid contained a double insert of the Myc fragment. Nevertheless it was used for p60 cloning as the double direct Myc repeat did not introduce a reading frame shift.

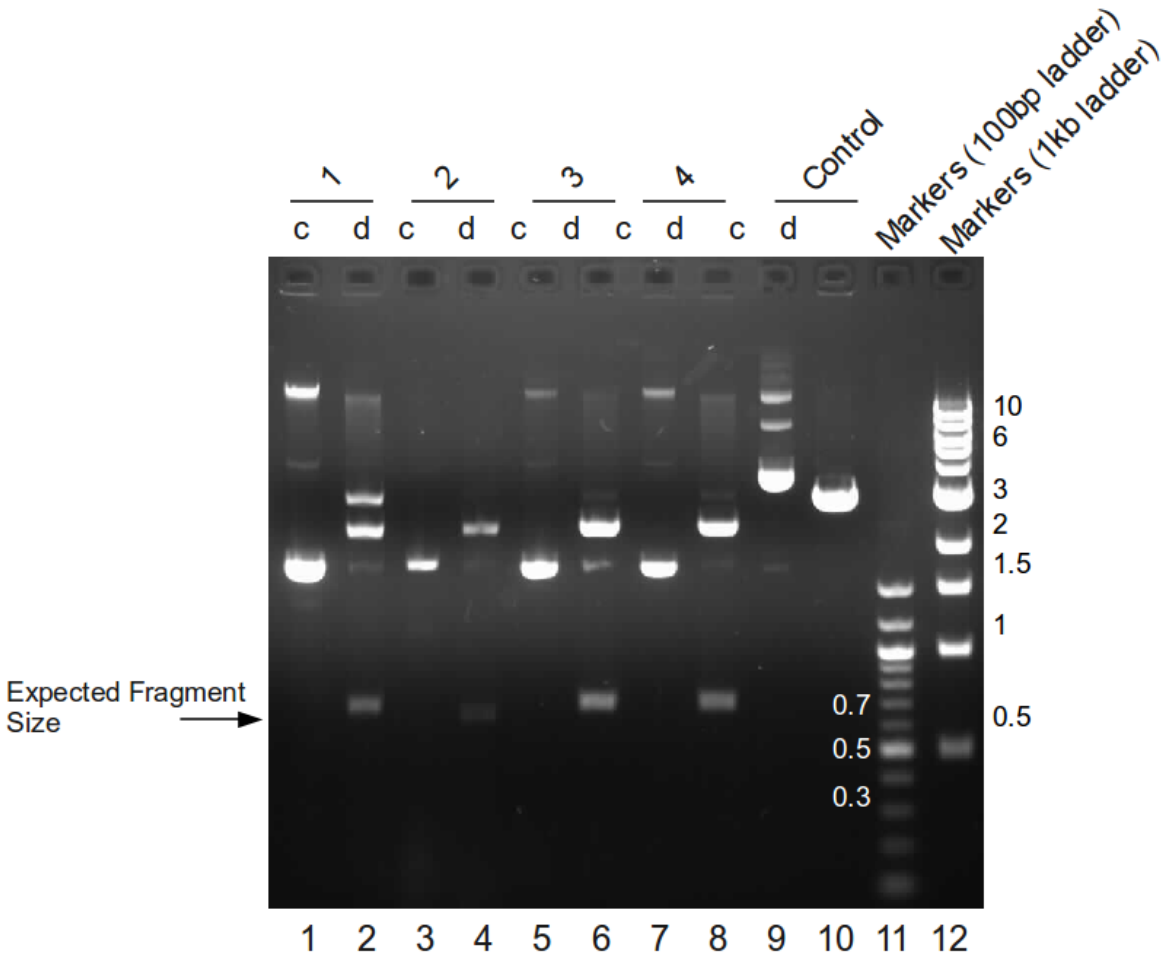


Figure 35: **Digestion of Possible Myc Integrants**.
Plasmids were digested with *Nde*I and *Pme*I to check for inserted fragments and separated on a 1 % agarose gel. c = circular, d = digested with *Nde*I and *Pme*I. Expected band size is 0.6kb. Marker is in kb.

4.4.3.3 PCR of p60 Fragments

The p60 fragments were synthesised by PCR. The basic protocol described in section 2.2.1.9 amplified all the fragments to sufficient concentrations of DNA when the number of cycles was increased to 30. The reactions were cleaned with a PCR clean up kit and used for In-Fusion reactions with pUCDM-Myc.

4.4.3.4 In-Fusion of Fragments with Vector

Linearised pUCDM-Myc digested with *Bam*HI and *Xba*I was combined with fragments of p60 and integrated using an In-Fusion reaction. As described previously, pUCDM is resistant to linearisation and a large number of colonies were seen on control plates. To prevent many minipreps being performed colony PCR using the PolH sequencing primer and a specific reverse primer was used to identify correct colonies as described in section 2.2.1.14. Once a fragment was identified the colony was grown overnight and a miniprep performed for analysis of plasmids by DNA sequencing.

4.4.3.5 *Lox*-Based Recombination and Bacmid Creation

As p60 was cloned into the donor vector of the MultiBac system it was necessary to recombine it with an acceptor vector to provide the Tn7 recombination element required for insertion into Baculoviral DNA. The various p60 constructs were combined with pFL from the iACEMBL derivatives if the MultiBac plasmids (section 2.2.1.17). pFL is an acceptor vector very similar to pFBDM, containing a *lox* site which allows combination of two plasmids containing identical sites into one on addition of the Cre enzyme. Recombined plasmids were selected for by plating onto agar containing both ampicillin and chloramphenicol and transforming into cells lacking the *pir*+ origin of replication. Colonies were grown overnight and the plasmid DNA extracted from cells. The correct combinations of plasmids was confirmed by restriction digestion. Recombined plasmids were then integrated into the baculoviral genome by transforming into EmBacY cells and blue/white selection performed to check integration into the bacmid (section 2.2.2.2).

4.4.3.6 Baculoviral Expression

The bacmid was transfected into adherent *Sf21* insect cells and a virus generated. This virus was amplified and the YFP time course followed, as described in section 2.2.2.2. Once YFP production had peaked a Western blot was performed to check for the presence of Myc-tagged protein, shown in Figure 36. Each construct had two separate viruses made, noted

as n and n'. Construct numbers refer to the constructs shown in Figure 31. Using anti-Myc antibody it was apparent that the full length version of p60 exists in several forms (lanes 8-11) whereas the truncated version 1 appears to have only one form (lanes 2-5). Only two constructs expressed, full length p60 and the most truncated version of the WD region, construct 1. It is thought that construct 5 may have degraded in cells rapidly due to the presence of the PEST box motif. All non-expressing constructs were prepared again from sequenced pUCDM but still did not express. Probing the blot with anti-p60 only visualised full-length protein, indicating that the recognised antigen is in the unstructured C-terminal tail. Therefore anti-Myc was used to visualise construct 1 in further experiments. As only construct 1 expressed it is hereafter known as p60_WD as it contains the WD region of the protein. Full length p60 is denoted p60_FL.

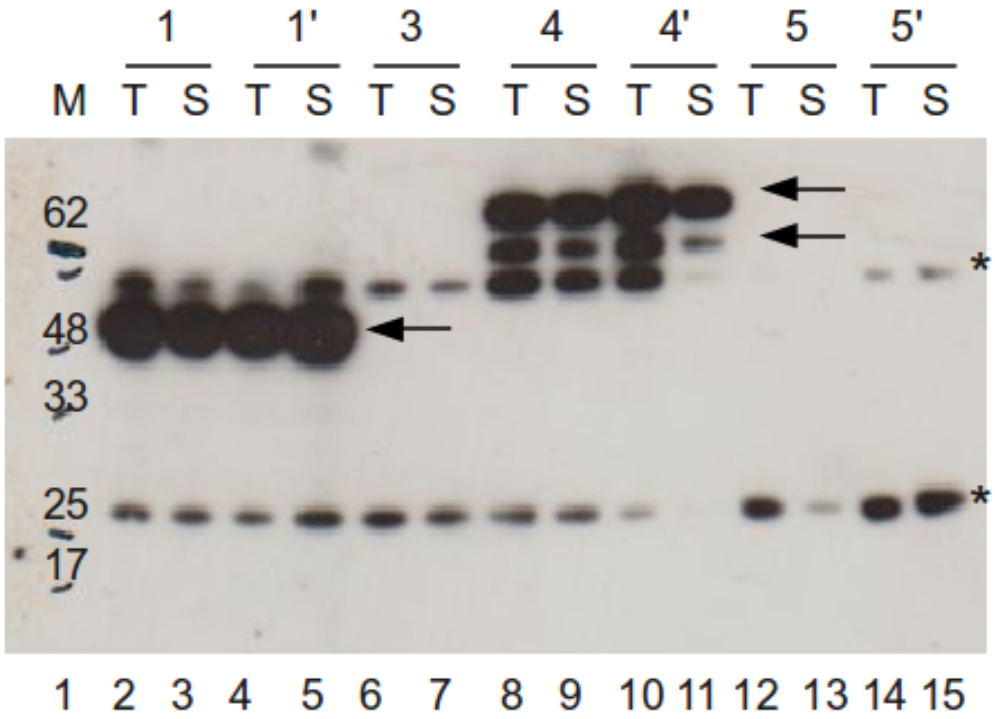


Figure 36: **Anti-Myc Visualisation of Myc-p60 Constructs.**
Western blot of of Myc-p60 constructs expressed in *Sf21* insect cells. Each construct was made into two viruses noted n and n'. Numbers above the gel refer to constructs as shown in Figure 31. Markers are in kDa. Asterisks (*) indicate non-specific antibody activity to an unknown protein . Arrows indicate expressed protein. M = Marker, T = Total cell lysate, S = Soluble lysate.

4.4.3.7 Purification of p60 _FL

The initial purification of p60_FL for pull-down experiments was performed by lysing cells using NP-40, clarifying the lysate by centrifugation and incubating with anti-Myc antibodies attached to beads, as described in section 2.2.3.4. The beads were harvested from the solution by centrifugation and washed. SDS-PAGE analysis of the purification is shown in Figure 37. The presence and gel position of p60_FL was confirmed by Western blotting (data not shown) and beads were then used for pull-down experiments. To remove p60_FL from the beads the protein was cleaved overnight with thrombin and then the thrombin removed by incubation with anti-thrombin beads. The concentration of protein pulled down using Myc-beads is low as the beads have a low binding capacity.

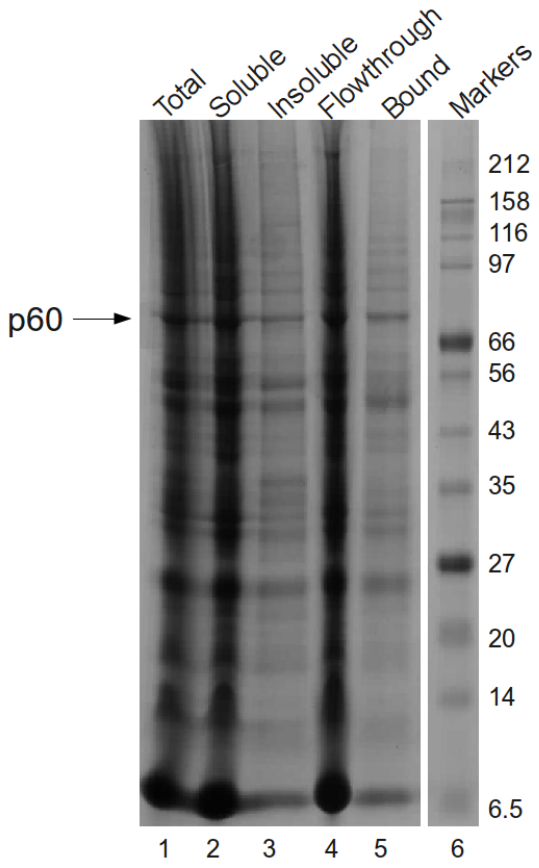


Figure 37: **Purification of Myc-p60_FL using Anti-Myc Beads**.
SDS-PAGE analysis protein samples collected during Myc-p60_FL purification. The presence of p60_FL was confirmed by Western blotting using anti-Myc and anti-p60 (data not shown). Markers are in kDa.

4.4.3.8 p60_WD Purification

Purification of Myc-tagged p60_WD (Construct 1 in Figure 31) was attempted using methods other than affinity purification as the Myc-beads are not re-usable and are prohibitively expensive in amounts sufficient for large scale purification. The Myc-tag was however used to visualise p60_WD. It was thought that as Construct 1 appears to only have one form on a Western blot (Figure 36, lanes 2-5) it would behave better than full length p60 during purification. Initial fractionation of the clarified lysate was done using a Q HP column (Figure 38).

Fractions A3-A9 were combined for further purification was attempted using a hydroxyapatite column, to which neither p60_WD nor its contaminants bound and then by gel filtration, shown in Figure 39. This separated p60_WD into several species, as seen by Western blot, including a heavy species containing DNA (peak with high absorbance at 260 nm), however these were not visible on a Coomassie-stained SDS-PAGE gel (Figure 39 lanes B15 and B14). A second ion exchange column was then run on gel filtration fractions B15 and B14 in an attempt to purify further and concentrate p60_WD but it did not bind well to the column and was at a very low concentration.

The amount of p60_WD produced in these cells is low and for maximum efficiency of future experiments it would be desirable to move the gene into a his-tagged vector for easier purification before optimising the virus for better expression. However the p60_WD purified on anti-Myc beads using small scale methods was sufficient for pull-down experiments.

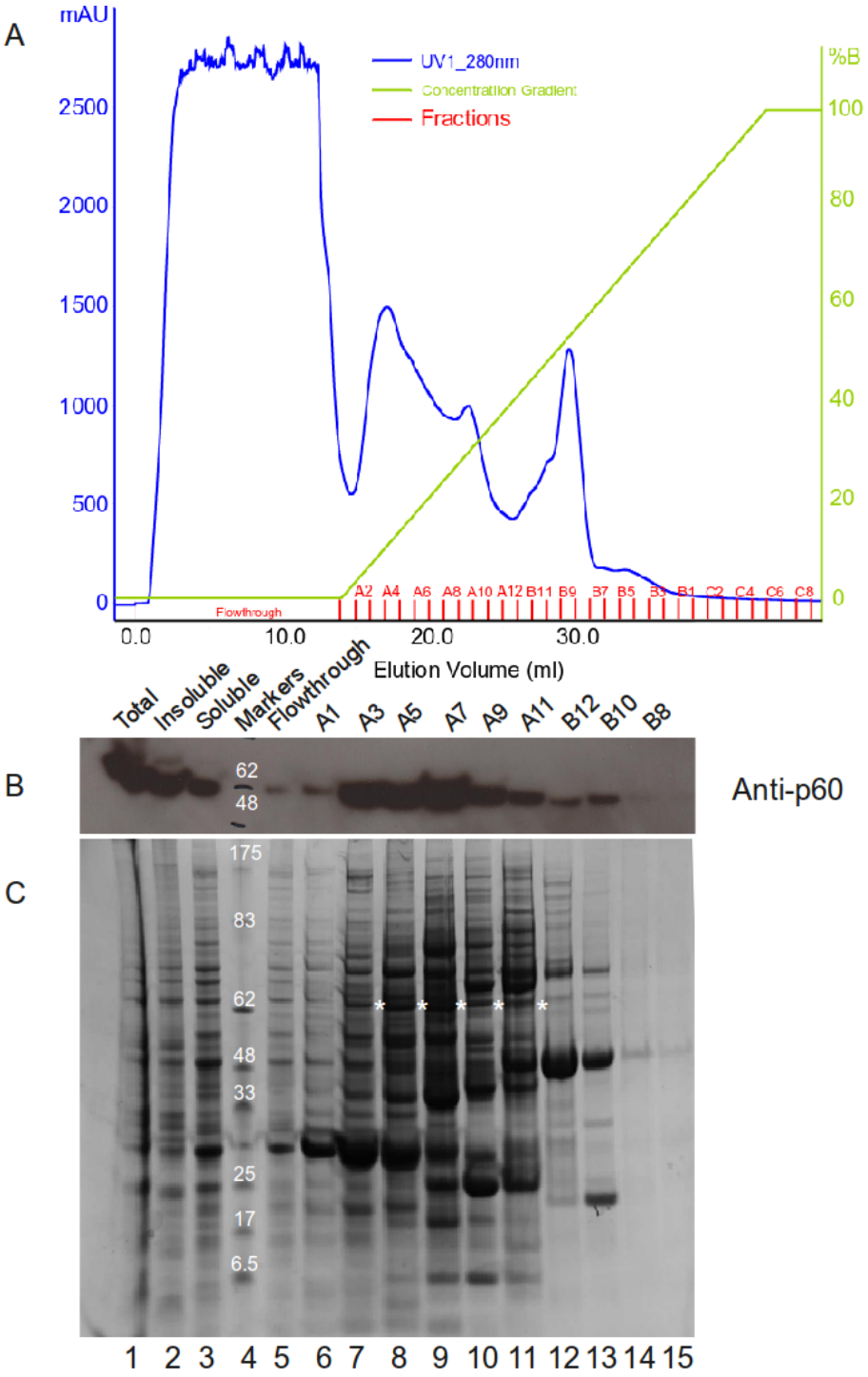


Figure 38: **Initial Purification of p60_WD.**
A: Elution profile of clarified whole cell extract containing p60_WD run on a 1 ml Q HP ion exchange column over 30 CV. 1 ml fractions were collected. B: Western blot showing anti-Myc visualisation of the fractions. C: SDS-PAGE gel analysis of fractions, suspected p60 bands are indicated by asterisks. Markers are in kDa.

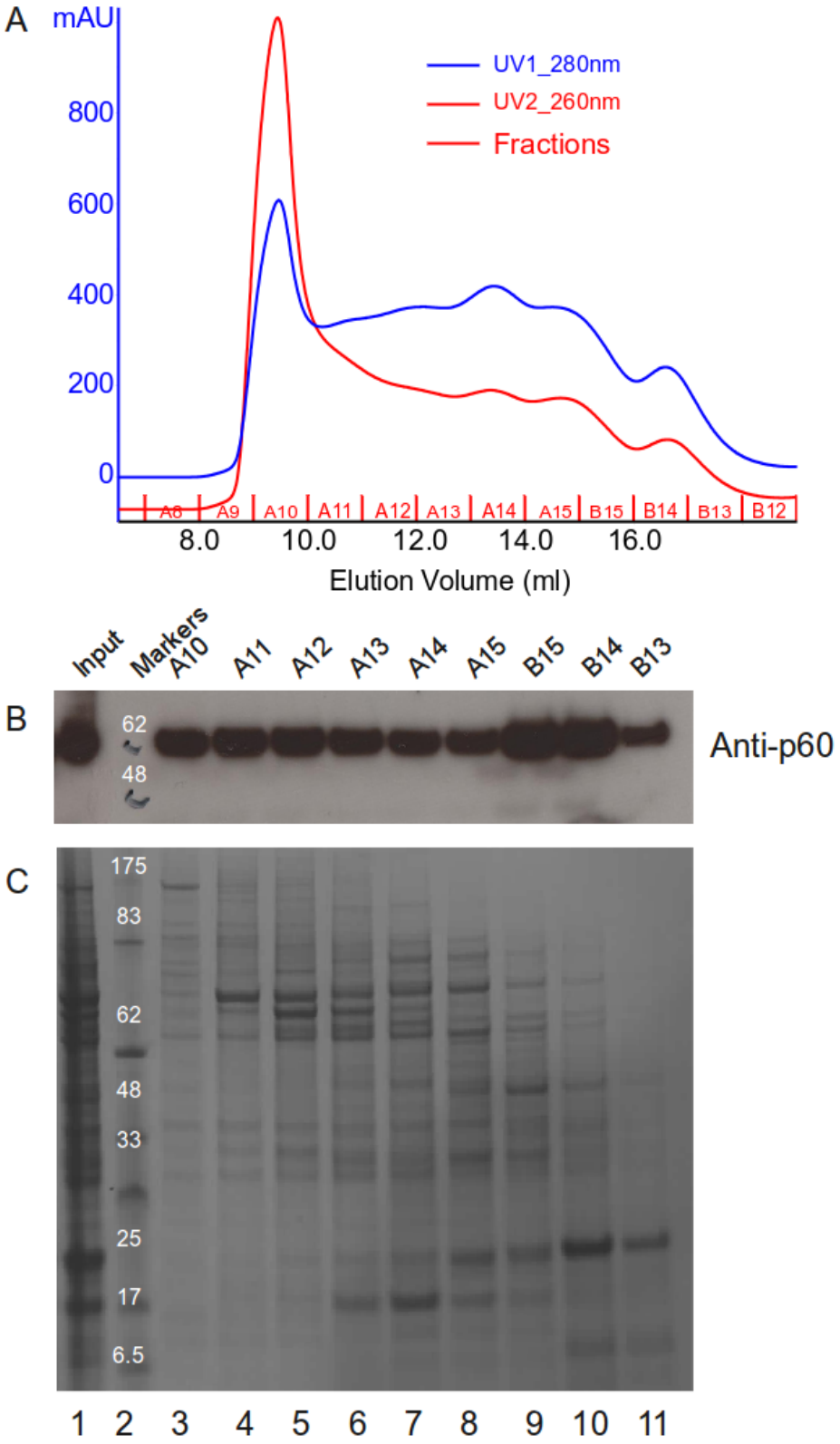


Figure 39: **Gel Filtration of p60_WD.**
A: Elution profile of Fractions A3-A9 from the Q HP column run on a S200 10/300 gel filtration column B: Western blot using anti-Myc to visualise p60_WD. C: SDS-PAGE analysis of fractions from column. Markers are in kDa.

4.5 p48

P48 had previously been produced in our laboratory to a very high purify using a protocol designed by N. Murzina and W. Zhang. Protein used in these experiments was kindly purified by S. Lejon, A. Murthy or W. Zhang, according to the protocol detailed in section 2.2.3.5. Purified protein was flash frozen in 10 % (v/v) glycerol and stored at -80^{o}C. This was defrosted on ice before use in assays. Previously frozen p48 can be crystallised (S. Lejon, personal communication), which indicates that it is not damaged by the freezing and defrosting process.

4.5.1 His-p48

His-tagged p48 was purified in the same way as normal p48 with the exception of the thrombin cleavage step, which was skipped. Where His-p48 was required bound to beads the purified protein was re-bound to Ni-NTA resin after purification and elution.

4.6 Conclusions

It has proved possible to express all the components of the CAF-1 complex in insect cell based baculovirus expression systems, albeit in low amounts for p150 and p60. Purification of p150ΔN was only possible on a nickel affinity resin before the protein began to aggregate, indicating that is it not very stable alone, possibly due to being largely unstructured. Work in the previous chapter indicates that this problem might be solved by addition of other components of the complex which stabilise the protein enough to make concentration of the complex possible. It may be possible in future to purify the other components of the complex separately then add them back to p150 to produce more concentrated, cleaner complex than is possible with co-expression. Shorter constructs of p150 have been successfully produced in *E. coli* to allow investigation of their binding properties.

E. coli expression of p60 was unsuccessful so constructs were cloned into plasmids for insect cell expression.The MultiBac system was utilised to create several p60 constructs, two of which were successfully expressed in insect cells. The reasons for the problems with pUCDM production and digestion are still not fully understood as when the work was repeated on a visit to the Berger lab using the same methods, cell lines and enzymes procured from the same company this resulted in much more efficient cloning. Despite these problems, cloning was eventually successful to produce Myc-tagged versions of full length p60 and a shorter construct lacking the PEST motif.

Western-blot analysis of the p60 constructs using an antibody to the Myc-tag showed that the full length protein is expressed in several different forms, where the shorter p60_WD only has one. When anti-p60 is used to visualise full length p60 only one band is seen on the Western blot, this is likely because the antibody does not recognise the modified versions of the protein. This theory is supported by the knowledge that the antigen is in the C-terminal tail, known to be an unstructured phosphorylation target (Smith and Stillman, 1991).

The small p60 construct p60_WD is produced as one form, which should be more amenable to purification. However at current levels of expression and without a cost-effective method for affinity purification this was not feasible. The most efficient method for further purification is likely to be to change the tag to one that allows affinity purification, such as a his-tag and optimisation of the virus to increase the yield. A high yield would be vital for structural methods such as crystallisation as relatively large amounts of protein are required.

The reasons for lack of expression of the other three p60 constructs are unclear. It is likely that the construct containing the PEST region was unstable (construct 5) and degraded. Constructs 2 and 3 contained the more stable WD region but were not expressed. This could also be due to stability issues or the virus not retaining the DNA. However the latter is less likely as several attempts were made to express the proteins using different viruses. This suggests the protein is simply not stable.

Chapter 5

Interactions Within the CAF-1 Complex

5.1 Introduction

The interactions within the CAF-1 complex are not well characterised biochemically. The interaction of the p150 subunit with p60 has previously been demonstrated (Smith and Stillman, 1991) and the region of interaction narrowed down (Kaufman *et al.*, 1995). The presence of p48 in the complex has also been reported (Verreault *et al.*, 1996), however neither the precise regions of the proteins involved nor the binding mode of the interactions between p150 and p60 or p48 have been conclusively identified (Takami *et al.*, 2007). The interaction between p48 and p60 has never been investigated.

The stoichiometry of the CAF-1 complex is also unknown. It is known that when phosphorylated, monomeric p150 binds PCNA and that dephosphorylated p150, which exists in a dimer/monomer equilibrium has chromatin assembly activity (Gerard *et al.*, 2006). The ratio of p60 and p48 within the CAF-1 complex is however unknown. This section describes investigation of the interactions between the subunits of CAF-1 using previously purified recombinant protein components, as described in earlier chapters.

5.2 p150 Dimerisation

5.2.1 p150ΔN

As p150 is known to dimerise (Quivy *et al.*, 2001) we assessed whether the constructs described here have this ability. As part of the purification trials gel filtration was performed on p150ΔN. This gave very low concentration protein peaks as it was not possible to concentrate p150ΔN (see section 4.2) and the elution profiles were not consistent between protein preparations. Generally, a high molecular weight peak containing aggregated protein was seen with a wide single peak containing p150ΔN elution at around 12 ml on a Sephadex S200 10/300 column (e.g. Figure 27, section 4.2 on page 95). Using a standard curve for the column (see section 2.2.3.3) an approximate molecular mass of 200 kDa was estimated for p150ΔN. This is likely to be an over estimate as CAF-1 has previously been seen to run on a gel filtration column at a much higher mass than expected (Smith and Stillman, 1989). Due to both the apparent mass of p150ΔN and the inconsistency between runs, which could be due to different amounts of monomer and dimer at the low protein concentration it is not possible to infer anything conclusive from this data.

Crosslinking experiments were trialled to investigate if a dimer of p150ΔN could be trapped. Purified p150ΔN after elution from Ni-NTA beads was crosslinked with BS^3, a specific amide crosslinker with a spacer of 11 Å, or formaldehyde, a non-specific amide-nitrogen crosslinker with minimal spacer. The crosslinked products were analysed using SDS-PAGE and Western blotting. This revealed the same crosslinking pattern for both crosslinkers. The results for BS^3 crosslinking are shown in Figure 40. There are three main bands indicated on the Figure of approximate sizes 70, 120 and >200 kDa. The smallest band is monomeric p150ΔN which is 75 kDa, as estimated from the raw sequence and which runs around 75 kDa when analysed by SDS-PAGE. The second band is consistent with dimeric protein (150 kDa). The third band is either an aggregate or a higher order oligomer, possibly a tetramer. As the higher band fades at lower concentrations of crosslinker it may be a result of non-specific crosslinking between existing dimers. This result suggests that p150ΔN dimerises as expected when produced in insect cells. However the proportion of the protein crosslinking is very

low (in comparison to proteins known to form dimers, e.g. histones H3 and H4), which may be explained by the low protein concentrations used in these experiments.

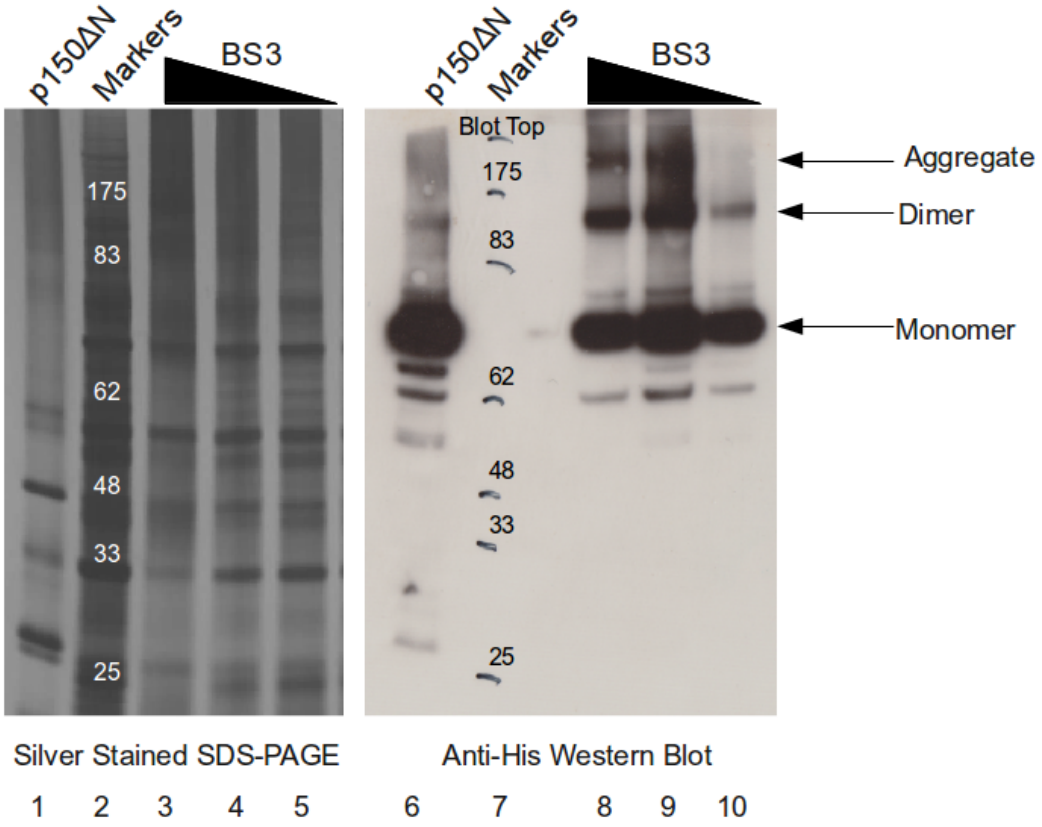


Figure 40: **Crosslinking of p150ΔN with BS^3.**
After elution from a Ni-NTA column (Lanes 1 and 6) p150ΔN was treated with at 0.15 (lanes 3 and 8), 0.06 (Lanes 4 and 9) and 0.03% (Lanes 5 and 10) (w/v) BS^3 crosslinker for 30 min. Samples were run using SDS-PAGE, transferred to a PVDF membrane and a Western blot was used to visualise His-tagged p150ΔN. Suspected stoichiometry is indicated for each band. Markers (Lanes 2 and 7) are in kDa, due to the SDS-PAGE being silver stained the markers are over-exposed.

5.2.2 p150 Construct C

As a representative of *E. coli* produced p150, construct C (residues 545-956) was examined. Gel filtration again did not give consistent results as peaks were at a similar elution volume to p150ΔN although p150_C is only 49_kDa in size. Crosslinking was repeated as for p150ΔN and gave a similar pattern with the monomer running around 55 kDa, suspected dimer at 83 kDa and larger aggregates. Less higher-order crosslinking is seen compared

to p150ΔN with high molecular weight species only seen at the highest concentration of crosslinker (compare Figure 40 lanes 8-10 with Figure 41 lanes 8-10). Some dimer can be seen in the un-crosslinked sample of p150_C, which is likely to be due to the formation of disulphide bridges between cysteine residues if the reducing agent was insufficient.

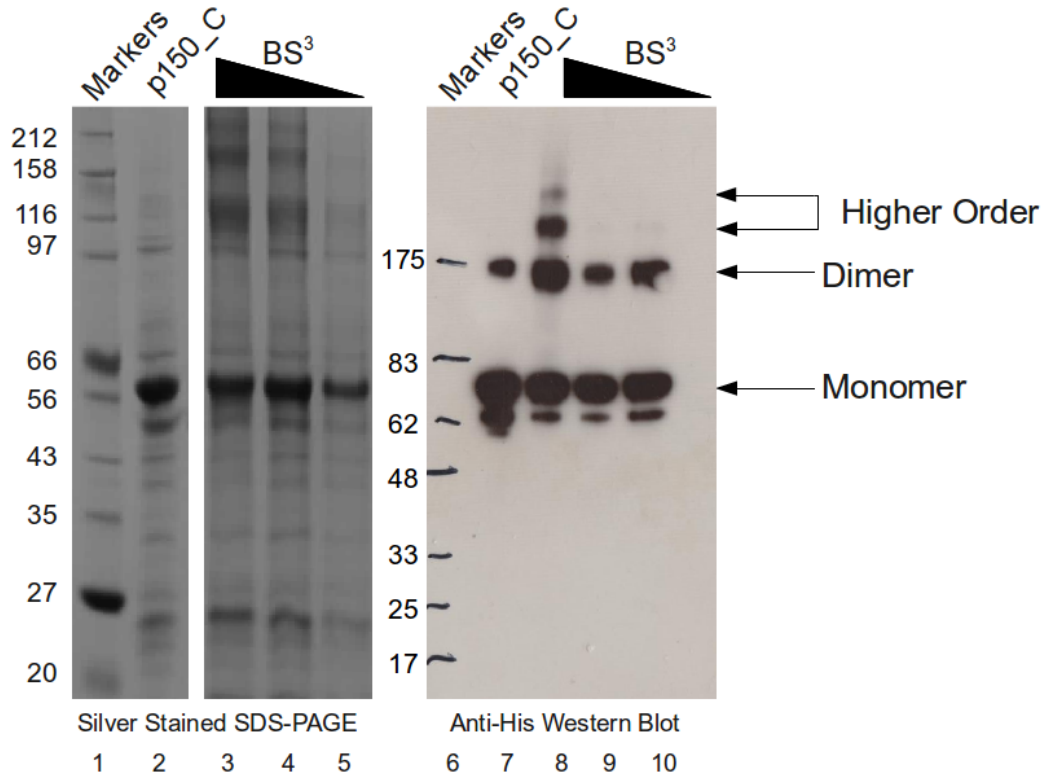


Figure 41: **Crosslinking of p150 Construct C with BS^3.**
After elution from a Ni-NTA column p150_C (lanes 2 and 7) was treated with 0.15 (lanes 3 and 8), 0.06 (lanes 4 and 9) and 0.03% (lanes 5 and 10) (w/v) BS^3 crosslinker for 30 min. Samples were separated by SDS-PAGE, transferred to a PVDF membrane and a Western blot was used to visualise His-tagged p150ΔN. Suspected stoichiometry is indicated for each band. Markers (lanes 1 and 6) are in kDa.

5.3 Interaction of p150 with p60

Full length p60 binds to p150ΔN as it is possible to co-purify untagged p60 by affinity purification of His-tagged p150ΔN from clarified lysate (section 3.3). To assess the interaction of the different p150 constructs with full length p60, pull-downs were performed between

the insect cell expressed p60 bound to anti-Myc beads and all expressed p150 constructs (Figure 42). Myc-p60 bound to beads was combined with different His-tagged truncated p150 constructs in solution and incubated at 4^{o}C for 2 h. The beads were washed to remove unbound protein and samples visualised by Western blotting. To control for different expression methods, full length p150 expressed in *E.coli* was also purified using the same methods as for the other constructs. Full length p150 is C-terminally His-tagged, which is not visualised by the anti-His antibody used, therefore anti-p150, which recognises an N-terminal epitope not present in p150ΔN was used (Smith and Stillman, 1991).

P60 only binds to protein that was expressed in insect cells (Figure 42, lanes 10 and 11). This indicates that p60 binds to something that is present when the protein is translated in eukaryotic cells, likely either a folded structural motif or a post-translational modification.

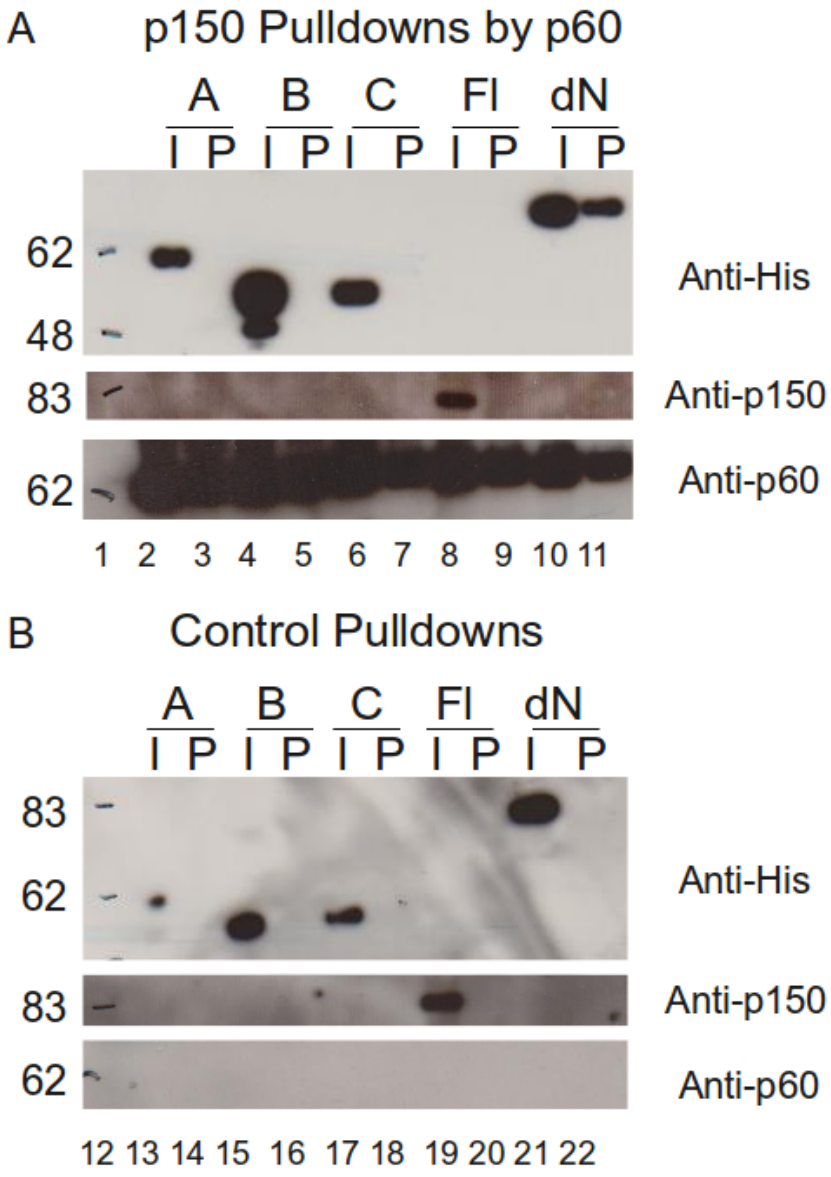


Figure 42: **Pull-down of His-p150 Constructs by Myc-p60_FL.**
A: p150 proteins were added to p60_FL bound to anti-Myc beads, incubated at 4°C, washed and visualised by Western blotting with antibodies as indicated on the right. Constructs of p150 A-C, produced in *E.coli* and p150ΔN (dN) purified from insect cells were visualised using anti-His antibody. Full length p150 (Fl) was visualised by an anti-p150 antibody. B: Control pull-downs were done using Sepharose 4B beads. I = Input, P= Pull-down. Molecular markers (lanes 1 and 12) are in kDa.

To determine whether p60 binds to a structure requiring eukaryotic chaperones to fold or a post-translational modification, p150ΔN expressed in insect cells was denatured. The protein was purified from clarified insect cell lysate using Ni-NTA beads in 8 M urea to

remove all structure and then washed back into native buffer (section 2.2.3). This protein was then either left on beads or eluted and pull-downs performed using p60 bound to beads. Native p150ΔN was used as a positive control (Figure 43, compare lane 2 with 4 and lane 6 with 8). This showed that p60_FL does not bind to denatured p150, indicating that it interacts with a structured region in p150 which requires eukaryotic expression systems to form.

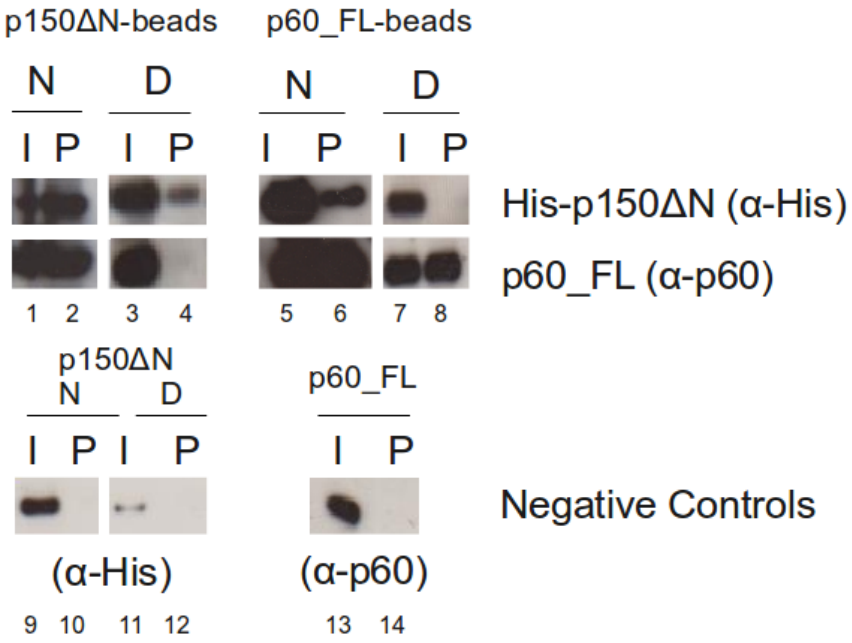


Figure 43: **Native and Denatured p150ΔN Pull-downs of p60_FL.**
Either His-p150ΔN, purified in a native state (N) and denatured in urea (D) or Myc-p60_FL were bound to beads and used to pull-down the other soluble protein. Western blotting was used to visualise the resulting bound protein with antibodies as indicated. Control pull-downs used clean Ni-NTA beads (binding to p150) and Sepharose 4B (binding to p60). I = Input, P= Pull-down.

To assess whether the binding of p60 to p150 is mediated by its WD region or through its unstructured tail, pull-downs were performed using p60_WD. The epitope recognised by anti-p60 is on the tail which is missing from this shorter construct, therefore anti-Myc was used to visualise p60_WD. Fortuitously, where p60_WD was eluted from beads by thrombin cleavage a small percentage of the protein appears be eluted without cleavage of the tag, allowing visualisation by anti-Myc. As shown in Figure 44, p60_WD behaves in the same way as p60_FL, suggesting that the WD region of p60 binds to a structural motif on p150.

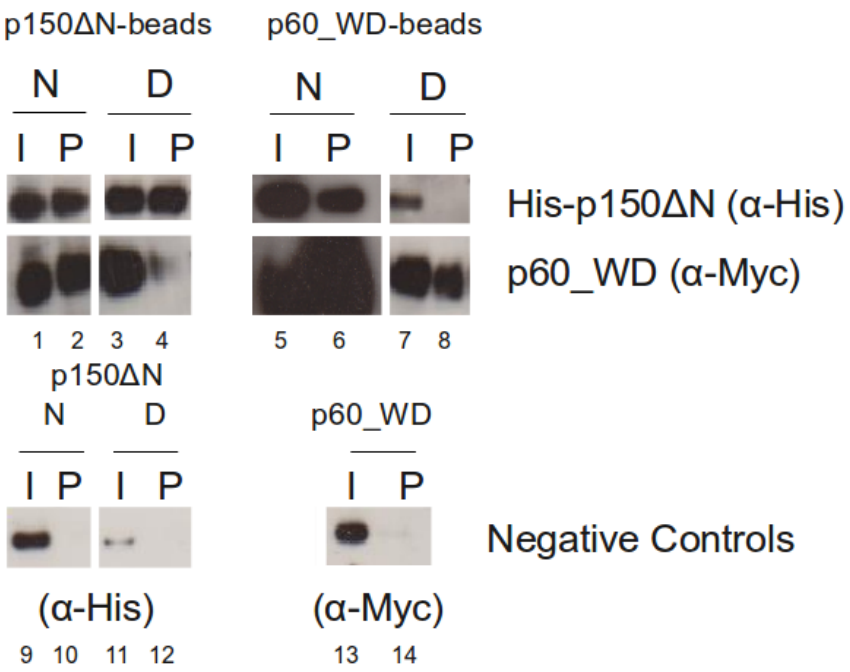


Figure 44: **Native and Denatured p150ΔN Pull-downs of p60_WD.**
Either native p150ΔN (N), p150ΔN denatured in urea (D) or Myc-p60_WD were bound to beads and used to pull-down the other soluble protein. Western blotting was used to visualise the bound protein with antibodies as indicated. Control pull-downs used clean Ni-NTA beads (binding to p150) and Sepharose 4B (binding to p60). I = Input, P= Pull-down.

5.4 Interaction of p150 with p48

To examine the interaction between p48 and p150 the *E. coli* expressed p150 constructs were used for pull-downs experiments. As both the p150 constructs and p48 are purified using the same tag we needed to cleave the His-tag off from one protein prior to the pull-downs. P48 was chosen to be in solution as its tag is cleaved off during the standard purification protocol (section 1.5.2.3) and the untagged protein can be visualised using anti-p48 antibodies, whereas attempts to cleave the tag from p150 constructs resulted in protein degradation (section 4.2).

Initially, p48 was seen to bind to Ni-NTA in the absence of any protein. Investigation of the conditions revealed that the presence of 10 mM imidazole and absence of glycerol from the reaction buffer prevents most p48 binding to the beads (His Pull-down Buffer, section

B). The pull-down is shown in Figure 45 and whilst a small amount of binding of p48 to beads can be observed in the negative control (Fig 45, lanes 13 and 14), it binds much more strongly to p150 (compare Figure 45, lanes 8, 10 and 14). The binding of p48 appears to be to a sequence motif in p150 as it binds to all the constructs including the one purified in urea (lanes 11 and 12). The results for anti-p48 on this blot are slightly over-exposed due to the antibody used having to be re-titrated after accidental desiccation, there is also evidence of a bubble having prevented optimal transfer of the p48 in lane 5.

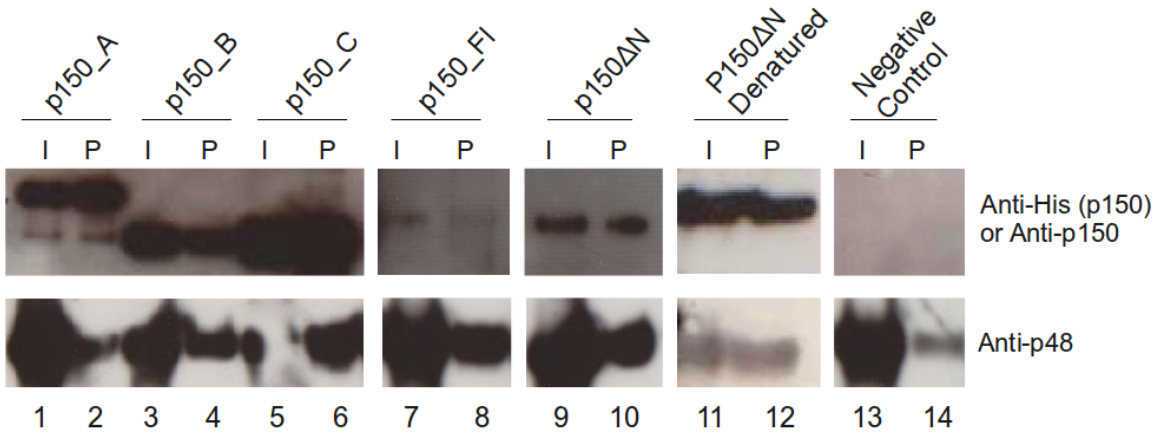


Figure 45: **Pull-down of p48 by p150 Constructs**.
Pull-downs were performed with p150 constructs bound to beads and tagless p48 in solution. Western blotting was used to visualise the result with antibodies as indicated. Constructs of p150 are denoted A-C, full length p150 expressed in *E.coli* is denoted Fl, p150ΔN expressed in insect cells is p150ΔN. p150ΔN Denatured was purified in urea from insect cells. Negative control (-) was clean Ni-NTA beads. I = Input, P= Pull-down.

5.4.1 Interaction of p48 with p150_C

All the constructs of p150 were seen to bind to p48. The smallest of these is construct C, which is still a relatively large protein at 49 kDa. For further investigation more protein was required so a large scale purification protocol was developed for p150_C. The concentration of KCl used in the purification was increased to 400 mM, the same as for purification of insect cell expressed p150ΔN as it increased protein yield. A denaturing purification was used (section 2.2.3.3) as p48 binds p150 purified in this manner and the purification removed

the majority of contaminating proteins, which appeared to decrease the affinity of p48 for p150_C. A typical purification is shown in Figure 46.

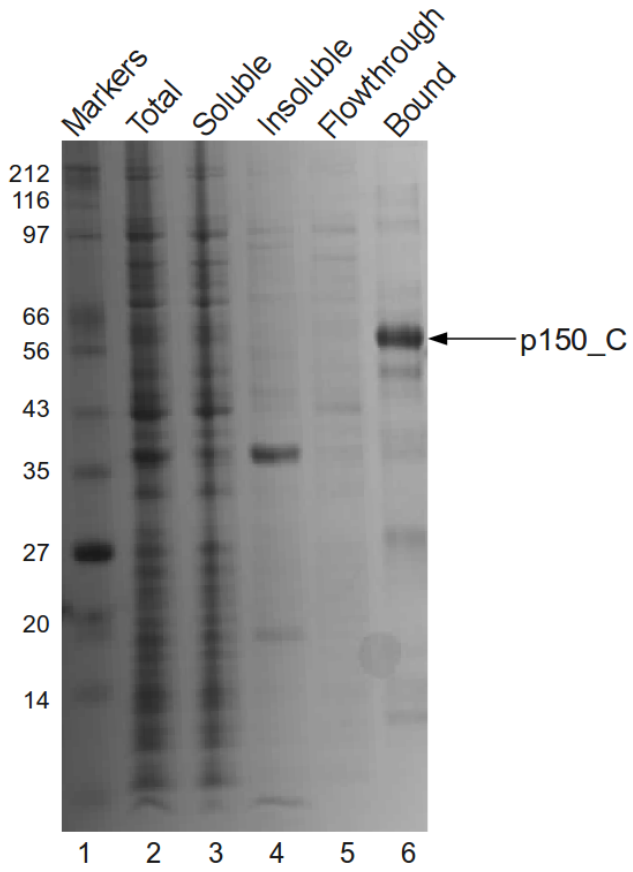


Figure 46: **SDS-PAGE Analysis of a Large Scale Purification of p150_C.**
p150_C was purified by adding urea to clarified lysate, binding to Ni-NTA beads and washing in decreasing concentrations of urea back to native conditions as described in section 2.2.3.

To locate the fragment of p150_C responsible for binding to p48 limited proteolysis was investigated. Excess p48 was added to p150_C bound to 0.5 ml Ni-NTA beads and washed with His Pull-down Buffer (section B) to remove surplus p48. The resulting complex eluted from the beads appears to be 1:1 p48 : p150_C by SDS-PAGE analysis (Figure 47, lanes 7-11).

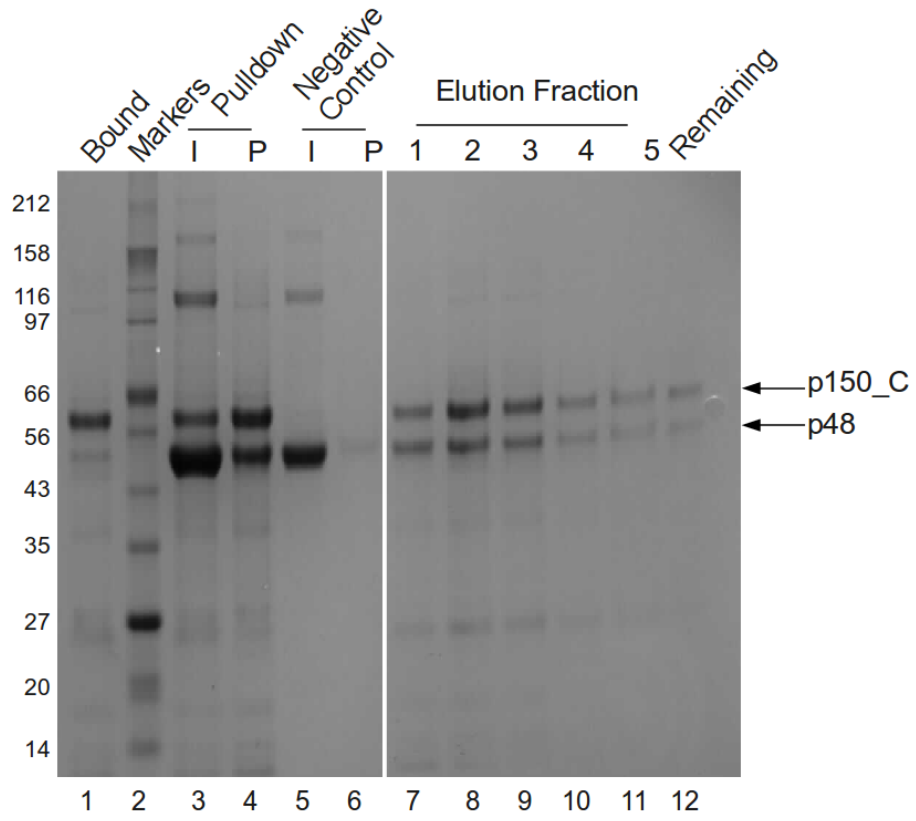


Figure 47: **Large Scale Preparation of the p150_C/p48 Complex.**
p150_C was purified under denaturing conditions bound to Ni-NTA beads (Bound). Purified p48 was then added to the beads and excess protein washed off. I= Input, P =pull-down, Remaining = proteins remaining on beads after elution. A control pull-down between p48 and Ni-NTA beads was also done (Negative Control). The protein was then eluted in 0.5 ml binding buffer containing 500 mM Imidazole. Markers are in kDa.

Limited proteolysis was attempted with a series of dilutions of subtilisin to find a concentration where p48 remained intact whilst digesting unstructured and non-protected parts of p150_C. Unfortunately, p48 was seen to be digested at similar rates to p150_C (data not shown). At the final subtilisin concentrations of 0.6 ng/µl neither p48 nor p150_C were digested whereas at higher concentrations both proteins were affected. It was not possible to use this technique to define a p150 interaction fragment. This is likely due to the structure of p48, a 7 blade beta propeller where the linker regions between the blades are readily accessible to protease and as such digest quickly.

5.4.1.1 Bioinformatic Investigations

Comparison of the binding region for p48 examined here (construct C, residues 545-956) with that previously published by Takami *et al.* (2007) (residues 442-642 when normalised to full length p150) gives a region of overlap of 97 amino acids. Manual examination of this region comparing it to other known p48 binding motifs revealed a RRK (arginine, arginine, lysine) sequence motif at residues 551-553, which was notable as there have recently been shown to be the critical residues in FOG1 required for interaction with p48 (Lejon *et al.*, 2011; Hong *et al.*, 2005).

	Sequence	**Region**
FOG1	MSRRKQXXP	1-9
p150	PERRKFGRM	549-557
Alignment	hXRRKXXXh	

Table 6: **Manual Alignment of the FOG1 Consensus Interaction Motif with a Putative Motif in p150.** h = hydrophobic residue, X = any amino acid

5.4.1.2 Competition Experiments

To assess if this motif was responsible for the binding of p150 to p48 a competition experiment was performed. The FOG1 peptide (residues 1-15, Lejon *et al.*, 2011) was mixed to ten times excess with p48 and used in a pull-down experiment with p150_C bound to Ni-NTA beads. As a control an identical amount of p48 diluted with buffer was also used. The results are shown in Figure 48. The addition of FOG1 clearly reduces the amount of p48 that binds to p150.

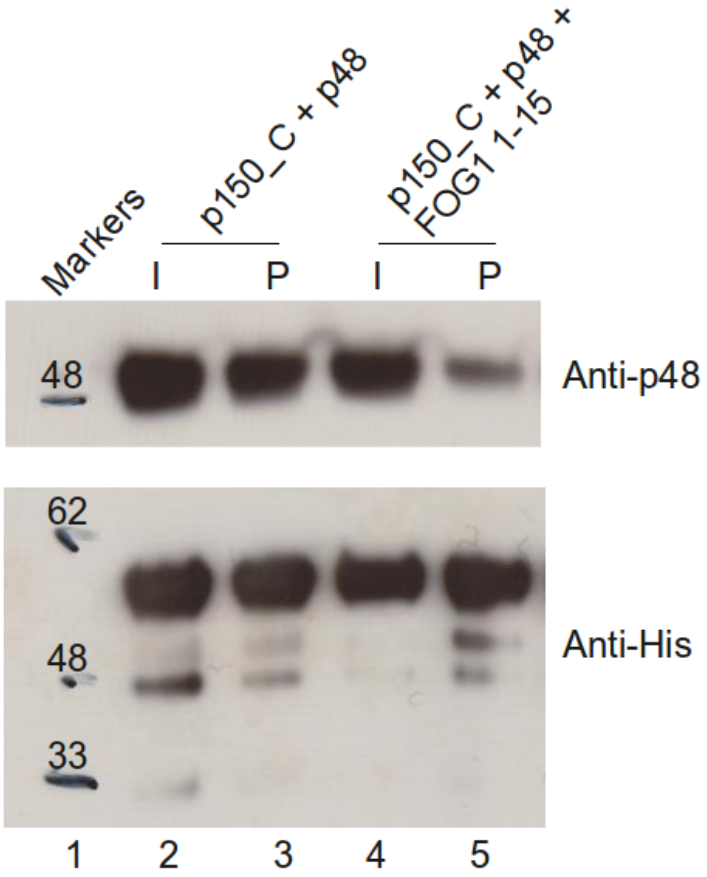


Figure 48: **Competition of FOG1 1-15 Peptide with p150 for Binding to p48.**
A pull-down was performed using Ni-NTA beads bound to p150_C with either p48 alone or p48 pre-mixed with 5 x excess of FOG1 peptide. The resulting proteins were analysed by Western blot using anti-p48 and anti-His antibodies to visualise p48 and p150_C respectively. Markers are in kDa.

5.4.1.3 Mutagenesis of the RRK Motif in p150_C

The essential residues for binding of FOG1 to p48 are the second arginine and the lysine of the RRK motif, mutation of either to a glycine or alanine respectively, completely removes binding to p48 (Hong *et al.*, 2005). Site directed mutagenesis of p150_C was undertaken to mutate these residues. Primers were designed to mutate R552A and K553G and a double mutation of R552GK553A (Appendix E) and the plasmid mutated by PCR using the Quik Change Mutagenesis Kit II (Agilent). Correct plasmids were transformed into Rosetta 2 pLysS cells and expressed as wild type p150_C.

5.4.1.4 Pull-downs of Mutated p150_C with p48

Protein was prepared bound to Ni-NTA beads as for wild type p150_C (see section 2.2.3.3.4). Purified p48 was added to each mutant and wild type p150_C (positive control) and a pull-down performed as in section 5.4. The result is shown in Figure 49. This reveals that mutation of arginine 552 to alanine has a negligible effect on the binding of p48 to p150_C whereas mutating lysine 553 to glycine diminishes the binding. However, surprisingly, the double mutant R552GK553A binds with an intensity similar to that of the full length protein.

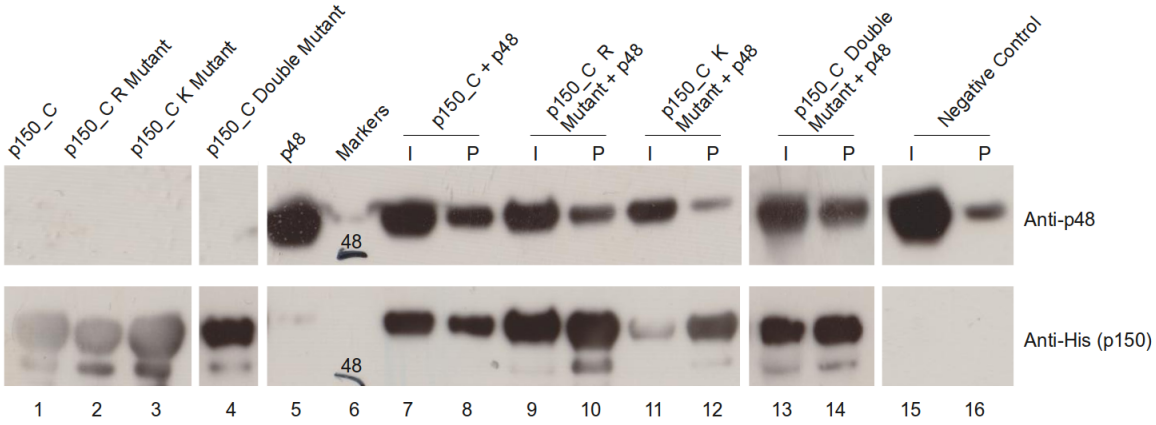


Figure 49: **Pull-downs of p48 by Mutant p150_C Proteins.**
Pull-downs were performed with His-p150_C constructs bound to beads and tagless p48 in solution. The results were analysed by Western blot using anti-p48 and anti-His antibodies to visualise p48 and p150_C, respectively. The negative control was clean Ni-NTA beads. I = Input, P= Pull-down.

5.5 Interaction of p60 with p48

To assess if p60 interacts with p48 in the absence of p150, a pull-down was performed as described in section 2.2.7 using p48 bound to Ni-NTA beads to bind p60 in solution. His-p48 beads were incubated with p60_FL at 4^{o}C then washed to remove unbound protein. No interaction is observed between p48 and p60 (Figure 50). It was only possible to perform this experiment in one direction as p48 runs on SDS-PAGE at the same size as rabbit IgG heavy chain, which is used to bind p60 to anti-Myc beads. This is also visualised by the same secondary antibody as used for p48, making it impossible to tell the two proteins apart.

Also it was not possible to show that the presence of p150 in this pull-down would cause p60 to bind His-p48 as it is not possible to remove the His-tag from p150.

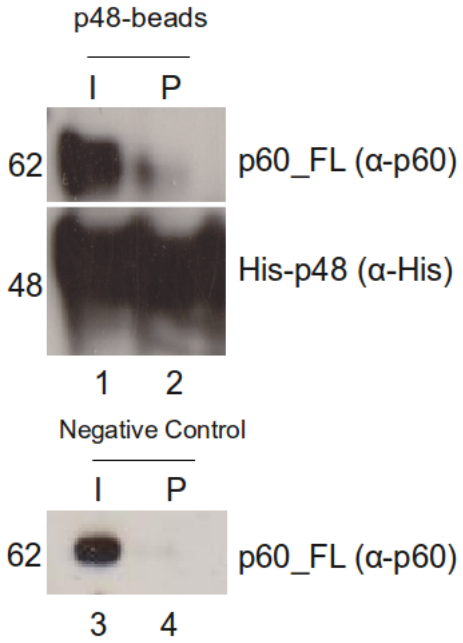


Figure 50: **Pull-downs of p60 by p48.**
Ni-NTA beads bound to p48 were mixed with full length p60 and a pull-down performed. The results were analysed by Western blot using anti-p60 and anti-His antibodies to visualise p60 and His-p48, respectively. Top: Pull-down of p60 in solution by His-p48 bound to beads. Bottom: Negative control of binding p60_FL to Ni-NTA beads. I = Input, P = Pull-down.

The lack of a detectable interaction between p60 and p48 is also seen during the purification of CAF-1. When the CAF-1 complex is purified by use of a tag on p48 (Figure 18, section 3.2.3, compare lane 2 to 6 and 7), p150ΔN is reliably pulled out of solution but p60 is lost.

5.6 Stoichiometry of CAF-1 complexes

To investigate the stoichiometry of the CAF-1 complex crosslinking was performed. Samples were crosslinked using both BS^3 and formaldehyde which were seen to give the same results. The results from BS^3crosslinking are shown in the figures as samples could be boiled without reducing the crosslinker, making SDS-PAGE and Western blot analysis more repeatable.

5.6.1 p150_C and p48

A Coomassie stained SDS-PAGE gel is shown in Figure 51. P150_C with and without p48 was crosslinked with increasing concentrations of BS^3. In both cases high molecular weight crosslinked species were observed, along with the original monomeric proteins as there was insufficient crosslinker to drive the reaction to completion. The pattern observed in the presence and absence of p48 is very similar (compare Figure 51 lane4 with lane 8). Due to the presence of contaminants in the purification of p150_C a Western blot was used to more accurately analyse the results.

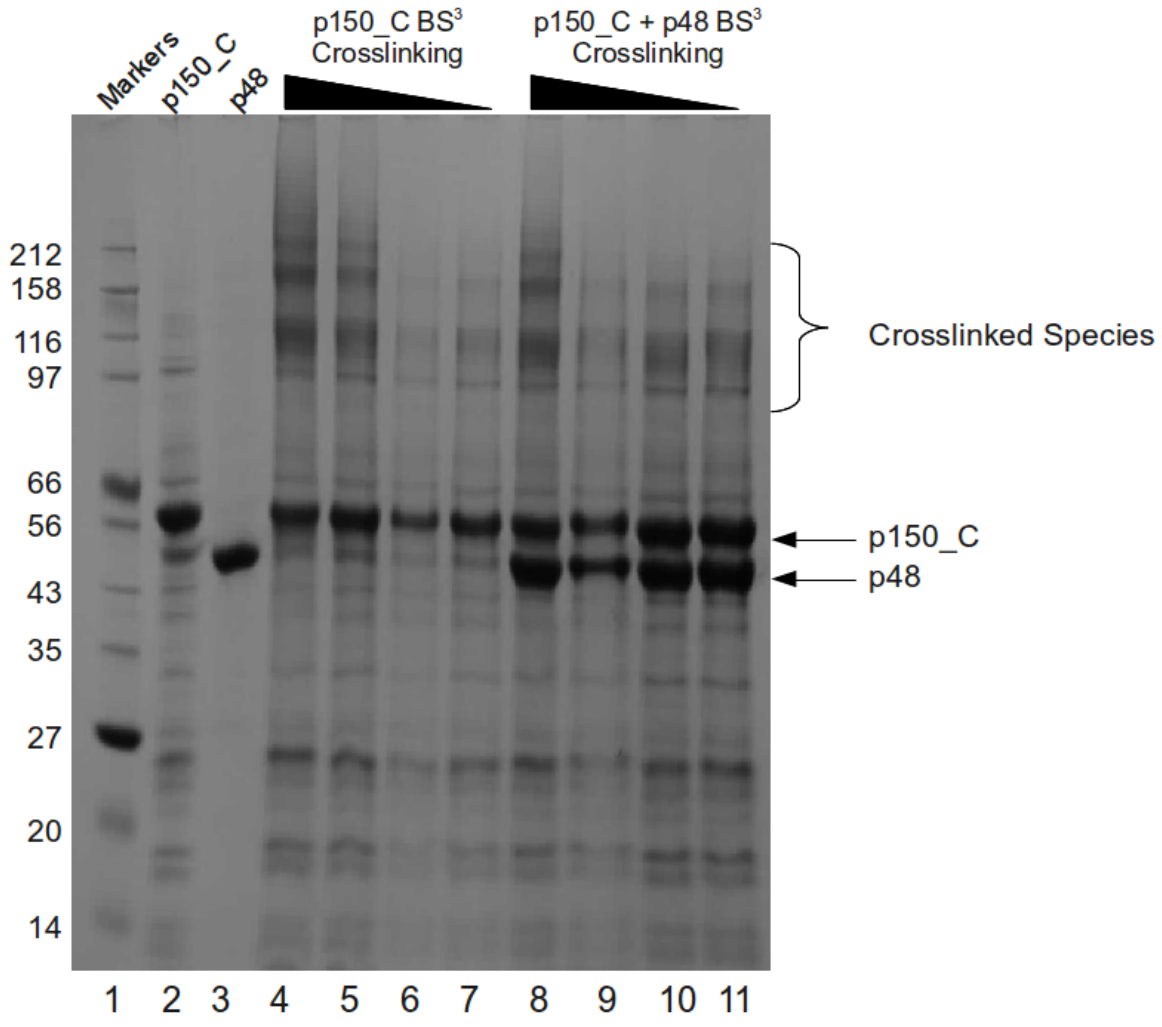


Figure 51: **Crosslinking of p150_C with p48 using BS^3.**
P150 in the presence or absence of p48 was incubated at 27°C for 30 min with 0.6, 0.15, 0.06 and 0.03% (w/v) crosslinker. The reaction was stopped by addition of Tris-HCl pH 8 to 0.1 M. Samples were analysed by SDS-PAGE. Markers are in kDa.

Western blots were performed on the samples containing the highest concentration of cross-linker, as initial experiments showed that the maximum amount of complex was trapped at this concentration (Figure 52). Lane 20 shows the overlay of the p150_C and p48 Western blots. The crosslinking of p150 to itself is as before, with some dimeric p150_C seen in the un-crosslinked protein. P48 alone does not crosslink strongly to itself. Crosslinking of p48 to p150_C produces a small overlap at around 110 kDa. This mass corresponds to either a p150_C monomer bound to a p48 monomer (predicted 97 kDa) or a p48 dimer. However crosslinking of p48 alone at the same crosslinker concentration (Figure 52, lanes 13, 18 and 26) only shows very faint crosslinking, indicating it is more likely to be a p150/p48 complex.

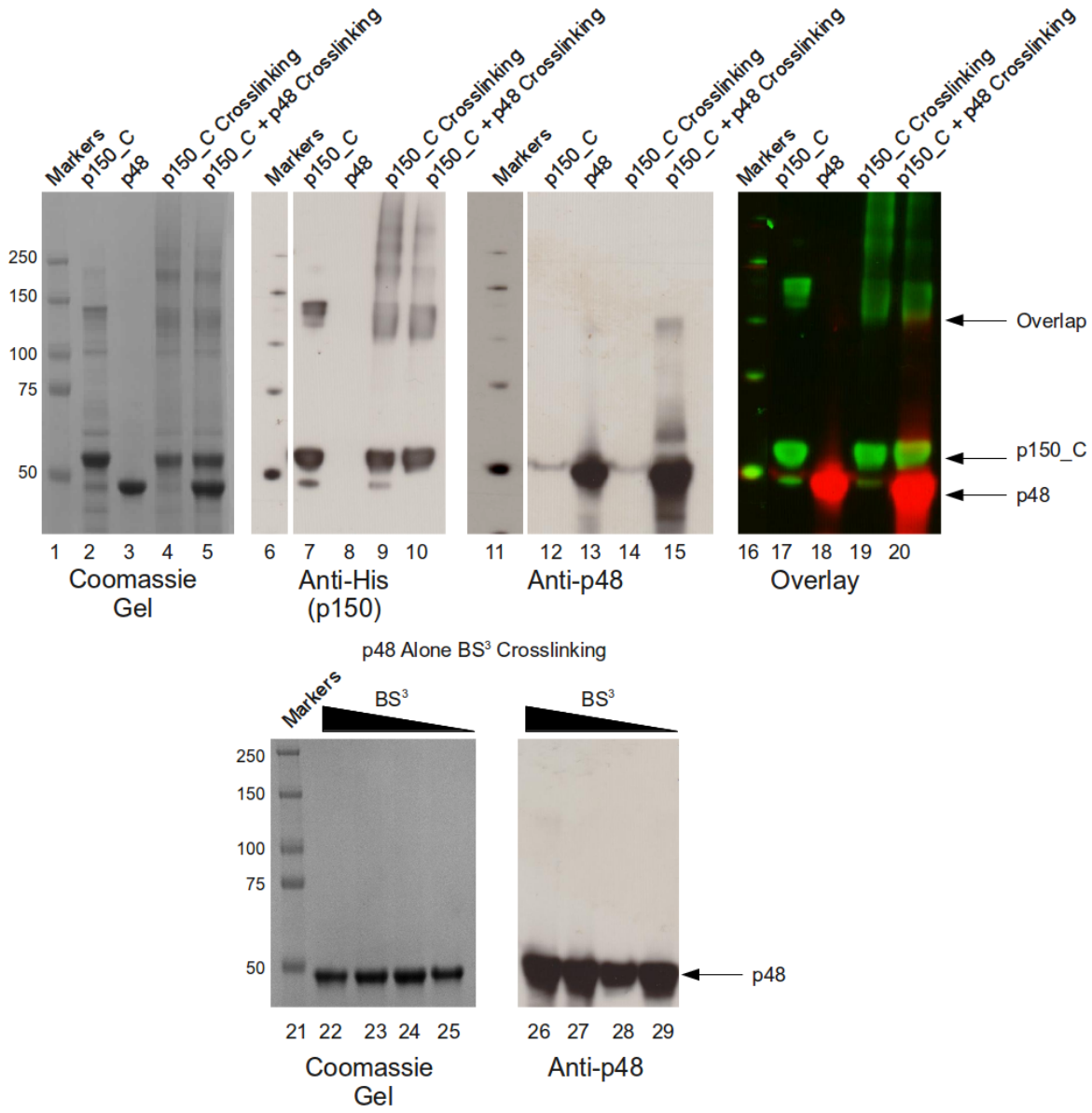


Figure 52: **Western Blot of 150_C Crosslinking with p48.**
Protein samples were crosslinked with 0.6 % (w/v) BS^3 were run on the same SDS-PAGE gel and transferred to a PVDF membrane. A Western blot was performed using the antibodies indicated at the bottom of the diagram. Western blots were aligned using markers and overlayed using GIMP and ImageJ for image processing, p150_C is green and p48 red. Markers are in kDa. Markers for the Western blots were taken from the blot at different exposure lengths to the protein samples and have been aligned based on overlays.

5.6.2 p150ΔN/p60 with and without p48

Insect cell p150ΔN/p60 was crosslinked alone, with the addition of p48 and with excess p48 to examine if p60 was displaced. Samples were crosslinked with increasing concentrations of BS^3 and analysed by Coomassie stained SDS-PAGE gel (Figure 53). p150ΔN/p60 appears to crosslink more efficiently than p150_C (compare Figure 53 lane 4 to Figure 51 lane 4). However, addition of p48 did not influence the results in a major way, with the majority of the protein remaining uncrosslinked, even at higher concentrations of crosslinker. Western blots were attempted to analyse the data (not shown), but as most of the protein either remained monomeric or was seen to be aggregated the results could not be readily interpreted.

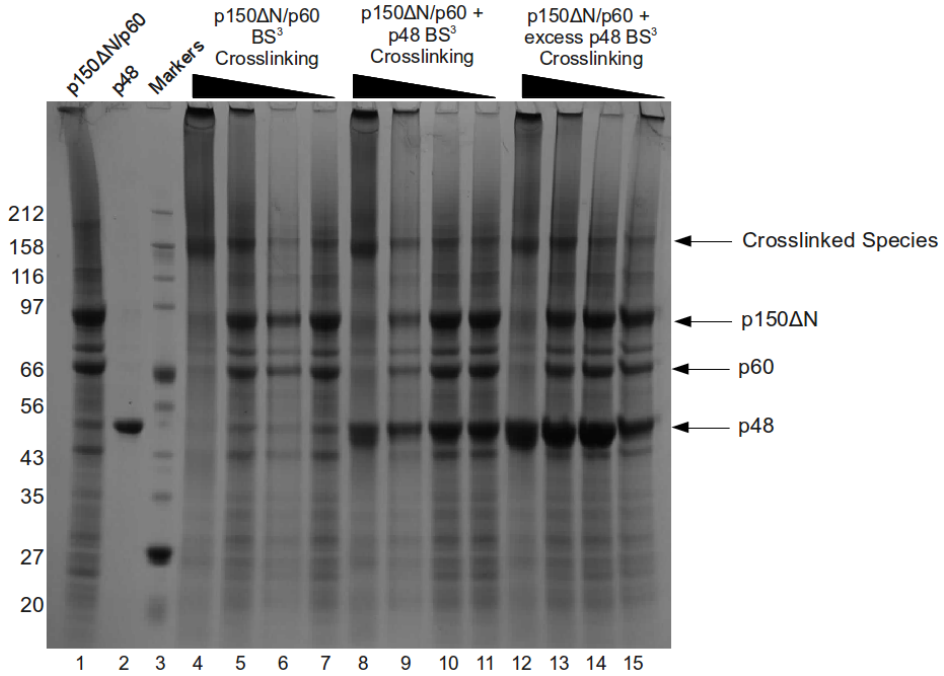


Figure 53: **SDS-PAGE of p150ΔN/p60 Crosslinking with and without p48.**
Ni-NTA purified p150ΔN/p60 in the presence and absence of p48 (at 1:1 ratio and in 3 fold excess) was treated with 0.6, 0.15, 0.06 and 0.03% (w/v) BS^3 crosslinker for 30 min at 27°C. The reaction was stopped by addition of Tris-HCl pH 8 to 0.1 M. Samples were analysed by Coomassie stained SDS-PAGE. Markers are in kDa.

5.7 Conclusions

The results in this chapter examine the interactions between subunits and the stoichiometry of the CAF-1 complex. Firstly the interactions between subunits were examined by pull-down experiments and the binding mode assessed using denatured protein, competition and mutagenesis experiments. Secondly, crosslinking reactions were used to look for high molecular weight complexes.

p150 has been shown to exist as a homodimer by crosslinking of both p150ΔN produced in insect cells and p150_C purified from *E. coli*, supporting the data from glycerol sedimentation (Quivy *et al.*, 2001). The preservation of the interaction in p150_C is to be noted as the purification was performed using 8 M urea, indicating that the interaction motif is likely to have refolded on return to native conditions. There is less high molecular weight crosslinking of p150_C compared to p150ΔN at the same concentrations of crosslinker, which may indicate less aggregation in the protein produced in *E. coli*. The decrease in high molecular weight aggregates may be due to the urea removing non-specifically interacting proteins from p150_C which otherwise increased protein aggregation.

Investigation of the interaction between p150 and p60 revealed that p60 binds exclusively to protein expressed in insect cells. Treatment of p150ΔN from insect cells with 8 M urea during purification also prevented the binding of p60. Treatment with urea removes structure that may require chaperones to refold effectively on return to native conditions. This indicates that p60 binds to a structural motif and that further work will require the use of eukaryotic expression systems. In particular, Takami *et al.* (2007) identify the region 624-714 as binding p60, which is predicted to have a small region of ordered structure (Figure 31, section 4.3). We intended to express this as p150 construct E, but it did not express in *E. coli* (section 4.3.4). However, this construct could be used for further work using insect cell expression.

The p60_WD construct appears to bind to p150ΔN in the same manner as full length protein, indicating that it is the WD region that facilitates the interaction between the proteins. This has not been investigated before, with the only previously characterised interaction of p60

being that to ASF1, which is mediated *via* the unstructured C-terminal tail (Tang *et al.*, 2006).

The interaction of p150 constructs with p48 was examined and showed that p48 binds to all the constructs produced in *E. coli* and insect cells, even when treated with urea. Limited proteolysis was attempted to locate a minimal binding fragment but this was unsuccessful. It may be possible to use other enzymes to improve this in future, however bioinformatic analysis of the sequence revealed a further line of enquiry.

Using information from Takami *et al.* (2007) and Lejon *et al.* (2011) a putative interaction motif was identified containing a RRK motif, through which FOG1 is known to bind p48. Competition of the FOG1 peptide with p150_C showed a decrease in the binding of p150 to p48 in the presence of the peptide. Mutagenesis of the residues thought to be key for this interaction (Hong *et al.*, 2005) showed a negligible decrease in binding to the R552G mutant but a more significant decrease when K553 was mutated to alanine. However, the double mutant appears to bind to p48 at a similar level to the control. The lack of a decrease on mutation of R552 can be explained given the new structure of H3 binding to Nurf55 (p55, the *D. melanogaster* homologue of p48) (Schmitges *et al.*, 2011). The H3 peptide binds to the same site as FOG1, but the second arginine of the RRK motif is threonine in H3, allowing the preceding arginine to bind in the cleft. The binding of the lysine residue in both peptides however is to a deep cleft which is required for the interaction. This is illustrated in Figure 54 - compare backbone position of cyan H3 to green FOG1.

The reversion on double mutation is less easy to understand. It may be that there is a second potential binding site in CAF-1 that is only employed on mutation of the first or, that the site investigated here is not responsible for binding. The fact that the decrease in binding was reliably observed in the lysine mutant makes this less likely. Further investigation of this using peptides in a similar manner to that employed in the investigation of the FOG1 interaction should provide greater insight into this problem. Either way, given the data from competition experiments it is very likely that p150 binds in the same cleft on p48 as the FOG1 peptide, indicating a more general role for this interaction site in the assembly of chromatin complexs.

There was no interaction observed between p60 and p48 in the absence of p150, using meth-

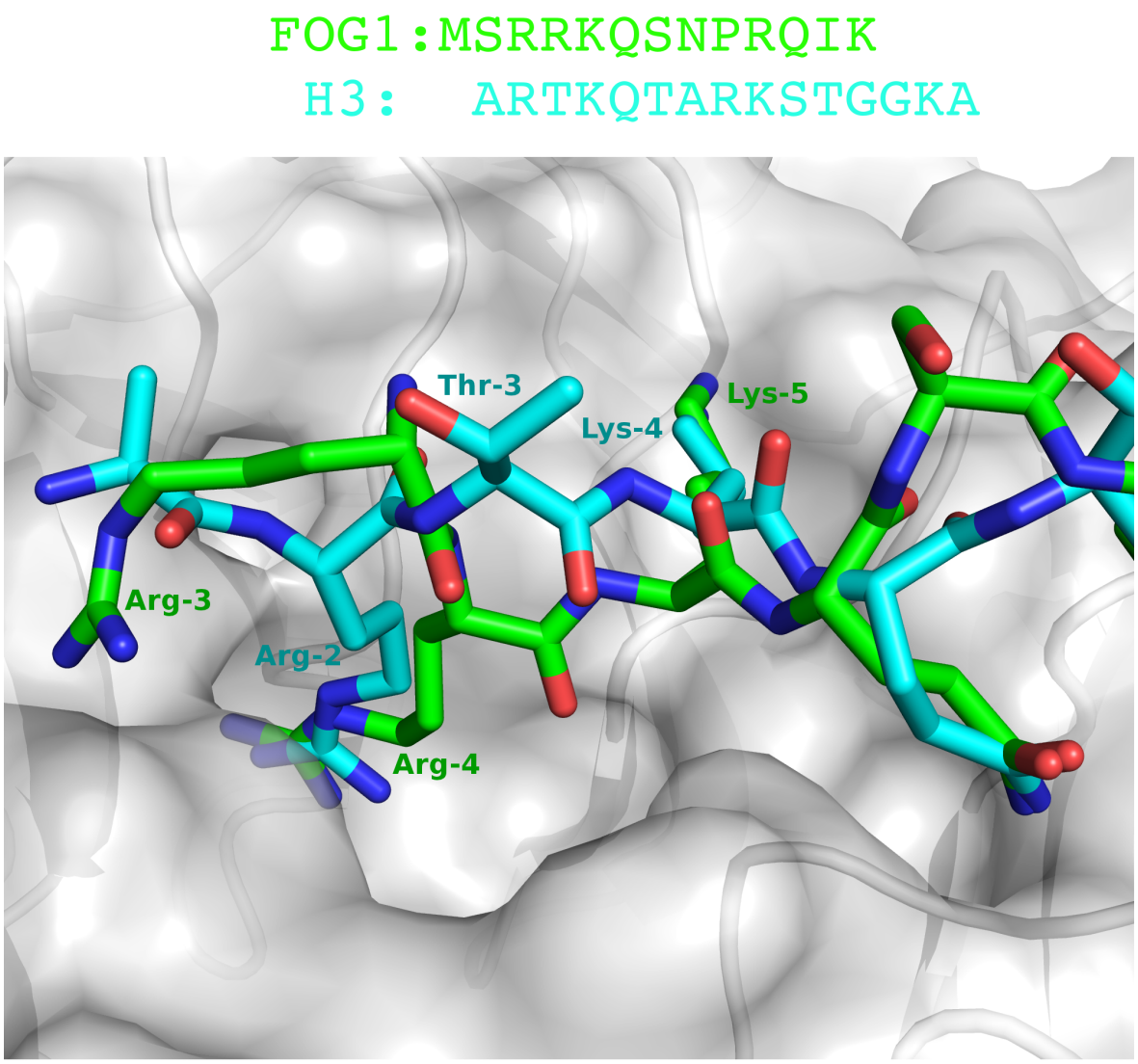


Figure 54: **Superimposition of FOG1 and H3 Peptides Binding to p48.** Superimposition of PDB ID 2XU7 (FOG1) and 2YBA (H3) Green: FOG1 peptide, visible sequence: MSRRKQS . Cyan: H3 fragment, visible sequence: ARTKQT . Figure aligned and constructed using PyMol

ods such as the pull-downs described above or indirectly observed from purifications of the CAF-1 complex. This has not been discussed in the literature and highlights the role of p150 in recruiting both proteins to sites of chromatin assembly. Whether both proteins can bind simultaneously to one molecule of p150 remains to be confirmed, but purification of the protein complex section 3.2) suggests a 1:1:1 ratio, assuming all proteins bind Coomassie dye in equal quantities. It appears unlikely that there are two pools of proteins, containing p150 bound to either p48 or p60 as p60 was seen during the p48-based calmodulin purification of the CAF-1 complex (section 3.2.3).

Based on the information above, the putative binding sites on p150 for itself, p60 and p48 are indicated in Figure 55. There is overlap between the binding site of p60 and the p150 dimerisation domain, which may suggest that the location of the p60 binding site of p60 corresponds to the p150 dimerisation domain.

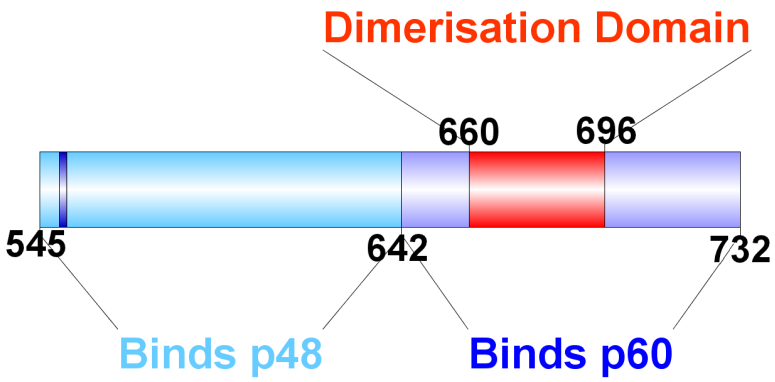


Figure 55: **Fragment of p150 That Binds p60 and p48.**
All three subunits of the CAF-1 complex bind over a 187 aa region. The putative p48 binding RRK motif is indicated in dark blue. Image constructed using information from (Takami *et al.*, 2007; Quivy *et al.*, 2001) using DOG 1.0 (Ren *et al.*, 2009).

The data from crosslinking of the CAF-1 complex was difficult to interpret due to what appears to be excessive crosslinking of p150, a protein rich in lysine residues masking other signals except at high concentrations of crosslinker. The p150 dimer is clearly seen even at the lowest concentrations of crosslinker. The aggregation of p150ΔN and heterogeneity of the p60 produced in the p150ΔN/p60 complex did not allow confident identification of

p150/60 crosslinking products. The crosslinking between p150_C and p48 produced one band of overlap at a size that could correspond to a dimer of p150, a dimer of p48 or a binary complex containing both proteins. Crosslinking the whole complex using p150ΔN/p60 and p48 produced two bands of overlap that could not be identified accurately due to the lack of coherent bands for p60. Further work with more homogeneous protein may provide better data.

Chapter 6

Summary and Perspectives

This thesis describes the production and purification of components of the CAF-1 complex and characterisation of the interactions between them. The original aims of this project were structural studies of the CAF-1 complex and its subunits (p150, p60 and p48). To achieve this it was necessary to produce recombinant CAF-1 reliably and stably at a quantity and purity suitable for biochemical and structural experiments. Whilst CAF-1 has been reliably produced at levels above what was previously possible it is not as homogeneous as is desired for structural work. This was due to a variety of factors.

Firstly, CAF-1 was cloned and produced as a complex (described in chapter 3). The trimeric complex was produced as a self-cleaving fusion protein in *Sf21* cells and purified by affinity resin. Further purification was undertaken but it was discovered that the complex could not be purified using the second affinity tag (calmodulin) and that the TEV protease, which cleaved the peptide into its constituent parts co-purified with the complex. Ion exchange revealed many impurities and low levels of expression which prevented further purification.

It was also seen that p150 was only extracted from CAF-1 expressing cells at high concentrations of KCl, indicating that it is strongly bound to chromatin. As p150 is known to be involved in heterochromatin formation as well as chromatin assembly it is possible that it is being sequestered *via* interactions with HP-1.

Expression of a binary complex containing a truncated form of p150 and p60, using the same expression system was also undertaken (described in chapter 3). The overall yield from this expression was higher than that for the whole CAF-1 complex, but ion exchange revealed that p60 was phosphorylated and eluted over a wide range. Analysis of the phosphorylation sites could be performed by mass spectrometry, although the current phosphorylation pattern may be due to over-expression in insect cells and not biologically relevant. A more homogeneous and concentrated sample that could be used for future experiments was prepared by extensive washing of the affinity resin. This sample is still not of a quality required for structural studies but provided insight into the interactions between the subunits discussed in chapter 5.

The ability to separate p150 from the other subunits of CAF-1 was observed during the purification of both complexes, indicating that p150ΔN was relatively stable when produced using the insect cell expression MultiBac system . The subunits were therefore produced individually with truncation constructs designed for p150 and p60 (described in chapter 4). Production of p150ΔN alone revealed a protein that could be purified by affinity purification but with a tendency to aggregate that made concentration of the protein impossible. The aggregation of p150 is thought to be due to the unstructured nature of the protein with exposed hydrophobic residues and highly charged patches such as the KER region tending to interact non-specifically with other proteins. Truncation constructs of p150 were produced in *E.coli* with some success, although the smaller constructs designed did not express, and the smallest was lethal on expression.

Four truncation constructs of p60 were designed for expression in insect cells, of which only one expressed, containing the WD region of the protein. Full length p60 protein was also expressed. The WD protein appeared to be more homogeneous as it did not exist in the two forms seen for full length protein. However, p60_WD could not be purified effectively in large quantities due to low expression and lack of a suitable purification tag.

The structure of p48 used in this study to examine the interactions between the subunits of CAF-1 was solved by crystallography in our laboratory (Lejon *et al.*, 2011).

As it was not possible to produce the CAF-1 fusion protein or its p150/p60 subunits at quantities and the quality required for structural studies, the project concentrated on characterising

the interaction between the subunits, discussed in chapter 5.

P150 is known to dimerise with itself (Quivy *et al.*, 2001). Crosslinking experiments confirmed this for both full length protein and construct_C, consisting of residues 545-596 of p150, the smallest construct that could be expressed and purified, as detailed in section 4.3. The minimal region for dimerisation of p150 has already been reported and the results show that the p150 constructs made here also dimerise when produced in *E. coli* and refolded after denaturation. It therefore may be possible to examine the structure of this motif by NMR as a labelled protein could be produced of a size amenable to study. The predicted structure of the p150 dimerisation domain involves a short helix of 7-12 aa followed by two more weakly predicted beta-strands (PSI-PRED, Jpred3). It is possible that the helices refold on removal of urea and facilitate the dimerisation of p150.

The interaction between p60 and p150 was seen to be mediated by the WD-region of p60 binding to a structured region of p150. The interaction is lost on treatment of p150 with urea. Further investigation of this will require p150 deletion mutants to be expressed in eukaryotic expression systems.

P48 was seen to bind to all p150 constructs, whether generated in insect cells or *E. coli*. Limited proteolysis was unsuccessful in locating minimal binding fragments but bioinformatic investigation of the interaction revealed a putative p48 interaction motif: R551 R552 K553, in p150, similar to that identified for the interaction of FOG1 with p48 (Lejon *et al.*, 2011). Competition binding experiments revealed that p150_C could be displaced from p48 by the FOG1 peptide. Also mutation of the critical K553 residue of the RRK motif greatly reduced p150 binding to p48. However, mutation of R552 and mutation of both R552 and K553 did not reduce binding. The retention of binding by the single R552 mutant is explained by the recently published p55/histone H3 structure (Schmitges *et al.*, 2011), where the binding of the H3 peptide to the same site as FOG1 is mediated by a RTK motif. In this structure the first arginine binds to the same cleft as the second arginine in the FOG1 RRK motif, with the threonine facing away from the molecule and contributing little to the binding. The reason for the binding of the double mutant to p48 is however a mystery. It is possible that mutation of the main binding residue allows the protein to utilise less favourable binding sites as there

are other RXK motifs within the sequence of p150_C.

Pull-down experiments showed no interaction between p48 and p60, indicating that they require the presence of p150 to assemble into the CAF-1 complex.

Investigation of the stoichiometry of the CAF-1 complex gave a complex picture. The cross-linking data indicate that p150 and p48 appear to bind for form a 1:1 complex but further conclusions could not be drawn. Judging from the intensity of the protein bands on SDS-PAGE analyses an equal ratio of components (1:1:1 or 2:2:2) is likely from purification of the complex, providing that all subunits bind the coomassie dye equally. To further assess the stoichiometry of the complex, experiments such as further analytical gel filtration and analytical ultra-centrifugation (AUC) may provide a more accurate insight into the complex as well as showing the state of the p150 monomer/dimer equilibrium when p60 and p48 are bound. However these methods require more homogeneous protein (particularly p60) in larger quantities than have been produced so far.

The issues encountered and overcome for the first time in this dissertation provide the basis for many possible future experiments in both the short and long term. The constructs produced show that the WD region of p60 is stable and binds p150. The addition of an affinity purification tag, such as a His-tag and subsequent optimisation of the virus produced, should provide a much more homogeneous protein that may be amenable to both crystallisation studies and further investigation of the interaction with p150. To this end, the fragments of p150, including those that did not express in *E. coli*, could also be expressed in insect cells or using other systems such as *in vitro* transcription translation and the interaction with p60 assessed. Once the minimal region of interaction of p150 for p60 is identified they could be co-expressed using the MultiBac system either in pFBDM or using both pUCDM and pFL for Cre-based recombination.

Further work investigating the interaction of p150 with p48 is ongoing. P150 peptides containing the putative p48 interaction region are currently being produced. If the peptides are shown to bind to p48 then the structure of a p48/p150 complex may be solved in a similar manner to that of p48/FOG1. This new role for the FOG1 binding site of p48 indicates that it may have a generic role for complex formation and explain how p48 interacts with so

many different complexes. This interaction is seen in the NuRD (FOG-1) and CAF-1 (p150) complexes. HAT1, which binds a histone H3/H4 dimer with p48 (Verreault *et al.*, 1998) also contains a RXK motif (I. Le Guillou, personal communication) its C-terminal tail and ERα, which binds to p48 as part of the Rb complex (Creekmore *et al.*, 2008) also contains a RRK motif, although the residues involved in either interaction have not yet been identified.

Once the interactions between the subunits of the complex have been more firmly established it would be possible to examine the proteins closely associated with CAF-1 in the context of replication dependent chromatin assembly, namely histones H3 and H4, ASF-1 and PCNA. The interaction between p150 and PCNA localises CAF-1 to the replication fork and has been well characterised (Ben-Shahar *et al.*, 2009) with work to solve the structure of the PCNA/p150 peptide ongoing within this laboratory (S. Lejon, unpublished data).

The interplay between CAF-1, histones and ASF1 is less well understood. Recent work in this laboratory shows that ASF1 can form a quaternary complex with p48 mediated *via* a histone H3/H4 dimer, although some re-arrangement of the H4 helix is required (Murzina *et al.*, 2008; English *et al.*, 2006). CAF-1 also interacts with ASF1 through the C-terminal tail of p60 (Malay *et al.*, 2008) at the opposite side to the histone binding site. A model of the putative CAF1/H3/H4/ASF1 complex is summarised in Figure 56 with the position of the FOG1 peptide indicating the putative p150 binding site. Whether all the interactions shown in the model are simultaneous or if binding of p60 and p48 to ASF1 *via* H3/H4 are mutually exclusive remains to be investigated.

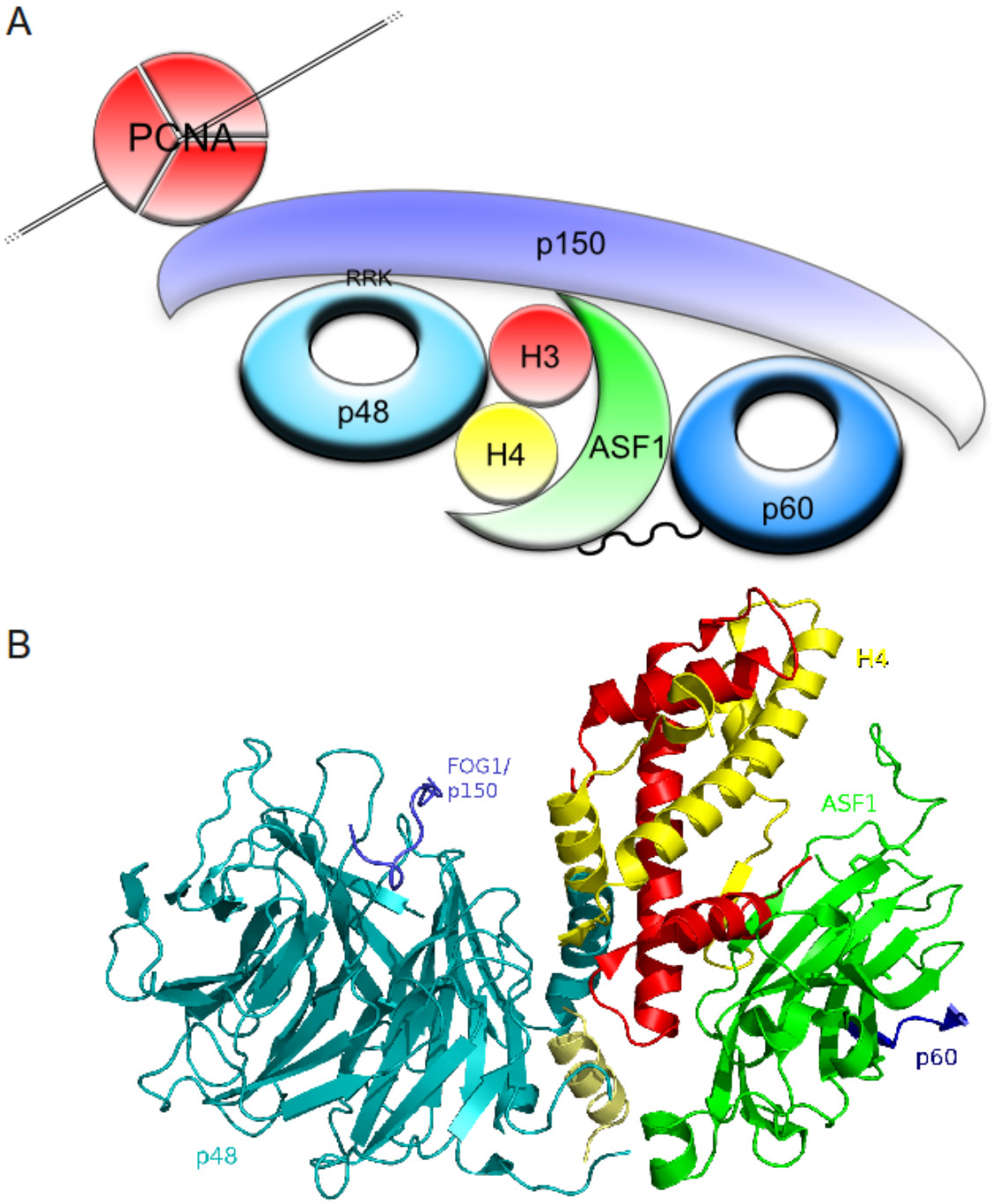


Figure 56: **Models of the Interactions Between CAF-1, H3/H4 and ASF1**
A: Cartoon illustrating interactions in the CAF-1 complex. P150 acts as a scaffold for assembly of the complex, locating the other proteins to sites of DNA replication via interaction with PCNA. B: Molecular model of the interactions. ASF1 is green, H3 red, H4 yellow, H4 peptide pale yellow, p48 aqua, p60 dark blue and FOG-1 peptide indicating p150 binding site is mid-blue. Aligned and visualised using PyMol, (p48/ASF1/H3/H4 model A Murzin, personal communication). p48/FOG1 PDBID: 2XU7, p46/H4 Fragment PDBID: 3CFS, ASF-1/H3/H4 PDBID: 2HUE and ASF-1/p60 PDBID: 2Z3F.

To study the movement of histones within the CAF-1 complex the interactions of p60 and p150 with histones should be investigated, in the presence and absence of ASF1 as this should shed light on one of the main questions of replication dependent assembly: Why do we only see a H3/H4 tetramer assembled onto DNA when nearly all known chaperones bind a H3/H4 dimer? It is possible that different subunits of the CAF-1 complex associate with separate dimers of the H3/H4 complex and deposit them onto replicated DNA. It is likely that these histones are provided by ASF1, however whether it acts as a shuttle or is directly involved in histone deposition awaits further investigation.

The mechanism by which CAF-1 deposits histones may not be unique. The tethering of p150 to the replication fork by PCNA provides replication dependency, however HIRA, which deposits H4 and H3.3 outside of replication binds to ASF1 using a similar mechanism to p60 (Tang *et al.*, 2006; Malay *et al.*, 2008). However the direct interaction with p48 exhibited by the WD motif of HIRA (Ahmad *et al.*, 2004) is not seen with p60, which may explain the lack of a p150-like scaffold for HIRA related deposition.

The work described here identifies and overcomes the major pitfalls of expression of the CAF-1 complex to provide information on the interactions within the complex. These experiments provide the groundwork for a variety of future experiments to dissect the movement and interactions that take place when histones are assembled onto newly replicated DNA.

References

Adkins M.W. and Tyler J.K. The histone chaperone Asf1p mediates global chromatin disassembly in vivo. *J Biol Chem*, 279(50):52069–52074, 2004.

Ahmad A., Takami Y. and Nakayama T. Wd dipeptide motifs and LXXLL motif of chicken HIRA are essential for interactions with the p48 subunit of chromatin assembly factor-1 and histone deacetylase-2 in vitro and in vivo. *Gene*, 342(1):125–136, 2004.

Alén C., Kent N.A., Jones H.S., O'Sullivan J., Aranda A. and Proudfoot N.J. A role for chromatin remodeling in transcriptional termination by RNA polymerase ii. *Mol Cell*, 10(6):1441–1452, 2002.

Amor-Gueret M. Bloom syndrome, genomic instability and cancer: the SOS-like hypothesis. *Cancer Lett*, 236(1):1–12, 2006.

Anderson J.D., Thaström A. and Widom J. Spontaneous access of proteins to buried nucleosomal DNA target sites occurs via a mechanism that is distinct from nucleosome translocation. *Mol Cell Biol*, 22(20):7147–7157, 2002.

Andrews A.J., Chen X., Zevin A., Stargell L.A. and Luger K. The histone chaperone Nap1 promotes nucleosome assembly by eliminating nonnucleosomal histone DNA interactions. *Mol Cell*, 37(6):834–842, 2010.

Andrews A.J., Downing G., Brown K., Park Y.J. and Luger K. A thermodynamic model for Nap1-histone interactions. *J Biol Chem*, 283(47):32412–32418, 2008.

Annunziato A.T. Split decision: what happens to nucleosomes during DNA replication? *J Biol Chem*, 280(13):12065–12068, 2005.

Arya G., Maitra A. and Grigoryev S.A. A structural perspective on the where, how, why, and what of nucleosome positioning. *J Biomol Struct Dyn*, 27(6):803–820, 2010.

Avvakumov N., Nourani A. and Côté J. Histone chaperones: modulators of chromatin marks. *Mol Cell*, 41(5):502–514, 2011.

Bannister A.J. and Kouzarides T. Reversing histone methylation. *Nature*, 436(7054):1103–1106, 2005.

Barak O., Lazzaro M.A., Lane W.S., Speicher D.W., Picketts D.J. and Shiekhattar R. Isolation of human NURF: a regulator of Engrailed gene expression. *EMBO J*, 22(22):6089–6100, 2003.

Bártová E., Krejcí J., Harnicarová A., Galiová G. and Kozubek S. Histone modifications and nuclear architecture: a review. *J Histochem Cytochem*, 56(8):711–721, 2008.

Basu R. and Zhang L.F. X chromosome inactivation: a silence that needs to be broken. *Genesis*, 49(11):821–834, 2011.

Batsché E., Yaniv M. and Muchardt C. The human SWI/SNF subunit Brm is a regulator of alternative splicing. *Nat Struct Mol Biol*, 13(1):22–29, 2006.

Beisel C. and Paro R. Silencing chromatin: comparing modes and mechanisms. *Nat Rev Genet*, 12(2):123–135, 2011.

Belotserkovskaya R., Oh S., Bondarenko V.A., Orphanides G., Studitsky V.M. and Reinberg D. FACT facilitates transcription-dependent nucleosome alteration. *Science*, 301(5636):1090–1093, 2003.

Ben-Shahar T.R., Castillo A.G., Osborne M.J., Borden K.L.B., Kornblatt J. and Verreault A. Two fundamentally distinct PCNA interaction peptides contribute to chromatin assembly factor 1 function. *Mol Cell Biol*, 29(24):6353–6365, 2009.

Berger I., Fitzgerald D.J. and Richmond T.J. Baculovirus expression system for heterologous multiprotein complexes. *Nat Biotechnol*, 22(12):1583–1587, 2004.

Bernstein B.E., Liu C.L., Humphrey E.L., Perlstein E.O. and Schreiber S.L. Global nucleosome occupancy in yeast. *Genome Biol*, 5(9):R62, 2004.

Bernstein B.E., Meissner A. and Lander E.S. The mammalian epigenome. *Cell*, 128(4):669–681, 2007.

Blossey R. and Schiessel H. The dynamics of the nucleosome: thermal effects, external forces and ATP. *FEBS J*, 278(19):3619–3632, 2011.

Bruning J.B. and Shamoo Y. Structural and thermodynamic analysis of human PCNA with peptides derived from DNA polymerase-delta p66 subunit and flap endonuclease-1. *Structure*, 12(12):2209–2219, 2004.

Burgers P.M.J. Polymerase dynamics at the eukaryotic DNA replication fork. *J Biol Chem*, 284(7):4041–4045, 2009.

Burgess R.J., Zhou H., Han J. and Zhang Z. A role for Gcn5 in replication-coupled nucleosome assembly. *Mol Cell*, 37(4):469–480, 2010.

Burkhard P., Stetefeld J. and Strelkov S.V. Coiled coils: a highly versatile protein folding motif. *Trends Cell Biol*, 11(2):82–88, 2001.

Cairns B.R. Chromatin remodeling complexes: strength in diversity, precision through specialization. *Curr Opin Genet Dev*, 15(2):185–190, 2005.

Cairns B.R. Chromatin remodeling: insights and intrigue from single-molecule studies. *Nat Struct Mol Biol*, 14(11):989–996, 2007.

Canzio D., Chang E.Y., Shankar S., Kuchenbecker K.M., Simon M.D., Madhani H.D., Narlikar G.J. and Al-Sady B. Chromodomain-mediated oligomerization of HP1 suggests a nucleosome-bridging mechanism for heterochromatin assembly. *Mol Cell*, 41(1):67–81, 2011.

Carey M., Li B. and Workman J.L. RSC exploits histone acetylation to abrogate the nucleosomal block to RNA polymerase II elongation. *Mol Cell*, 24(3):481–487, 2006.

Chai B., Huang J., Cairns B.R. and Laurent B.C. Distinct roles for the RSC and Swi/Snf ATP-dependent chromatin remodelers in DNA double-strand break repair. *Genes Dev*, 19(14):1656–1661, 2005.

Chambeyron S. and Bickmore W.A. Chromatin decondensation and nuclear reorganization of the HoxB locus upon induction of transcription. *Genes Dev*, 18(10):1119–1130, 2004.

Chen C.C., Carson J.J., Feser J., Tamburini B., Zabaronick S., Linger J. and Tyler J.K. Acetylated lysine 56 on histone H3 drives chromatin assembly after repair and signals for the completion of repair. *Cell*, 134(2):231–243, 2008.

Collins N., Poot R.A., Kukimoto I., García-Jiménez C., Dellaire G. and Varga-Weisz P.D. An ACF1-ISWI chromatin-remodeling complex is required for DNA replication through heterochromatin. *Nat Genet*, 32(4):627–632, 2002.

Corona D.F., Eberharter A., Budde A., Deuring R., Ferrari S., Varga-Weisz P., Wilm M., Tamkun J. and Becker P.B. Two histone fold proteins, CHRAC-14 and CHRAC-16, are developmentally regulated subunits of chromatin accessibility complex (CHRAC). *EMBO J*, 19(12):3049–3059, 2000.

Corpet A. and Almouzni G. Making copies of chromatin: the challenge of nucleosomal organization and epigenetic information. *Trends Cell Biol*, 19(1):29–41, 2009.

Costanzi C. and Pehrson J.R. Histone macroH2A1 is concentrated in the inactive X chromosome of female mammals. *Nature*, 393(6685):599–601, 1998.

Creekmore A.L., Walt K.A., Schultz-Norton J.R., Ziegler Y.S., McLeod I.X., Yates J.R. and Nardulli A.M. The role of retinoblastoma-associated proteins 46 and 48 in estrogen receptor alpha mediated gene expression. *Mol Cell Endocrinol*, 291(1-2):79–86, 2008.

Cui F. and Zhurkin V.B. Distinctive sequence patterns in metazoan and yeast nucleosomes: implications for linker histone binding to AT-rich and methylated DNA. *Nucleic Acids Res*, 37(9):2818–2829, 2009.

Das C. and Tyler J.K. Histone exchange and histone modifications during transcription and aging. *Biochim Biophys Acta*, 1819(3-4):332–342, 2012.

Del Rosario B.C. and Pemberton L.F. Nap1 links transcription elongation, chromatin assembly, and messenger RNP complex biogenesis. *Mol Cell Biol*, 28(7):2113–2124, 2008.

Delmas V., Stokes D.G. and Perry R.P. A mammalian DNA-binding protein that contains a chromodomain and an SNF2/SWI2-like helicase domain. *Proc Natl Acad Sci U S A*, 90(6):2414–2418, 1993.

Diffley J.F.X. Quality control in the initiation of eukaryotic DNA replication. *Philos Trans R Soc Lond B Biol Sci*, 366(1584):3545–3553, 2011.

Dignam J.D. Preparation of extracts from higher eukaryotes. *Methods Enzymol*, 182:194–203, 1990.

Dion M.F., Kaplan T., Kim M., Buratowski S., Friedman N. and Rando O.J. Dynamics of replication-independent histone turnover in budding yeast. *Science*, 315(5817):1405–1408, 2007.

Dong H., Lin W., Zhang C.K., Xiong H., Fu G., Jin W.R., Chen R., Chen Z., Qi Z.T. and Huang G.M. Genomic sequence and expression analyses of human chromatin assembly factor 1 p150 gene. *Gene*, 264(2):187–196, 2001.

Donham 2nd D.C., Scorgie J.K. and Churchill M.E.A. The activity of the histone chaperone yeast Asf1 in the assembly and disassembly of histone h3/h4-dna complexes. *Nucleic Acids Res*, 39(13):5449–5458, 2011.

Dror V. and Winston F. The Swi/Snf chromatin remodeling complex is required for ribosomal dna and telomeric silencing in Saccharomyces cerevisiae. *Mol Cell Biol*, 24(18):8227–8235, 2004.

Dudasova Z., Dudas A. and Chovanec M. Non-homologous end-joining factors of saccharomyces cerevisiae. *FEMS Microbiol Rev*, 28(5):581–601, 2004.

Eberharter A., Ferrari S., Längst G., Straub T., Imhof A., Varga-Weisz P., Wilm M. and Becker P.B. Acf1, the largest subunit of CHRAC, regulates ISWI-induced nucleosome remodelling. *EMBO J*, 20(14):3781–3788, 2001.

Eitoku M., Sato L., Senda T. and Horikoshi M. Histone chaperones: 30 years from isolation to elucidation of the mechanisms of nucleosome assembly and disassembly. *Cell Mol Life Sci*, 65(3):414–444, 2008.

Endo M., Ishikawa Y., Osakabe K., Nakayama S., Kaya H., Araki T., ichi Shibahara K., Abe K., Ichikawa H., Valentine L., Hohn B. and Toki S. Increased frequency of homologous recombination and T-DNA integration in Arabidopsis CAF-1 mutants. *EMBO J*, 25(23):5579–5590, 2006.

English C.M., Adkins M.W., Carson J.J., Churchill M.E.A. and Tyler J.K. Structural basis for the histone chaperone activity of Asf1. *Cell*, 127(3):495–508, 2006.

English C.M., Maluf N.K., Tripet B., Churchill M.E.A. and Tyler J.K. ASF1 binds to a heterodimer of histones H3 and H4: a two-step mechanism for the assembly of the H3-H4 heterotetramer on DNA. *Biochemistry*, 44(42):13673–13682, 2005.

Enomoto S. and Berman J. Chromatin assembly factor 1 contributes to the maintenance, but not the re-establishment, of silencing at the yeast silent mating loci. *Genes Dev*, 12(2):219–232, 1998.

Erdel F. and Rippe K. Chromatin remodelling in mammalian cells by ISWI-type complexes–where, when and why? *FEBS J*, 278(19):3608–3618, 2011.

Escargueil A.E., Soares D.G., Salvador M., Larsen A.K. and Henriques J.A.P. What histone code for DNA repair? *Mutat Res*, 658(3):259–270, 2008.

Eskeland R., Eberharter A. and Imhof A. HP1 binding to chromatin methylated at H3K9 is enhanced by auxiliary factors. *Mol Cell Biol*, 27(2):453–465, 2007.

Exner V., Taranto P., Schonrock N., Gruissem W. and Hennig L. Chromatin assembly factor CAF-1 is required for cellular differentiation during plant development. *Development*, 133(21):4163–4172, 2006.

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. Regulation of HP1-chromatin binding by histone H3 methylation and phosphorylation. *Nature*, 438(7071):1116–1122, 2005.

Fitzgerald D.J., Berger P., Schaffitzel C., Yamada K., Richmond T.J. and Berger I. Protein complex expression by using multigene baculoviral vectors. *Nat Methods*, 3(12):1021–1032, 2006.

Fitzgerald D.J., Schaffitzel C., Berger P., Wellinger R., Bieniossek C., Richmond T.J. and Berger I. Multiprotein expression strategy for structural biology of eukaryotic complexes. *Structure*, 15(3):275–279, 2007.

Furuyama T., Dalal Y. and Henikoff S. Chaperone-mediated assembly of centromeric chromatin in vitro. *Proc Natl Acad Sci U S A*, 103(16):6172–6177, 2006.

Gaillard P.H., Martini E.M., Kaufman P.D., Stillman B., Moustacchi E. and Almouzni G. Chromatin assembly coupled to DNA repair: a new role for chromatin assembly factor 1. *Cell*, 86(6):887–896, 1996.

Game J.C. and Kaufman P.D. Role of saccharomyces cerevisiae chromatin assembly factor-i in repair of ultraviolet radiation damage in vivo. *Genetics*, 151(2):485–497, 1999.

Gangaraju V.K. and Bartholomew B. Mechanisms of ATP dependent chromatin remodeling. *Mutat Res*, 618(1-2):3–17, 2007.

Gaszner M. and Felsenfeld G. Insulators: exploiting transcriptional and epigenetic mechanisms. *Nat Rev Genet*, 7(9):703–713, 2006.

Gdula D.A., Sandaltzopoulos R., Tsukiyama T., Ossipow V. and Wu C. Inorganic pyrophosphatase is a component of the Drosophila nucleosome remodeling factor complex. *Genes Dev*, 12(20):3206–3216, 1998.

Gerard A., Koundrioukoff S., Ramillon V., Sergere J.C., Mailand N., Quivy J.P. and Almouzni G. The replication kinase Cdc7-Dbf4 promotes the interaction of the p150 subunit of chromatin assembly factor 1 with proliferating cell nuclear antigen. *EMBO Rep*, 7(8):817–823, 2006.

Glowczewski L., Waterborg J.H. and Berman J.G. Yeast chromatin assembly complex 1 protein excludes nonacetylatable forms of histone H4 from chromatin and the nucleus. *Mol Cell Biol*, 24(23):10180–10192, 2004.

Godde J.S., Nakatani Y. and Wolffe A.P. The amino-terminal tails of the core histones and the translational position of the TATA box determine TBP/TFIIA association with nucleosomal DNA. *Nucleic Acids Res*, 23(22):4557–4564, 1995.

Green C.M. and Almouzni G. Local action of the chromatin assembly factor CAF-1 at sites of nucleotide excision repair in vivo. *EMBO J*, 22(19):5163–5174, 2003.

Green E.M., Antczak A.J., Bailey A.O., Franco A.A., Wu K.J., Yates J.R. and Kaufman P.D. Replication-independent histone deposition by the HIR complex and Asf1. *Curr Biol*, 15(22):2044–2049, 2005.

Groth A. Replicating chromatin: a tale of histones. *Biochem Cell Biol*, 87(1):51–63, 2009.

Groth A., Corpet A., Cook A.J.L., Roche D., Bartek J., Lukas J. and Almouzni G. Regulation of replication fork progression through histone supply and demand. *Science*, 318(5858):1928–1931, 2007a.

Groth A., Rocha W., Verreault A. and Almouzni G. Chromatin challenges during DNA replication and repair. *Cell*, 128(4):721–733, 2007b.

Guillemette B., Bataille A.R., GÃ©vry N., Adam M., Blanchette M., Robert F. and Gaudreau L. Variant histone H2A.Z is globally localized to the promoters of inactive yeast genes and regulates nucleosome positioning. *PLoS Biol*, 3(12):e384, 2005.

Gulbis J.M., Kelman Z., Hurwitz J., O'Donnell M. and Kuriyan J. Structure of the C-terminal region of p21(WAF1/CIP1) complexed with human PCNA. *Cell*, 87(2):297–306, 1996.

Hake S.B. and Allis C.D. Histone H3 variants and their potential role in indexing mammalian genomes: the "H3 barcode hypothesis". *Proc Natl Acad Sci U S A*, 103(17):6428–6435, 2006.

Hakimi M.A., Bochar D.A., Schmiesing J.A., Dong Y., Barak O.G., Speicher D.W., Yokomori K. and Shiekhattar R. A chromatin remodelling complex that loads cohesin onto human chromosomes. *Nature*, 418(6901):994–998, 2002.

Hamiche A., Sandaltzopoulos R., Gdula D.A. and Wu C. ATP-dependent histone octamer sliding mediated by the chromatin remodeling complex NURF. *Cell*, 97(7):833–842, 1999.

Hansen K.H., Bracken A.P., Pasini D., Dietrich N., Gehani S.S., Monrad A., Rappsilber J., Lerdrup M. and Helin K. A model for transmission of the H3K27me3 epigenetic mark. *Nat Cell Biol*, 10(11):1291–1300, 2008.

Happel N. and Doenecke D. Histone H1 and its isoforms: contribution to chromatin structure and function. *Gene*, 431(1-2):1–12, 2009.

Hartley P.D. and Madhani H.D. Mechanisms that specify promoter nucleosome location and identity. *Cell*, 137(3):445–458, 2009.

Hassan A.H., Neely K.E. and Workman J.L. Histone acetyltransferase complexes stabilize SWI/SNF binding to promoter nucleosomes. *Cell*, 104(6):817–827, 2001.

Havas K., Whitehouse I. and Owen-Hughes T. Atp-dependent chromatin remodeling activities. *Cell Mol Life Sci*, 58(5-6):673–682, 2001.

Hengstschläger M., Rosner M., Fountoulakis M. and Lubec G. Regulation of PCNA and CAF-1 expression by the two tuberous sclerosis gene products. *Biochem Biophys Res Commun*, 307(3):737–742, 2003.

Henikoff S. and Ahmad K. Assembly of variant histones into chromatin. *Annu Rev Cell Dev Biol*, 21:133–153, 2005.

Henikoff S., Furuyama T. and Ahmad K. Histone variants, nucleosome assembly and epigenetic inheritance. *Trends Genet*, 20(7):320–326, 2004.

Hoek M., Myers M.P. and Stillman B. An analysis of CAF-1 interacting proteins reveals dynamic and direct interactions with the KU complex and 14-3-3 proteins an analysis of CAF-1 interacting proteins reveals dynamic and direct interactions with the KU complex and 14-3-3 proteins. *J Biol Chem*, 286:10876–10887, 2011.

Hoek M. and Stillman B. Chromatin assembly factor 1 is essential and couples chromatin assembly to DNA replication in vivo. *Proc Natl Acad Sci U S A*, 100(21):12183–12188, 2003.

Hong W., Nakazawa M., Chen Y.Y., Kori R., Vakoc C.R., Rakowski C. and Blobel G.A. FOG-1 recruits the NuRD repressor complex to mediate transcriptional repression by GATA-1. *EMBO J*, 24(13):2367–2378, 2005.

Houlard M., Berlivet S., Probst A.V., Quivy J.P., Hery P., Almouzni G. and Gerard M. CAF-1 is essential for heterochromatin organization in pluripotent embryonic cells. *PLoS Genet*, 2(11):e181, 2006.

Huang J. and Laurent B.C. A role for the RSC chromatin remodeler in regulating cohesion of sister chromatid arms. *Cell Cycle*, 3(8):973–975, 2004.

Huang S., Zhou H., Katzmann D., Hochstrasser M., Atanasova E. and Zhang Z. Rtt106p is a histone chaperone involved in heterochromatin-mediated silencing. *Proc Natl Acad Sci U S A*, 102(38):13410–13415, 2005.

Ishii S., Koshiyama A., Hamada F.N., Nara T.Y., Iwabata K., Sakaguchi K. and Namekawa S.H. Interaction between Lim15/Dmc1 and the homologue of the large subunit of CAF-1: a molecular link between recombination and chromatin assembly during meiosis. *FEBS J*, 275(9):2032–2041, 2008.

Ito T., Bulger M., Kobayashi R. and Kadonaga J.T. Drosophila NAP-1 is a core histone chaperone that functions in ATP-facilitated assembly of regularly spaced nucleosomal arrays. *Mol Cell Biol*, 16(6):3112–3124, 1996.

Ito T., Levenstein M.E., Fyodorov D.V., Kutach A.K., Kobayashi R. and Kadonaga J.T. ACF consists of two subunits, Acf1 and ISWI, that function cooperatively in the ATP-dependent catalysis of chromatin assembly. *Genes Dev*, 13(12):1529–1539, 1999.

Jackson V. Deposition of newly synthesized histones: hybrid nucleosomes are not tandemly arranged on daughter dna strands. *Biochemistry*, 27(6):2109–2120, 1988.

Jackson V. In vivo studies on the dynamics of histone-DNA interaction: evidence for nucleosome dissolution during replication and transcription and a low level of dissolution independent of both. *Biochemistry*, 29(3):719–731, 1990.

Jackson V. and Chalkley R. A reevaluation of new histone deposition on replicating chromatin. *J Biol Chem*, 256(10):5095–5103, 1981.

Jasencakova Z., Scharf A.N.D., Ask K., Corpet A., Imhof A., Almouzni G. and Groth A. Replication stress interferes with histone recycling and predeposition marking of new histones. *Mol Cell*, 37(5):736–743, 2010.

Jenuwein T. and Allis C.D. Translating the histone code. *Science*, 293(5532):1074–1080, 2001.

Jiao R., Bachrati C.Z., Pedrazzi G., Kuster P., Petkovic M., Li J.L., Egli D., Hickson I.D. and Stagljar I. Physical and functional interaction between the Bloom's syndrome gene product and the largest subunit of chromatin assembly factor 1. *Mol Cell Biol*, 24(11):4710–4719, 2004.

Jiao R., Harrigan J.A., Shevelev I., Dietschy T., Selak N., Indig F.E., Piotrowski J., Janscak P., Bohr V.A. and Stagljar I. The Werner syndrome protein is required for recruitment of chromatin assembly factor 1 following DNA damage. *Oncogene*, 26(26):3811–3822, 2007.

Jin J., Cai Y., Li B., Conaway R.C., Workman J.L., Conaway J.W. and Kusch T. In and out: histone variant exchange in chromatin. *Trends Biochem Sci*, 30(12):680–687, 2005.

Johnson S.M., Tan F.J., McCullough H.L., Riordan D.P. and Fire A.Z. Flexibility and constraint in the nucleosome core landscape of caenorhabditis elegans chromatin. *Genome Res*, 16(12):1505–1516, 2006.

Kamakaka R.T., Bulger M., Kaufman P.D., Stillman B. and Kadonaga J.T. Postreplicative chromatin assembly by Drosophila and human chromatin assembly factor 1. *Mol Cell Biol*, 16(3):810–817, 1996.

Kang B., Pu M., Hu G., Wen W., Dong Z., Zhao K., Stillman B. and Zhang Z. Phosphorylation of H4 Ser 47 promotes HIRA-mediated nucleosome assembly. *Genes Dev*, 25(13):1359–1364, 2011.

Kang J.G., Hamiche A. and Wu C. GAL4 directs nucleosome sliding induced by NURF. *EMBO J*, 21(6):1406–1413, 2002.

Kaufman P.D., Cohen J.L. and Osley M.A. Hir proteins are required for position-dependent gene silencing in Saccharomyces cerevisiae in the absence of chromatin assembly factor 1. *Mol Cell Biol*, 18(8):4793–4806, 1998.

Kaufman P.D., Kobayashi R., Kessler N. and Stillman B. The p150 and p60 subunits of chromatin assembly factor I: a molecular link between newly synthesized histones and DNA replication. *Cell*, 81(7):1105–1114, 1995.

Kaufman P.D., Kobayashi R. and Stillman B. Ultraviolet radiation sensitivity and reduction of telomeric silencing in Saccharomyces cerevisiae cells lacking chromatin assembly factor-1. *Genes Dev*, 11(3):345–357, 1997.

Kaya H., Shibahara K.I., Taoka K.I., Iwabuchi M., Stillman B. and Araki T. FASCIATA genes for chromatin assembly factor-1 in Arabidopsis maintain the cellular organization of apical meristems. *Cell*, 104(1):131–142, 2001.

Keller C. and Krude T. Requirement of Cyclin/Cdk2 and protein phosphatase 1 activity for chromatin assembly factor 1-dependent chromatin assembly during DNA synthesis. *J Biol Chem*, 275(45):35512–35521, 2000.

Kim H.J., Seol J.H., Han J.W., Youn H.D. and Cho E.J. Histone chaperones regulate histone exchange during transcription. *EMBO J*, 26(21):4467–4474, 2007.

Kireeva M.L., Walter W., Tchernajenko V., Bondarenko V., Kashlev M. and Studitsky V.M. Nucleosome remodeling induced by RNA polymerase II: loss of the H2A/H2B dimer during transcription. *Mol Cell*, 9(3):541–552, 2002.

Kirik A., Pecinka A., Wendeler E. and Reiss B. The chromatin assembly factor subunit FASCIATA1 is involved in homologous recombination in plants. *Plant Cell*, 18(10):2431–2442, 2006.

Kitagawa H., Fujiki R., Yoshimura K., Mezaki Y., Uematsu Y., Matsui D., Ogawa S., Unno K., Okubo M., Tokita A., Nakagawa T., Ito T., Ishimi Y., Nagasawa H., Matsumoto T., Yanagisawa J. and Kato S. The chromatin-remodeling complex WINAC targets a nuclear receptor to promoters and is impaired in williams syndrome. *Cell*, 113(7):905–917, 2003.

Konev A.Y., Tribus M., Park S.Y., Podhraski V., Lim C.Y., Emelyanov A.V., Vershilova E., Pirrotta V., Kadonaga J.T., Lusser A. and Fyodorov D.V. CHD1 motor protein is required for deposition of histone variant H3.3 into chromatin in vivo. *Science*, 317(5841):1087–1090, 2007.

Kong L., Yu X.P., Bai X.H., Zhang W.F., Zhang Y., Zhao W.M., Jia J.H., Tang W., Zhou Y.B. and ju Liu C. RbAp48 is a critical mediator controlling the transforming activity of human papillomavirus type 16 in cervical cancer. *J Biol Chem*, 282(36):26381–26391, 2007.

Korber P. and Hörz W. In vitro assembly of the characteristic chromatin organization at the yeast PHO5 promoter by a replication-independent extract system. *J Biol Chem*, 279(33):35113–35120, 2004.

Kornberg R.D. Structure of chromatin. *Annu Rev Biochem*, 46:931–954, 1977.

Kornberg R.D. and Stryer L. Statistical distributions of nucleosomes: nonrandom locations by a stochastic mechanism. *Nucleic Acids Res*, 16(14A):6677–6690, 1988.

Kouzarides T. Chromatin modifications and their function. *Cell*, 128(4):693–705, 2007.

Krishna T.S., Kong X.P., Gary S., Burgers P.M. and Kuriyan J. Crystal structure of the eukaryotic DNA polymerase processivity factor PCNA. *Cell*, 79(7):1233–1243, 1994.

Krude T. Chromatin assembly factor 1 (CAF-1) colocalizes with replication foci in HeLa cell nuclei. *Exp Cell Res*, 220(2):304–311, 1995.

Kukimoto I., Elderkin S., Grimaldi M., Oelgeschläger T. and Varga-Weisz P.D. The histone-fold protein complex CHRAC-15/17 enhances nucleosome sliding and assembly mediated by ACF. *Mol Cell*, 13(2):265–277, 2004.

Kwon S.H. and Workman J.L. The changing faces of HP1: From heterochromatin formation and gene silencing to euchromatic gene expression: HP1 acts as a positive regulator of transcription. *Bioessays*, 33(4):280–289, 2011.

Längst G., Bonte E.J., Corona D.F. and Becker P.B. Nucleosome movement by CHRAC and ISWI without disruption or trans-displacement of the histone octamer. *Cell*, 97(7):843–852, 1999.

Lassle M., Richter A. and Knippers R. Comparison of replicative and non-replicative chromatin assembly pathways in HeLa cell extracts. *Biochim Biophys Acta*, 1132(1):1–10, 1992.

Le S., Davis C., Konopka J.B. and Sternglanz R. Two new S-phase-specific genes from Saccharomyces cerevisiae. *Yeast*, 13(11):1029–1042, 1997.

Lee C.K., Shibata Y., Rao B., Strahl B.D. and Lieb J.D. Evidence for nucleosome depletion at active regulatory regions genome-wide. *Nat Genet*, 36(8):900–905, 2004.

Lee S.B., Lee C.F., Ou D.S.C., Dulal K., Chang L.H., Ma C.H., Huang C.F., Zhu H., Lin Y.S. and Juan L.J. Host-viral effects of chromatin assembly factor 1 interaction with HCMV IE2. *Cell Res*, 2011.

Lee S.B., Ou D.S.C., Lee C.F. and Juan L.J. Gene-specific transcriptional activation mediated by the p150 subunit of the chromatin assembly factor 1. *J Biol Chem*, 284(21):14040–14049, 2009.

Lee W., Tillo D., Bray N., Morse R.H., Davis R.W., Hughes T.R. and Nislow C. A high-resolution atlas of nucleosome occupancy in yeast. *Nat Genet*, 39(10):1235–1244, 2007.

Lejon S., Thong S.Y., Murthy A., AlQarni S., Murzina N.V., Blobel G.A., Laue E.D. and Mackay J.P. Insights into association of the NuRD complex with FOG-1 from the crystal structure of an RbAp48/FOG-1 complex. *J Biol Chem*, 286(2):1196–1203, 2011.

Leong M.M., Milstein C. and Pannell R. Luminescent detection method for immunodot, Western, and Southern blots. *J Histochem Cytochem*, 34(12):1645–1650, 1986.

LeRoy G., Orphanides G., Lane W.S. and Reinberg D. Requirement of RSF and FACT for transcription of chromatin templates in vitro. *Science*, 282(5395):1900–1904, 1998.

Levchenko V. and Jackson V. Histone release during transcription: NAP1 forms a complex with H2A and H2B and facilitates a topologically dependent release of H3 and H4 from the nucleosome. *Biochemistry*, 43(9):2359–2372, 2004.

Lewis J.D., Meehan R.R., Henzel W.J., Maurer-Fogy I., Jeppesen P., Klein F. and Bird A. Purification, sequence, and cellular localization of a novel chromosomal protein that binds to methylated DNA. *Cell*, 69(6):905–914, 1992.

Li C. and Jin J. DNA replication licensing control and rereplication prevention. *Protein Cell*, 1(3):227–236, 2010.

Li J.J. and Kelly T.J. Simian virus 40 DNA replication in vitro. *Proc Natl Acad Sci U S A*, 81(22):6973–6977, 1984.

Li M.Z. and Elledge S.J. Harnessing homologous recombination in vitro to generate recombinant DNA via SLIC. *Nat Methods*, 4(3):251–256, 2007.

Li Y., DeFatta R., Anthony C., Sunavala G. and Benedetti A.D. A translationally regulated Tousled kinase phosphorylates histone H3 and confers radioresistance when overexpressed. *Oncogene*, 20(6):726–738, 2001.

Linger J. and Tyler J.K. The yeast histone chaperone chromatin assembly factor 1 protects against double-strand DNA-damaging agents. *Genetics*, 171(4):1513–1522, 2005.

Lorch Y., Cairns B.R., Zhang M. and Kornberg R.D. Activated RSC-nucleosome complex and persistently altered form of the nucleosome. *Cell*, 94(1):29–34, 1998.

Lorch Y., Davis B. and Kornberg R.D. Chromatin remodeling by DNA bending, not twisting. *Proc Natl Acad Sci U S A*, 102(5):1329–1332, 2005.

Lorch Y., Zhang M. and Kornberg R.D. Histone octamer transfer by a chromatin-remodeling complex. *Cell*, 96(3):389–392, 1999.

Luger K., Mäder A.W., Richmond R.K., Sargent D.F. and Richmond T.J. Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature*, 389(6648):251–260, 1997.

Lusser A. and Kadonaga J.T. Chromatin remodeling by ATP-dependent molecular machines. *Bioessays*, 25(12):1192–1200, 2003.

Macdonald N., Welburn J.P.I., Noble M.E.M., Nguyen A., Yaffe M.B., Clynes D., Moggs J.G., Orphanides G., Thomson S., Edmunds J.W., Clayton A.L., Endicott J.A. and Mahadevan L.C. Molecular basis for the recognition of phosphorylated and phosphoacetylated histone H3 by 14-3-3. *Mol Cell*, 20(2):199–211, 2005.

Maison C. and Almouzni G. HP1 and the dynamics of heterochromatin maintenance. *Nat Rev Mol Cell Biol*, 5(4):296–304, 2004.

Malay A.D., Umehara T., Matsubara-Malay K., Padmanabhan B. and Yokoyama S. Crystal structures of fission yeast histone chaperone Asf1 complexed with the Hip1 B-domain or the Cac2 C terminus. *J Biol Chem*, 283(20):14022–14031, 2008.

Margueron R. and Reinberg D. The Polycomb complex PRC2 and its mark in life. *Nature*, 469(7330):343–349, 2011.

Marheineke K. and Krude T. Nucleosome assembly activity and intracellular localization of human CAF-1 changes during the cell division cycle. *J Biol Chem*, 273(24):15279–15286, 1998.

Martin C. and Zhang Y. Mechanisms of epigenetic inheritance. *Curr Opin Cell Biol*, 19(3):266–272, 2007.

Martini E., Roche D.M., Marheineke K., Verreault A. and Almouzni G. Recruitment of phosphorylated chromatin assembly factor 1 to chromatin after UV irradiation of human cells. *J Cell Biol*, 143(3):563–575, 1998.

Mascolo M., Vecchione M.L., Ilardi G., Scalvenzi M., Molea G., Benedetto M.D., Nugnes L., Siano M., Rosa G.D. and Staibano S. Overexpression of chromatin assembly factor-1/p60 helps to predict the prognosis of melanoma patients. *BMC Cancer*, 10:63, 2010.

Masumoto H., Hawke D., Kobayashi R. and Verreault A. A role for cell-cycle-regulated histone H3 lysine 56 acetylation in the DNA damage response. *Nature*, 436(7048):294–298, 2005.

Mavrich T.N., Ioshikhes I.P., Venters B.J., Jiang C., Tomsho L.P., Qi J., Schuster S.C., Albert I. and Pugh B.F. A barrier nucleosome model for statistical positioning of nucleosomes throughout the yeast genome. *Genome Res*, 18(7):1073–1083, 2008a.

Mavrich T.N., Jiang C., Ioshikhes I.P., Li X., Venters B.J., Zanton S.J., Tomsho L.P., Qi J., Glaser R.L., Schuster S.C., Gilmour D.S., Albert I. and Pugh B.F. Nucleosome organization in the Drosophila genome. *Nature*, 453(7193):358–362, 2008b.

Meijsing S.H. and Ehrenhofer-Murray A.E. The silencing complex SAS-I links histone acetylation to the assembly of repressed chromatin by CAF-I and Asf1 in Saccharomyces cerevisiae. *Genes Dev*, 15(23):3169–3182, 2001.

Mello J.A., Sillje H.H.W., Roche D.M.J., Kirschner D.B., Nigg E.A. and Almouzni G. Human asf1 and CAF-1 interact and synergize in a repair-coupled nucleosome assembly pathway. *EMBO Rep*, 3(4):329–334, 2002.

Miller A., Yang B., Foster T. and Kirchmaier A.L. Proliferating cell nuclear antigen and ASF1 modulate silent chromatin in Saccharomyces cerevisiae via lysine 56 on histone H3. *Genetics*, 179(2):793–809, 2008.

Mito Y., Henikoff J.G. and Henikoff S. Histone replacement marks the boundaries of cis-regulatory domains. *Science*, 315(5817):1408–1411, 2007.

Mizuguchi G., Shen X., Landry J., Wu W.H., Sen S. and Wu C. ATP-driven exchange of histone H2AZ variant catalyzed by SWR1 chromatin remodeling complex. *Science*, 303(5656):343–348, 2004.

Mizuguchi G., Tsukiyama T., Wisniewski J. and Wu C. Role of nucleosome remodeling factor NURF in transcriptional activation of chromatin. *Mol Cell*, 1(1):141–150, 1997.

Moggs J.G., Grandi P., Quivy J.P., Jonsson Z.O., Hubscher U., Becker P.B. and Almouzni G. A CAF-1-PCNA-mediated chromatin assembly pathway triggered by sensing DNA damage. *Mol Cell Biol*, 20(4):1206–1218, 2000.

Moldovan G.L., Pfander B. and Jentsch S. PCNA, the maestro of the replication fork. *Cell*, 129(4):665–679, 2007.

Monson E.K., de Bruin D. and Zakian V.A. The yeast Cac1 protein is required for the stable inheritance of transcriptionally repressed chromatin at telomeres. *Proc Natl Acad Sci U S A*, 94(24):13081–13086, 1997.

Munakata T., Adachi N., Yokoyama N., Kuzuhara T. and Horikoshi M. A human homologue of yeast anti-silencing factor has histone chaperone activity. *Genes Cells*, 5(3):221–233, 2000.

Murzina N., Verreault A., Laue E. and Stillman B. Heterochromatin dynamics in mouse cells: interaction between chromatin assembly factor 1 and HP1 proteins. *Mol Cell*, 4(4):529–540, 1999.

Murzina N.V., Pei X.Y., Zhang W., Sparkes M., Vicente-Garcia J., Pratap J.V., McLaughlin S.H., Ben-Shahar T.R., Verreault A., Luisi B.F. and Laue E.D. Structural basis for the recognition of histone H4 by the histone-chaperone RbAp46. *Structure*, 16(7):1077–1085, 2008.

Nabatiyan A. and Krude T. Silencing of chromatin assembly factor 1 in human cells leads to cell death and loss of chromatin assembly during DNA synthesis. *Mol Cell Biol*, 24(7):2853–2862, 2004.

Nabatiyan A., Szuts D. and Krude T. Induction of CAF-1 expression in response to DNA strand breaks in quiescent human cells. *Mol Cell Biol*, 26(5):1839–1849, 2006.

Nakatani Y., Ray-Gallet D., Quivy J.P., Tagami H. and Almouzni G. Two distinct nucleosome assembly pathways: dependent or independent of DNA synthesis promoted by histone H3.1 and H3.3 complexes. *Cold Spring Harb Symp Quant Biol*, 69:273–280, 2004.

Naryzhny S.N., Zhao H. and Lee H. Proliferating cell nuclear antigen (PCNA) may function as a double homotrimer complex in the mammalian cell. *J Biol Chem*, 280(14):13888–13894, 2005.

Nathan D. and Crothers D.M. Bending and flexibility of methylated and unmethylated EcoRI DNA. *J Mol Biol*, 316(1):7–17, 2002.

Natsume R., Eitoku M., Akai Y., Sano N., Horikoshi M. and Senda T. Structure and function of the histone chaperone CIA/ASF1 complexed with histones H3 and H4. *Nature*, 446(7133):338–341, 2007.

Nelson H.C., Finch J.T., Luisi B.F. and Klug A. The structure of an oligo(dA).oligo(dT) tract and its biological implications. *Nature*, 330(6145):221–226, 1987.

Neumann H., Hancock S.M., Buning R., Routh A., Chapman L., Somers J., Owen-Hughes T., van Noort J., Rhodes D. and Chin J.W. A method for genetically installing site-specific acetylation in recombinant histones defines the effects of H3 K56 acetylation. *Mol Cell*, 36(1):153–163, 2009.

Nicolas E., Morales V., Magnaghi-Jaulin L., Harel-Bellan A., Richard-Foy H. and Trouche D. RbAp48 belongs to the histone deacetylase complex that associates with the retinoblastoma protein. *J Biol Chem*, 275(13):9797–9804, 2000.

Nishino Y., Eltsov M., Joti Y., Ito K., Takata H., Takahashi Y., Hihara S., Frangakis A.S., Imamoto N., Ishikawa T. and Maeshima K. Human mitotic chromosomes consist predominantly of irregularly folded nucleosome fibres without a 30-nm chromatin structure. *EMBO J*, 31(7):1644–1653, 2012.

Noma K., Allis C.D. and Grewal S.I. Transitions in distinct histone H3 methylation patterns at the heterochromatin domain boundaries. *Science*, 293(5532):1150–1155, 2001.

Okano S., Lan L., Caldecott K.W., Mori T. and Yasui A. Spatial and temporal cellular responses to single-strand breaks in human cells. *Mol Cell Biol*, 23(11):3974–3981, 2003.

Ono T., Kaya H., Takeda S., Abe M., Ogawa Y., Kato M., Kakutani T., Scheid O.M., Araki T. and ichi Shibahara K. Chromatin assembly factor 1 ensures the stable maintenance of silent chromatin states in Arabidopsis. *Genes Cells*, 11(2):153–162, 2006.

Orphanides G., LeRoy G., Chang C.H., Luse D.S. and Reinberg D. FACT, a factor that facilitates transcript elongation through nucleosomes. *Cell*, 92(1):105–116, 1998.

Pacifico F., Paolillo M., Chiappetta G., Crescenzi E., Arena S., Scaloni A., Monaco M., Vascotto C., Tell G., Formisano S. and Leonardi A. RbAp48 is a target of nuclear factor-kappaB activity in thyroid cancer. *J Clin Endocrinol Metab*, 92(4):1458–1466, 2007.

Park Y.J., Chodaparambil J.V., Bao Y., McBryant S.J. and Luger K. Nucleosome assembly protein 1 exchanges histone H2A-H2B dimers and assists nucleosome sliding. *J Biol Chem*, 280(3):1817–1825, 2005.

Parthun M.R., Widom J. and Gottschling D.E. The major cytoplasmic histone acetyltransferase in yeast: links to chromatin replication and histone metabolism. *Cell*, 87(1):85–94, 1996.

Perpelescu M., Nozaki N., Obuse C., Yang H. and Yoda K. Active establishment of centromeric CENP-A chromatin by RSF complex. *J Cell Biol*, 185(3):397–407, 2009.

Phelan M.L., Schnitzler G.R. and Kingston R.E. Octamer transfer and creation of stably remodeled nucleosomes by human SWI-SNF and its isolated ATPases. *Mol Cell Biol*, 20(17):6380–6389, 2000.

Phillips J.E. and Corces V.G. CTCF: master weaver of the genome. *Cell*, 137(7):1194–1211, 2009.

Poleshko A., Einarson M.B., Shalginskikh N., Zhang R., Adams P.D., Skalka A.M. and Katz R.A. Identification of a functional network of human epigenetic silencing factors. *J Biol Chem*, 285:422–433, 2010.

Polo S.E., Roche D. and Almouzni G. New histone incorporation marks sites of UV repair in human cells. *Cell*, 127(3):481–493, 2006.

Polo S.E., Theocharis S.E., Klijanienko J., Savignoni A., Asselain B., Vielh P. and Almouzni G. Chromatin assembly factor-1, a marker of clinical value to distinguish quiescent from proliferating cells. *Cancer Res*, 64(7):2371–2381, 2004.

Poot R.A., Bozhenok L., van den Berg D.L.C., Steffensen S., Ferreira F., Grimaldi M., Gilbert N., Ferreira J. and Varga-Weisz P.D. The Williams syndrome transcription factor interacts with pcna to target chromatin remodelling by ISWI to replication foci. *Nat Cell Biol*, 6(12):1236–1244, 2004.

Prilusky J., Felder C.E., Zeev-Ben-Mordehai T., Rydberg E.H., Man O., Beckmann J.S., Silman I. and Sussman J.L. FoldIndex: a simple tool to predict whether a given protein sequence is intrinsically unfolded. *Bioinformatics*, 21(16):3435–3438, 2005.

Probst A.V., Dunleavy E. and Almouzni G. Epigenetic inheritance during the cell cycle. *Nat Rev Mol Cell Biol*, 10(3):192–206, 2009.

Prochasson P., Florens L., Swanson S.K., Washburn M.P. and Workman J.L. The HIR corepressor complex binds to nucleosomes generating a distinct protein/DNA complex resistant to remodeling by SWI/SNF. *Genes Dev*, 19(21):2534–2539, 2005.

Pusarla R.H. and Bhargava P. Histones in functional diversification. core histone variants. *FEBS J*, 272(20):5149–5168, 2005.

Qian Y.W., Wang Y.C., Hollingsworth R.E., Jones D., Ling N. and Lee E.Y. A retinoblastoma-binding protein related to a negative regulator of Ras in yeast. *Nature*, 364(6438):648–652, 1993.

Quivy J.P., Gérard A., Cook A.J.L., Roche D. and Almouzni G. The HP1-p150/CAF-1 interaction is required for pericentric heterochromatin replication and S-phase progression in mouse cells. *Nat Struct Mol Biol*, 15:972 – 979, 2008.

Quivy J.P., Grandi P. and Almouzni G. Dimerization of the largest subunit of chromatin assembly factor 1: importance in vitro and during Xenopus early development. *EMBO J*, 20(8):2015–2027, 2001.

Quivy J.P., Roche D., Kirschner D., Tagami H., Nakatani Y. and Almouzni G. A CAF-1 dependent pool of HP1 during heterochromatin duplication. *EMBO J*, 23(17):3516–3526, 2004.

Rai T.S., Puri A., McBryan T., Hoffman J., Tang Y., Pchelintsev N.A., van Tuyn J., Marmorstein R., Schultz D.C. and Adams P.D. Human CABIN1 is a functional member of the human HIRA/UBN1/ASF1a histone H3.3 chaperone complex. *Mol Cell Biol*, 31(19):4107–4118, 2011.

Ramírez J. and Hagman J. The Mi-2/NuRD complex: a critical epigenetic regulator of hematopoietic development, differentiation and cancer. *Epigenetics*, 4(8):532–536, 2009.

Ransom M., Dennehey B.K. and Tyler J.K. Chaperoning histones during DNA replication and repair. *Cell*, 140(2):183–195, 2010.

Ray-Gallet D., Quivy J.P., Silljé H.W.W., Nigg E.A. and Almouzni G. The histone chaperone Asf1 is dispensable for direct de novo histone deposition in Xenopus egg extracts. *Chromosoma*, 116(5):487–496, 2007.

Rea S., Eisenhaber F., O'Carroll D., Strahl B.D., Sun Z.W., Schmid M., Opravil S., Mechtler K., Ponting C.P., Allis C.D. and Jenuwein T. Regulation of chromatin structure by site-specific histone H3 methyltransferases. *Nature*, 406(6796):593–599, 2000.

Reese B.E., Bachman K.E., Baylin S.B. and Rountree M.R. The methyl-CpG binding protein MBD1 interacts with the p150 subunit of chromatin assembly factor 1. *Mol Cell Biol*, 23(9):3226–3236, 2003.

Reinke H. and Hörz W. Anatomy of a hypersensitive site. *Biochim Biophys Acta*, 1677(1-3):24–29, 2004.

Ren J., Wen L., Gao X., Jin C., Xue Y. and Yao X. DOG 1.0: illustrator of protein domain structures. *Cell Res*, 19(2):271–273, 2009.

Rippe K., Schrader A., Riede P., Strohner R., Lehmann E. and Längst G. DNA sequence- and conformation-directed positioning of nucleosomes by chromatin-remodeling complexes. *Proc Natl Acad Sci U S A*, 104(40):15635–15640, 2007.

Robinson P.J.J., Fairall L., Huynh V.A.T. and Rhodes D. EM measurements define the dimensions of the "30-nm" chromatin fiber: evidence for a compact, interdigitated structure. *Proc Natl Acad Sci U S A*, 103(17):6506–6511, 2006.

Robinson P.J.J. and Rhodes D. Structure of the '30 nm' chromatin fibre: a key role for the linker histone. *Curr Opin Struct Biol*, 16(3):336–343, 2006.

Rocha W. and Verreault A. Clothing up DNA for all seasons: Histone chaperones and nucleosome assembly pathways. *FEBS Lett*, 582(14):1938–1949, 2008.

Sancar A., Lindsey-Boltz L.A., Unsal-KaÃ§maz K. and Linn S. Molecular mechanisms of mammalian DNA repair and the DNA damage checkpoints. *Annu Rev Biochem*, 73:39–85, 2004.

Sanematsu F., Takami Y., Barman H.K., Fukagawa T., Ono T., Shibahara K.I. and Nakayama T. Asf1 is required for viability and chromatin assembly during DNA replication in vertebrate cells. *J Biol Chem*, 281(19):13817–13827, 2006.

Sarraf S.A. and Stancheva I. Methyl-cpg binding protein MBD1 couples histone H3 methylation at lysine 9 by SETDB1 to DNA replication and chromatin assembly. *Mol Cell*, 15(4):595–605, 2004.

Savitsky P., Bray J., Cooper C.D.O., Marsden B.D., Mahajan P., Burgess-Brown N.A. and Gileadi O. High-throughput production of human proteins for crystallization: the SGC experience. *J Struct Biol*, 172(1):3–13, 2010.

Schalch T., Duda S., Sargent D.F. and Richmond T.J. X-ray structure of a tetranucleosome and its implications for the chromatin fibre. *Nature*, 436(7047):138–141, 2005.

Schmitges F.W., Prusty A.B., Faty M., Stützer A., Lingaraju G.M., Aiwazian J., Sack R., Hess D., Li L., Zhou S., Bunker R.D., Wirth U., Bouwmeester T., Bauer A., Ly-Hartig

N., Zhao K., Chan H., Gu J., Gut H., Fischle W., Müller J. and Thomä N.H. Histone methylation by PRC2 is inhibited by active chromatin marks. *Mol Cell*, 42(3):330–341, 2011.

Schones D.E., Cui K., Cuddapah S., Roh T.Y., Barski A., Wang Z., Wei G. and Zhao K. Dynamic regulation of nucleosome positioning in the human genome. *Cell*, 132(5):887–898, 2008.

Schönrock N., Exner V., Probst A., Gruissem W. and Hennig L. Functional genomic analysis of CAF-1 mutants in Arabidopsis thaliana. *J Biol Chem*, 281(14):9560–9568, 2006.

Schwabish M.A. and Struhl K. Asf1 mediates histone eviction and deposition during elongation by RNA polymerase II. *Mol Cell*, 22(3):415–422, 2006.

Segal E. and Widom J. Poly(dA:dT) tracts: major determinants of nucleosome organization. *Curr Opin Struct Biol*, 19(1):65–71, 2009a.

Segal E. and Widom J. Poly(dA:dT) tracts: major determinants of nucleosome organization. *Curr Opin Struct Biol*, 19(1):65–71, 2009b.

Segal M.R. Re-cracking the nucleosome positioning code. *Stat Appl Genet Mol Biol*, 7(1):Article14, 2008.

Sharp J.A., Franco A.A., Osley M.A. and Kaufman P.D. Chromatin assembly factor i and Hir proteins contribute to building functional kinetochores in S. cerevisiae. *Genes Dev*, 16(1):85–100, 2002.

Shibahara K. and Stillman B. Replication-dependent marking of DNA by PCNA facilitates CAF-1-coupled inheritance of chromatin. *Cell*, 96(4):575–585, 1999.

Shivaswamy S., Bhinge A., Zhao Y., Jones S., Hirst M. and Iyer V.R. Dynamic remodeling of individual nucleosomes across a eukaryotic genome in response to transcriptional perturbation. *PLoS Biol*, 6(3):e65, 2008.

Shogren-Knaak M., Ishii H., Sun J.M., Pazin M.J., Davie J.R. and Peterson C.L. Histone H4-K16 acetylation controls chromatin structure and protein interactions. *Science*, 311(5762):844–847, 2006.

Smith M.M. Centromeres and variant histones: what, where, when and why? *Curr Opin Cell Biol*, 14(3):279–285, 2002.

Smith S. and Stillman B. Purification and characterization of CAF-I, a human cell factor required for chromatin assembly during dna replication in vitro. *Cell*, 58(1):15–25, 1989.

Smith S. and Stillman B. Immunological characterization of chromatin assembly factor I, a human cell factor required for chromatin assembly during DNA replication in vitro. *J Biol Chem*, 266(18):12041–12047, 1991.

Sobel R.E., Cook R.G., Perry C.A., Annunziato A.T. and Allis C.D. Conservation of deposition-related acetylation sites in newly synthesized histones H3 and H4. *Proc Natl Acad Sci U S A*, 92(4):1237–1241, 1995.

Sogo J.M., Stahl H., Koller T. and Knippers R. Structure of replicating simian virus 40 minichromosomes. the replication fork, core histone segregation and terminal structures. *J Mol Biol*, 189(1):189–204, 1986.

Song J.J., Garlick J.D. and Kingston R.E. Structural basis of histone H4 recognition by p55. *Genes Dev*, 22(10):1313–1318, 2008.

Song Y., He F., Xie G., Guo X., Xu Y., Chen Y., Liang X., Stagljar I., Egli D., Ma J. and Jiao R. CAF-1 is essential for Drosophila development and involved in the maintenance of epigenetic memory. *Dev Biol*, 311(1):213–222, 2007.

Staibano S., Mascolo M., Mancini F.P., Kisslinger A., Salvatore G., Benedetto M.D., Chieffi P., Altieri V., Prezioso D., Ilardi G., Rosa G.D. and Tramontano D. Overexpression of chromatin assembly factor-1 (CAF-1) p60 is predictive of adverse behaviour of prostatic cancer. *Histopathology*, 54(5):580–589, 2009.

Staibano S., Mignogna C., Muzio L.L., Mascolo M., Salvatore G., Benedetto M.D., Califano L., Rubini C. and Rosa G.D. Chromatin assembly factor-1 (CAF-1)-mediated regulation of cell proliferation and DNA repair: a link with the biological behaviour of squamous cell carcinoma of the tongue? *Histopathology*, 50(7):911–919, 2007.

Staynov D.Z. The controversial 30 nm chromatin fibre. *Bioessays*, 30(10):1003–1009, 2008.

Stillman B. Chromatin assembly during SV40 DNA replication in vitro. *Cell*, 45(4):555–565, 1986.

Stokes D.G. and Perry R.P. DNA-binding and chromatin localization properties of CHD1. *Mol Cell Biol*, 15(5):2745–2753, 1995.

Stokes D.G., Tartof K.D. and Perry R.P. CHD1 is concentrated in interbands and puffed regions of drosophila polytene chromosomes. *Proc Natl Acad Sci U S A*, 93(14):7137–7142, 1996.

Szenker E., Ray-Gallet D. and Almouzni G. The double face of the histone variant H3.3. *Cell Res*, 21(3):421–434, 2011.

Taddei A., Roche D., Sibarita J.B., Turner B.M. and Almouzni G. Duplication and maintenance of heterochromatin domains. *J Cell Biol*, 147(6):1153–1166, 1999.

Tagami H., Ray-Gallet D., Almouzni G. and Nakatani Y. Histone H3.1 and H3.3 complexes mediate nucleosome assembly pathways dependent or independent of DNA synthesis. *Cell*, 116(1):51–61, 2004.

Takami Y., Ono T., Fukagawa T., ichi Shibahara K. and Nakayama T. Essential role of chromatin assembly factor-1-mediated rapid nucleosome assembly for DNA replication and cell division in vertebrate cells. *Mol Biol Cell*, 18(1):129–141, 2007.

Talbert P.B. and Henikoff S. Histone variants–ancient wrap artists of the epigenome. *Nat Rev Mol Cell Biol*, 11(4):264–275, 2010.

Tamaru H. Confining euchromatin/heterochromatin territory: jumonji crosses the line. *Genes Dev*, 24(14):1465–1478, 2010.

Tan B.C.M., Chien C.T., Hirose S. and Lee S.C. Functional cooperation between FACT and MCM helicase facilitates initiation of chromatin DNA replication. *EMBO J*, 25(17):3975–3985, 2006.

Tang L., Nogales E. and Ciferri C. Structure and function of SWI/SNF chromatin remodeling complexes and mechanistic implications for transcription. *Prog Biophys Mol Biol*, 102(2-3):122–128, 2010.

Tang Y., Poustovoitov M.V., Zhao K., Garfinkel M., Canutescu A., Dunbrack R., Adams P.D. and Marmorstein R. Structure of a human ASF1a-HIRA complex and insights into specificity of histone chaperone complex assembly. *Nat Struct Mol Biol*, 13(10):921–929, 2006.

Tchénio T., Casella J.F. and Heidmann T. A truncated form of the human CAF-1 p150 subunit impairs the maintenance of transcriptional gene silencing in mammalian cells. *Mol Cell Biol*, 21(6):1953–1961, 2001.

Thaström A., Lowary P.T., Widlund H.R., Cao H., Kubista M. and Widom J. Sequence motifs and free energies of selected natural and non-natural nucleosome positioning DNA sequences. *J Mol Biol*, 288(2):213–229, 1999.

Torigoe S.E., Urwin D.L., Ishii H., Smith D.E. and Kadonaga J.T. Identification of a rapidly formed nonnucleosomal histone-DNA intermediate that is converted into chromatin by ACF. *Mol Cell*, 43(4):638–648, 2011.

Tran H.G., Steger D.J., Iyer V.R. and Johnson A.D. The chromo domain protein chd1p from budding yeast is an ATP-dependent chromatin-modifying factor. *EMBO J*, 19(10):2323–2331, 2000.

Travers A., Caserta M., Churcher M., Hiriart E. and Mauro E.D. Nucleosome positioning–what do we really know? *Mol Biosyst*, 5(12):1582–1592, 2009.

Tremethick D.J. Higher-order structures of chromatin: the elusive 30 nm fiber. *Cell*, 128(4):651–654, 2007.

Trowitzsch S., Bieniossek C., Nie Y., Garzoni F. and Berger I. New baculovirus expression tools for recombinant protein complex production. *J Struct Biol*, 172:45–54, 2010.

Tsukiyama T., Palmer J., Landel C.C., Shiloach J. and Wu C. Characterization of the imitation switch subfamily of ATP-dependent chromatin-remodeling factors in Saccharomyces cerevisiae. *Genes Dev*, 13(6):686–697, 1999.

Tsukiyama T. and Wu C. Purification and properties of an ATP-dependent nucleosome remodeling factor. *Cell*, 83(6):1011–1020, 1995.

Tsukuda T., Fleming A.B., Nickoloff J.A. and Osley M.A. Chromatin remodelling at a DNA double-strand break site in Saccharomyces cerevisiae. *Nature*, 438(7066):379–383, 2005.

Tyler J.K., Adams C.R., Chen S.R., Kobayashi R., Kamakaka R.T. and Kadonaga J.T. The RCAF complex mediates chromatin assembly during DNA replication and repair. *Nature*, 402(6761):555–560, 1999.

Tyler J.K., Collins K.A., Prasad-Sinha J., Amiott E., Bulger M., Harte P.J., Kobayashi R. and Kadonaga J.T. Interaction between the Drosophila CAF-1 and ASF1 chromatin assembly factors. *Mol Cell Biol*, 21(19):6574–6584, 2001.

Uwada J., Tanaka N., Yamaguchi Y., Uchimura Y., Shibahara K.I., Nakao M. and Saitoh H. The p150 subunit of CAF-1 causes association of SUMO2/3 with the DNA replication foci. *Biochem Biophys Res Commun*, 391(1):407–413, 2009.

Varga-Weisz P.D., Wilm M., Bonte E., Dumas K., Mann M. and Becker P.B. Chromatin-remodelling factor CHRAC contains the ATPases ISWI and topoisomerase ii. *Nature*, 388(6642):598–602, 1997.

Vary J.C., Gangaraju V.K., Qin J., Landel C.C., Kooperberg C., Bartholomew B. and Tsukiyama T. Yeast Isw1p forms two separable complexes in vivo. *Mol Cell Biol*, 23(1):80–91, 2003.

Verreault A., Kaufman P.D., Kobayashi R. and Stillman B. Nucleosome assembly by a complex of CAF-1 and acetylated histones H3/H4. *Cell*, 87(1):95–104, 1996.

Verreault A., Kaufman P.D., Kobayashi R. and Stillman B. Nucleosomal DNA regulates the core-histone-binding subunit of the human Hat1 acetyltransferase. *Curr Biol*, 8(2):96–108, 1998.

Wade P.A., Jones P.L., Vermaak D. and Wolffe A.P. A multiple subunit Mi-2 histone deacetylase from Xenopus laevis cofractionates with an associated Snf2 superfamily ATPase. *Curr Biol*, 8(14):843–846, 1998.

Warbrick E. The puzzle of PCNA's many partners. *Bioessays*, 22(11):997–1006, 2000.

Wei Y., Yu L., Bowen J., Gorovsky M.A. and Allis C.D. Phosphorylation of histone H3 is required for proper chromosome condensation and segregation. *Cell*, 97(1):99–109, 1999.

Williams S.K., Truong D. and Tyler J.K. Acetylation in the globular core of histone H3 on lysine-56 promotes chromatin disassembly during transcriptional activation. *Proc Natl Acad Sci U S A*, 105(26):9000–9005, 2008.

Wippo C.J., Israel L., Watanabe S., Hochheimer A., Peterson C.L. and Korber P. The RSC chromatin remodelling enzyme has a unique role in directing the accurate positioning of nucleosomes. *EMBO J*, 30(7):1277–1288, 2011.

Worcel A., Han S. and Wong M.L. Assembly of newly replicated chromatin. *Cell*, 15(3):969–977, 1978.

Wu C., Bassett A. and Travers A. A variable topology for the 30-nm chromatin fibre. *EMBO Rep*, 8(12):1129–1134, 2007.

Xiao A., Li H., Shechter D., Ahn S.H., Fabrizio L.A., Erdjument-Bromage H., Ishibe-Murakami S., Wang B., Tempst P., Hofmann K., Patel D.J., Elledge S.J. and Allis C.D. WSTF regulates the H2A.X DNA damage response via a novel tyrosine kinase activity. *Nature*, 457(7225):57–62, 2009.

Xiao H., Sandaltzopoulos R., Wang H.M., Hamiche A., Ranallo R., Lee K.M., Fu D. and Wu C. Dual functions of largest NURF subunit NURF301 in nucleosome sliding and transcription factor interactions. *Mol Cell*, 8(3):531–543, 2001.

Xu M., Long C., Chen X., Huang C., Chen S. and Zhu B. Partitioning of histone H3-H4 tetramers during DNA replication-dependent chromatin assembly. *Science*, 328(5974):94–98, 2010.

Yamane K., Mizuguchi T., Cui B., Zofall M., ichi Noma K. and Grewal S.I.S. Asf1/HIRA facilitate global histone deacetylation and associate with HP1 to promote nucleosome occupancy at heterochromatic loci. *Mol Cell*, 41(1):56–66, 2011.

Yang Q., Rasmussen S.A. and Friedman J.M. Mortality associated with Down's syndrome in the USA from 1983 to 1997: a population-based study. *Lancet*, 359(9311):1019–1025, 2002.

Yarragudi A., Miyake T., Li R. and Morse R.H. Comparison of ABF1 and RAP1 in chromatin opening and transactivator potentiation in the budding yeast Saccharomyces cerevisiae. *Mol Cell Biol*, 24(20):9152–9164, 2004.

Ye X., Franco A.A., Santos H., Nelson D.M., Kaufman P.D. and Adams P.D. Defective S phase chromatin assembly causes DNA damage, activation of the S phase checkpoint, and S phase arrest. *Mol Cell*, 11(2):341–351, 2003.

Yu L. and Morse R.H. Chromatin opening and transactivator potentiation by RAP1 in Saccharomyces cerevisiae. *Mol Cell Biol*, 19(8):5279–5288, 1999.

Yuan G.C., Liu Y.J., Dion M.F., Slack M.D., Wu L.F., Altschuler S.J. and Rando O.J. Genome-scale identification of nucleosome positions in S. cerevisiae. *Science*, 309(5734):626–630, 2005.

Zeng W., Ball A.R. and Yokomori K. HP1: heterochromatin binding proteins working the genome. *Epigenetics*, 5(4):287–292, 2010.

Zhang Y., Ng H.H., Erdjument-Bromage H., Tempst P., Bird A. and Reinberg D. Analysis of the NuRD subunits reveals a histone deacetylase core complex and a connection with DNA methylation. *Genes Dev*, 13(15):1924–1935, 1999.

Zink D., Sadoni N. and Stelzer E. Visualizing chromatin and chromosomes in living cells. *Methods*, 29(1):42–50, 2003.

Zlatanova J., Seebart C. and Tomschik M. Nap1: taking a closer look at a juggler protein of extraordinary skills. *FASEB J*, 21(7):1294–1310, 2007.

Appendix A

Abbreviations

aa	Amino Acid
Amp	Ampicillin
BME	Beta Mercaptoethanol
BSA	Bovine Serum Albumin
bp	base pair
CAF-1	Chromatin Assembly Factor 1
Cam	Chloramphenicol
CFP	Cyan Fluorescent Protein
CV	Column Volumes
DNAPC	DNA Polymerase Complex
DTT	Dithiothreitol
EDTA	Ethylenediaminetetraacetic acid
EtOH	Ethanol

Gent	Genetamycin
GST	Glutathione S-Transferase
HCMV	Human Cytomegalovirus
HRP	Horseradish Peroxidase
IPTG	Isopropyl Thiogalactoside
Kan	Kanamycin
kDa	Kilodalton
kb	kilobase
MeOH	Methanol
MNase	Micrococcal Nuclease
MOPS	3-(N-morpholino)propanesulfonic acid
NP-40	Nonylphenyl Polyethylene Glycol, NP-40 alternative (Calbiochem)
nt	nucleotides
OD	Optical Density (Apparent absorbance at 600 nm)
p48	RbAp48 (Retinoblastoma Associated protein of 48 kDa)
PCNA	Proliferating Cell Nuclear Antigen
RNAse A	Ribonuclease A
SDS	Sodium Dodecyl Sulphate
SDS-PAGE	Sodium Dodecyl Sulphate - PolyAcrylamide Gel Electrophoresis
SV40	Simian Virus 40
Tet	Tetracyclin

TEV	Tobacco Etch Virus
Tris-HCl	Tris(hydroxymethyl)aminomethane, titrated to pH with HCl
UV	Ultraviolet
XGal	5-bromo-4-chloro-3-indolyl-beta-D-galactopyranoside
YFP	Yellow Fluorescent Protein

Appendix B

Buffers and Solutions

p48-His Buffer	20 mM Tris-HCl, pH 8.0, 150 mM NaCl
3x SDS-PAGE Sample Buffer	40 % (v/v) glycerol, 4 % (w/v) lithium dodecyl sulphate, 4 % (w/v) Ficoll-400, 0.8 M triethanolamine-Cl pH 7.6, 0.025 % (w/v) phenol red, 0.025 % (w/v) coomassie G250, 2 mM EDTA disodium
5x DNA Loading Buffer	0.25 % (w/v) Bromophenol Blue, 0.25 % (w/v) Xylene cyanol FF, 15 % (w/v) Ficoll-400
10x Protease Inhibitor Solution	1 Roche protease inhibitor set tablet (EDTA-free) re-suspended in 5 ml MilliQ H_2O
Alkaline Lysis Buffer	50 mM Glucose, 25 mM Tris-HCl, pH 8.0, 10 mM EDTA
CAF Calmodulin Elution Buffer	50 mM Tris-HCl, pH 8.0, 400 mM KCl, 5 mM EDTA

p48-His Buffer	20 mM Tris-HCl, pH 8.0, 150 mM NaCl
CAF E20/40/100/500	50 mM Tris-HCl, pH 8.0, 100 mM KCl, 5 mM Beta-mercaptoethanol , 20/40/100/250/500 mM Imidazole
CAF GF Buffer	50 mM Tris-HCl, pH 8.0 400 mM KCl
CAF High Salt Buffer	50 mM Tris-HCl, pH 7.5, 1 M KCl, 5 mM Imidazole
CAF IX Buffer A	50 mM Tris-HCl, pH 8.0
CAF IX Buffer B	50 mM Tris-HCl, pH 8.0, 1 M KCl
CAF Lysis Buffer	50 mM Tris-HCl, pH 7.5, 400 mM KCl, 0.5 % (v/v) NP-40
CAF Wash Buffer	50 mM Tris-HCl, pH 7.5, 100 mM KCl, 5 mM Imidazole
Enhancer Solution	0.58 mM para-hydroxy coumarin acid (11 mg in 10 ml 100 % (v/v) nDEMISE)
HA Buffer A	50 mM Tris-HCl, pH 8.0, 200 mM KCl, 10 mM K_2HPO_4
HA Buffer B	50 mM Tris-HCl, pH 8.0, 200 mM KCl, 0.5 M K_2HPO_4
His Pull-down Buffer	50 mM Tris-HCl, pH 8.0, 200 mM KCl, 10 mM Imidazole

p48-His Buffer	20 mM Tris-HCl, pH 8.0, 150 mM NaCl
LB	0.5 % (w/v) Yeast Extract, 1 % (w/v) Tryptone, 1 % (w/v) NaCl (autoclaved)
LB-Agar	LB + 1.5 % (w/v) Agar (autoclaved)
Luminol	0.1 M Tris-HCl, pH 8.6, 1.4 mM luminol
Myc Pull-down Buffer	50 mM Tris-HCl, pH 8.0, 200 mM KCl, 0.1 % (v/v) NP-40
NaOH/SDS Buffer	0.2M NaOH, 1% (w/v) SDS
Nu-PAGE Running Buffer	50 mM MOPS, 50 mM Tris base, 0.1 %(w/v) SDS, 1 mM EDTA, pH 7.7
p48-His Lysis Buffer	20 mM Tris-HCl, pH 8.0, 150 mM NaCl, 10 mM $MgCl_2$, 1 mM Beta-mercaptoethanol
p48-His Dialysis Buffer	10 mM Tris-HCl, pH 8.0, 100 mM NaCl, 2.5 mM $CaCl_2$
p48-His Buffer A	20 mM Tris-HCl, pH 7.5
p48-His Buffer B	20 mM Tris-HCl, pH 7.5, 1 M NaCl
PBS	137 mM NaCl, 2.7 mM KCl, 10 mM NaH_2PO_4, 2 mM KH_2PO_4, pH 7.4
Phenol:CIA Extraction Solution	25:24:1 mixture of Tris-buffered Phenol:Chloroform:Isoamyl Alcohol

p48-His Buffer	20 mM Tris-HCl, pH 8.0, 150 mM NaCl
TE	10 mM Tris-HCl, pH 7.5, 1 mM EDTA
TES Buffer	0.1 M NaCl, 10 mM Tris-HCl, pH 8.0, 1 mM EDTA
TFB1	30 mM potassium acetate, 10 mM calcium chloride, 50 mM manganese chloride, 100 mM rubdinium chloride, 15 % (v/v) glycerol. Adjusted to pH 5.8 with 1 M acetic acid and filter sterilised.
TFB2	10 mM MOPS pH 6.5, 75 mM calcium chloride, 10 mM rubdinium chloride, 15 % (v/v) glycerol. Adjusted to pH 6.5 with 1 M potassium hydroxide and filter sterilised.
Transfer Buffer	20 % (v/v) Methanol, 33 mM Tris-base, 16 mM glycine
50x TAE	2 M Tris-Acetate, 50 mM EDTA

Appendix C

E.coli Genotypes

Cell Line	Genotype	Source
BW23474	F$^-$ *Δ(argF-lac)*169 *rpo*S(Am) *rob*A1 *cre*C510 *hsd*R514 *uid*A(MluI)::*pir*-116, *end*A(BT333) *rec*A1	Gift from I Berger (EMBL, Grenoble, France)
BW23473	F$^-$ Δ(argF-lac)169 *rpo*S(Am) *rob*A1 *cre*C510 *hsd*R514 Δ*uid*A3::pir$^+$ *end*A(BT333) *rec*A1	Gift from I Berger (EMBL, Grenoble, France)
Top10	F$^-$*mcr*A Δ(*mrr-hsd*RMS-*mcr*BC) φ80d*lac*ZΔM15 ΔlacX74 *rec*A1 *ara*D139 Δ(*ara-leu*)7697 *gal*U *gal*K *rps*L (Str^R) *end*A1 *nup*G	Invitrogen
Rosetta2 pLysS	F$^-$*omp*T *hsd*SB($r_B^-m_B^-$) *gal dcm* (DE3) pLysSRARE2 (Cam^R)	Novagen
BacY	F$^-$*mcr*A Δ(*mrr-hsd*RMS-*mcr*BC) φ80d*lac*ZΔM15 ΔlacX74 *rec*A1 *end*A1 *ara*D139 Δ(*ara-leu*)7697 *gal*U *gal*K λ$^-$*rps*L *nup*G/bMON14272/pMON7124 *yfp*$^+$	Gift from I Berger (EMBL, Grenoble, France)

Appendix D

Plasmids Used in this Study

Plasmid	Resistance Gene	Promoter	Tag	Gene(s)	Source
pFL	Amp/Gent	PolH/p10	N/A	N/A	I Berger
pUCDM	Cam	PolH/p10	N/A	N/A	I Berger
pGEXT2	Amp	T7	His	GST 1-232	In House
pNIC-Bsa4	Kan	T7	His	N/A	SGC Oxford
pBAC-Caf1	Gent	PolH/p10	His and Strep	p150 304-938, p60 1-599 and p48 1-425	This Study
pFBDM-p60-His-p150ΔN	Amp/Gent	PolH/p10	His	p150 304-938 and p60 1-599	In House
pFBDM-His-p150ΔN	Amp/Gent	PolH/p10	His	p150 302-955	In House

Plasmid	Resistance Gene	Promoter	Tag	Gene(s)	Source
pNIC-Bsa4-p150_A	Kan	T7	His	p150 436-956	This Study
pNIC-Bsa4-p150_B	Kan	T7	His	p150 465-913	This Study
pNIC-Bsa4-p150_C	Kan	T7	His	p150 545-956	This Study
pNIC-Bsa4-p150_D	Kan	T7	His	p150 545-770	This Study
pGEX-2T-p150_E	Amp	T7	GST	p150 652-732	This Study
pUCDM-Myc	Cam	PolH/p10	Myc	N/A	This Study
pUCDM-Myc-p60_1 (pUCDM-p60_WD)	Cam	PolH/p10	Myc	p60 1-386	This Study
pUCDM-Myc-p60_2	Cam	PolH/p10	Myc	p60 1-400	This Study
pUCDM-Myc-p60_3	Cam	PolH/p10	Myc	p60 1-409	This Study

Plasmid	Resistance Gene	Promoter	Tag	Gene(s)	Source
pUCDM-Myc-p60_4 (pUCDM-p60_FL)	Cam	PolH/p10	Myc	p60 1-559 (FL)	This Study
pUCDM-Myc-p60_5	Cam	PolH/p10	Myc	p60 387-559	This Study
pFBDM-His-p48	Amp/Gent	PolH/p10	His	p48 1-425	In House
pNIC-Bsa4-p150_C RG mutant	Kan	T7	His	p150 545-956 R552G	This Study
pNIC-Bsa4-p150_C KA mutant	Kan	T7	His	p150 545-956 K553A	This Study

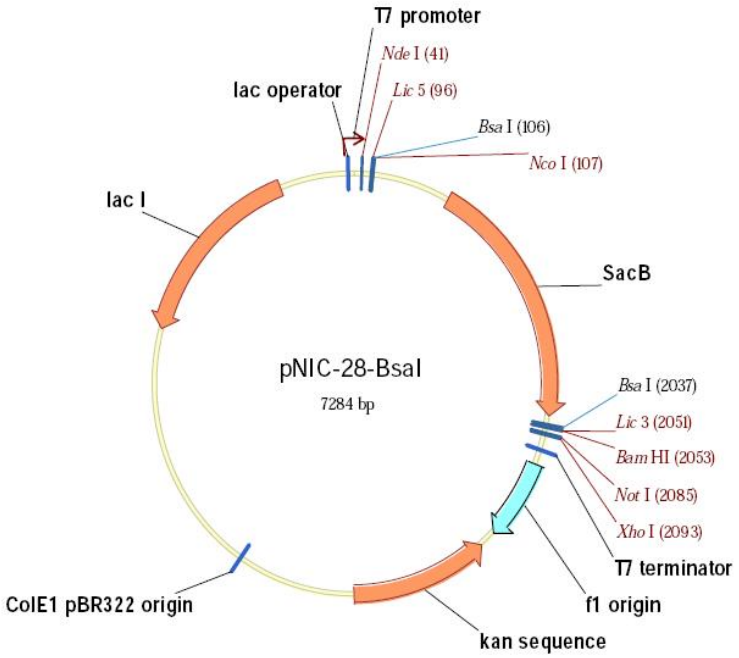


Figure 57: **Map of pNIC-Bsa4**

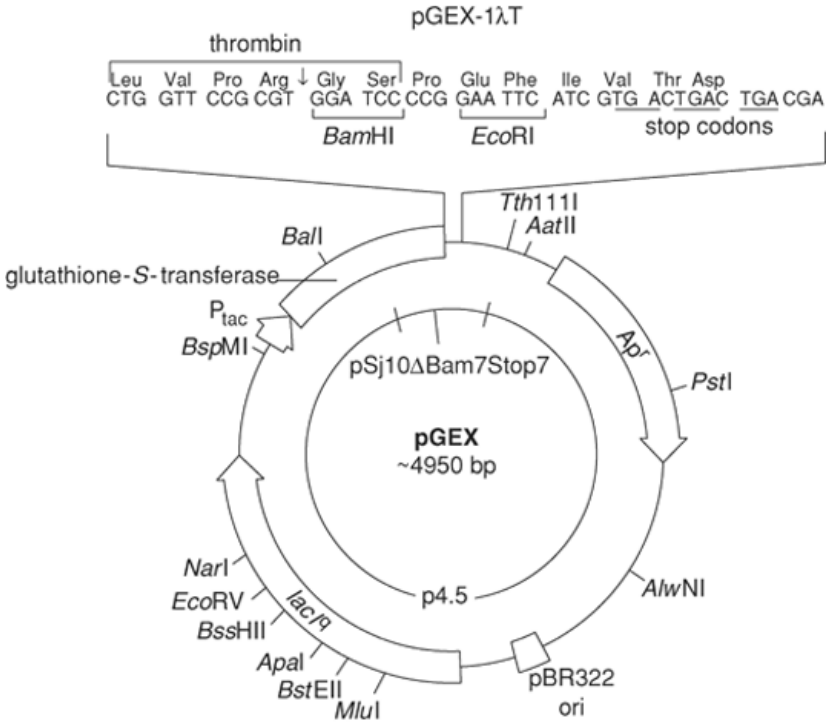


Figure 58: **Map of pGEX-2T**

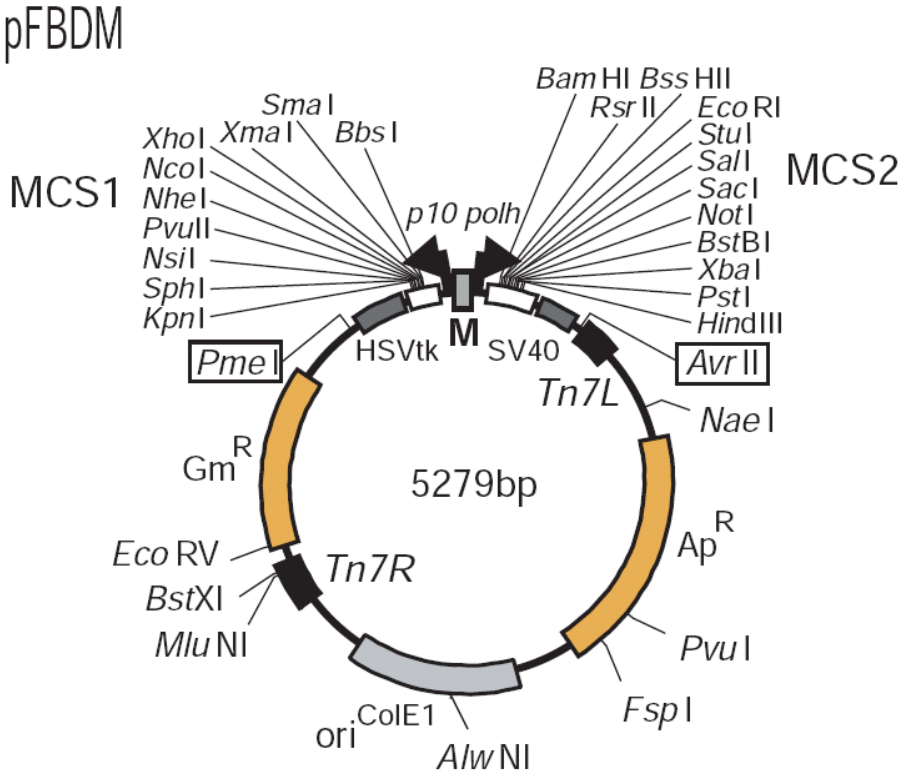


Figure 59: **Map of pFBDM**

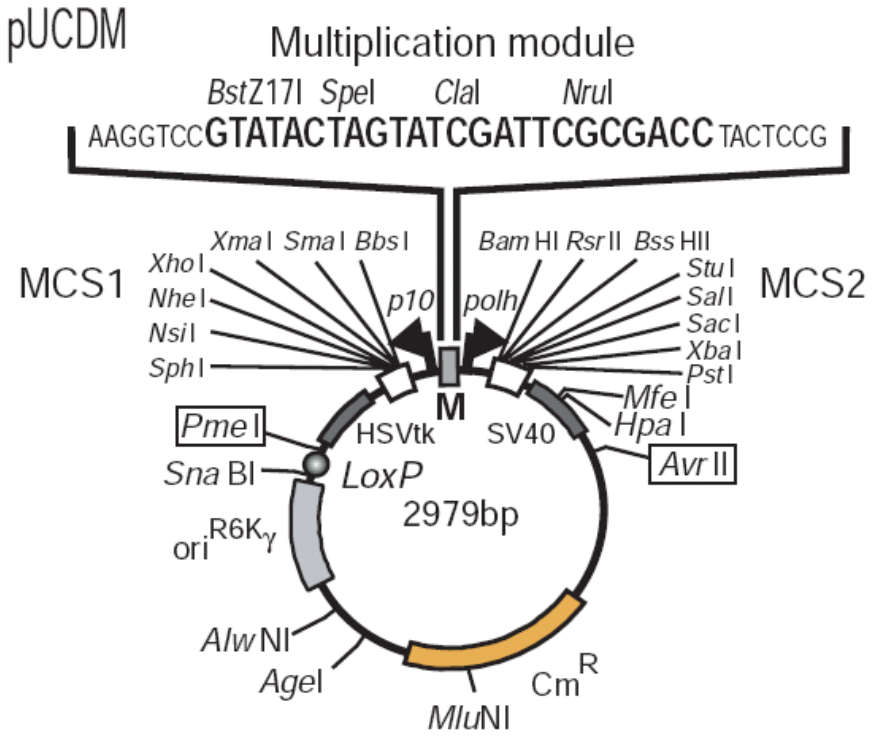


Figure 60: **Map of pUCDM**

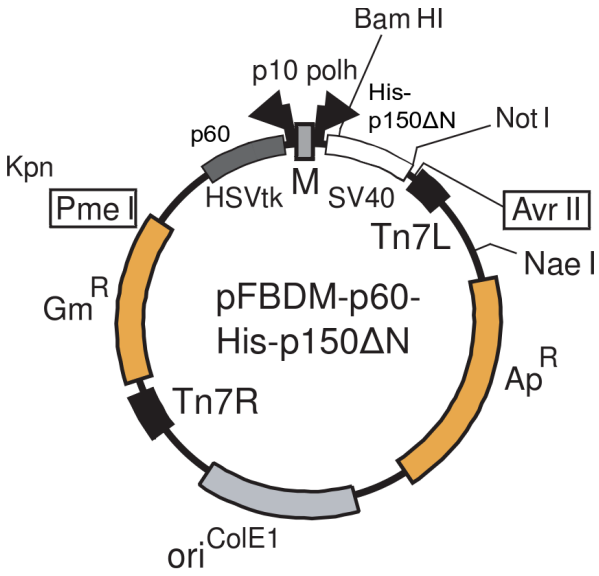


Figure 61: **Map of pFBDM-p60-His-p150ΔN**

Appendix E

Primers

E.1 LIC Primers for E. coli Expression in *BsaI* Digested pNIC-Bsa4

Primer Design

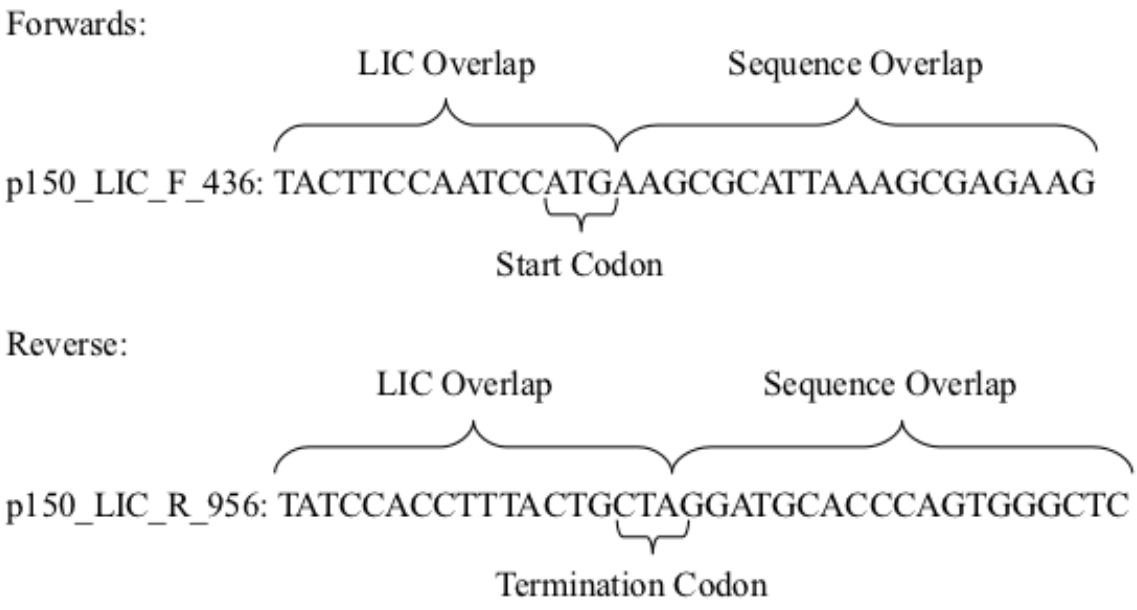
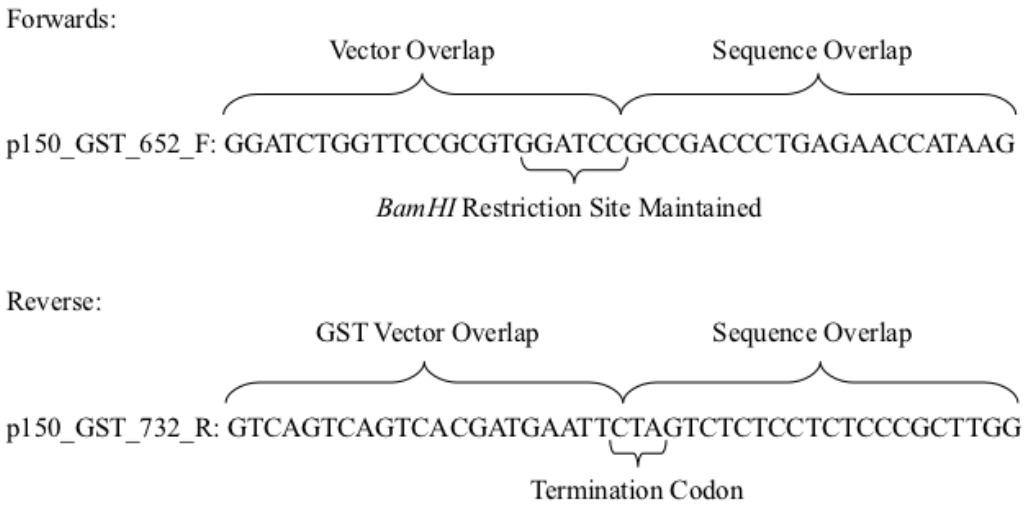


Table of Primers

Primer Name	Sequence
p150_LIC_F_436	TACTTCCAATCCATGAAGCGCATTAAAGCGAGAAG
p150_LIC_F_465	TACTTCCAATCCATGGGGAAGTTTGCCCCCTTG
p150_LIC_F_545	TACTTCCAATCCATGGGCGACGGTGTTCCCGAG
p150_LIC_R_770	TATCCACCTTTACTGCTAGGTGTGGTTGCTGAGCAGTC
p150_LIC_R_912	TATCCACCTTTACTGCTAGTCGCCCTCCTCCTCTTC
p150_LIC_R_956	TATCCACCTTTACTGCTAGGATGCACCCAGTGGGCTC
p60_LIC_F _1	TACTTCCAATCCATGAAAGTCATCACTTGTG
p60_LIC_F _387	TACTTCCAATCCATGAAGCCAGTTTTGAAC
p60_LIC_R _386	TATCCACCTTAACTGTCACTCTTTCAAAGGAATTC
p60_LIC_R _400	TATCCACCTTTACTGTCATTTCTTTGCTGTATC
p60_LIC_R _409	TATCCACCTTTACTGTCAAGACCCTCGATGTGT

E.2 SLIC Cloning into *BamHI*/*EcoRI* Digested pGEX2T



E.3 SLIC Cloning into *BamHI*/*XbaI* Digested pUCDM Using the In-Fusion System

Primer Design

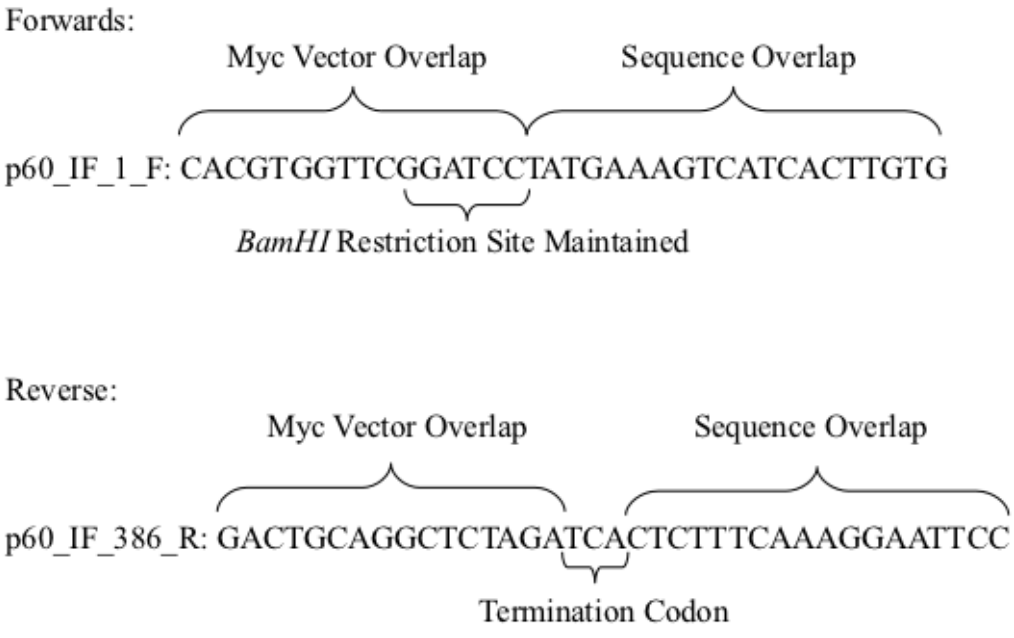
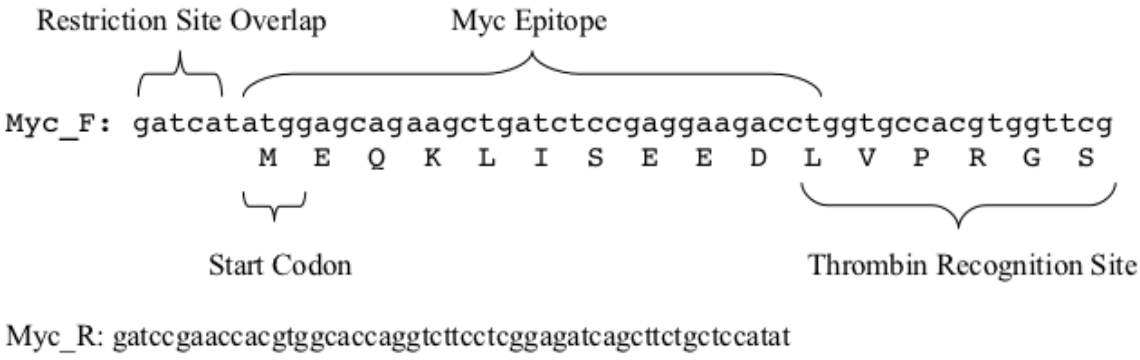


Table of Primers

Primer Name	Sequence
p60_IF_1_F	CACGTGGTTCGGATCCTATGAAAGTCATCACTTGTG
p60_IF_387_F	CACGTGGTTCGGATCCTAAGCCAGTTTTGAACATG
p60_IF_386_R	GACTGCAGGCTCTAGATCACTCTTTCAAAGGAATTCC
p60_IF_400_R	GACTGCAGGCTCTAGATCATTTCTTTGCTGTATCAGG
p60_IF_409_R	GACTGCAGGCTCTAGATCAAGACCCTCGATGTGTCTG
p60_IF_559_R	GACTGCAGGCTCTAGATCAAGGGTCCAGACTTTC

E.4 Oligos for Annealing to Create Myc Insert



Myc_R: gatccgaaccacgtggcaccaggtcttcctcggagatcagcttctgctccatat

E.5 Mutagenesis Primers for p150_C

Mutated base pairs are in bold and underlined.

Name	Sequence
p150C_R552G_F	GACGGTGTTCCCGAGAGG**G**GGAAGTTTGGCAGGATG
p150C_R552G_R	CATCCTGCCAAACTTCC**C**CCTCTCGGGAACACCGTC
p150C_K553A_F	GTTCCCGAGAGGAGG**GC**GTTTGGCAGGATGAAG
p150C_K553A_R	CTTCATCCTGCCAAAC**GC**CCTCCTCTCGGGAAC
p150C_RK552GA_F	GCGACGGTGTTCCCGAGAGG**G**GG**GC**GTTTGGCAGGATGAAGCTCC
p150C_RK552GA_R	GGAGCTTCATCCTGCCAAA**CG**CC**C**CCTCTCGGGAACACCGTCGC

Appendix F

Sequences

p150△N (305-938) (*Homo sapiens*, NCBI Q13111)

MPPKQHSSTS PFPTSTPLRR ITKKFVKGST EKNKLRLQRD QERLGKQLKL RAEREEKEKL
KEEAKRAKEE AKKKKEEEKE LKEKERREKR EKDEKEKAEK QRLKEERRKE RQEALEAKLE
EKRKKEEEKR LREEEKRIKA EKAEITRFFQ KPKTPQAPKT LAGSCGKFAP FEIKEHMVLA
PRRRTAFHPD LCSQLDQLLQ QQSGEFSFLK DLKGRQPLRS GPTHVSTRNA DIFNSDVVIV
ERGKGDGVPE RRKFGRMKLL QFCENHRPAY WGTWNKKTAL IRARDPWAQD TKLLDYEVDS
DEEWEEEEPG ESLSHSEGDD DDDMGEDEDE DDGFFVPHGY LSEDEGVTEE CADPENHKVR
QKLKAKEWDE FLAKGKRFRV LQPVKIGCVW AADRDCAGDD LKVLQQFAAC FLETLPAQEE
QTPKASKRER RDEQILAQLL PLLHGNVNGS KVIIREFQEH CRRGLLSNHT GSPRSPSTTY
LHTPTPSEDA AIPSKSRLKR LISENSVYEK RPDFRMCWYV HPQVLQSFQQ EHLPVPCQWS
YVTSVPSAPK EDSGSVPSTG PSQGTPISLK RKSAGSMCIT QFMKKRRHDG QIGAEDMDGF
QADTEEEEEE EGDCMIVDVP DAAEVQAPCG AASGAGGGVG VDTGKATLTA SPLGAS

p60 (*Homo sapiens*, NCBI Q13112)

MKVITSEIAW HNKEPVYSLD FQHGTAGRIH RLASAGVDTN VRIWKVEKGP DGKAIVEFLS
NLARHTKAVN VVRFSPTGEI LASGGDDAVI LLWKVNDNKE PEQIAFQDED EAQLNKENWT
VVKTLRGHLE DVYDICWATD GNLMASASVD NTAIIWDVSK GQKISIFNEH KSYVQGVTWD
PLGQYVATLS CDRVLRVYSI QKKRVAFNVS KMLSGIGAEG EARSYRMFHD DSMKSFFRRL
SFTPDGSLLL TPAGCVESGE NVMNTTYVFS RKNLKRPIAH LPCPGKATLA VRCCPVYFEL
RPVVETGVEL MSLPYRLVFA VASEDSVLLY DTQQSFPFGY VSNIHYHTLS DISWSSDGAF
LAISSTDGYC SFVTFEKDEL GIPLKEKPVL NMRTPDTAKK TKSQTHRGSS PGPRPVEGTP
ASRTQDPSSP GTTPPQARQA PAPTVIRDPP SITPAVKSPL PGPSEEKTLQ PSSQNTKAHP
SRRVTLNTLQ AWSKTTPRRI NLTPLKTDTP PSSVPTSVIS TPSTEEIQSE TPGDAQGSPP
ELKRPRLDEN KGGTESLDP

p48 (*Homo sapiens*, NCBI Q09028)

MADKEAAFDD AVEERVINEE YKIWKKNTPF LYDLVMTHAL EWPSLTAQWL PDVTRPEGKD
FSIHRLVLGT HTSDEQNHLV IASVQLPNDD AQFDASHYDS EKGEFGGFGS VSGKIEIEIK
INHEGEVNRA RYMPQNPCII ATKTPSSDVL VFDYTKHPSK PDPSGECNPD LRLRGHQKEG
YGLSWNPNLS GHLLSASDDH TICLWDISAV PKEGKVVDAK TIFTGHTAVV EDVSWHLLHE
SLFGSVADDQ KLMIWDTRSN NTSKPSHSVD AHTAEVNCLS FNPYSEFILA TGSADKTVAL
WDLRNLKLKL HSFESHKDEI FQVQWSPHNE TILASSGTDR RLNVWDLSKI GEEQSPEDAE
DGPPELLFIH GGHTAKISDF SWNPNEPWVI CSVSEDNIMQ VWQMAENIYN DEDPEGSVDP
EGQGS

CAF-1 Poly Protein (cleaved)

TEV Protease

MGMGHHHHHH GESLFKGPRD YNPISSTICH LTNESDGHTT SLYGIGFGPF IITNKHLFRR
NNGTLLVQSL HGVFKVKNTT TLQQHLIDGR DMIIIRMPKD FPPFPQKLKF REPQREERIC
LVTTNFQTKS MSSMVSDTSC TFPSSDGIFW KHWIQTKDGQ CGSPLVSTRD GFIVGIHSAS
NFTNTNNYFT SVPKNFMELL TNQEAQQWVS GWRLNADSVL WGGHKVFMSK PEEPFQPVKE
ATQLMNELVY SQGNHENLYF Q

p150

GVHHHHHHHH HHGPLVPRGS QDPSKQHSST SPFPTSTPLR RITKKFVKGS TEKNKLRLQR
DQERLGKQLK LRAEREEKEK LKEEAKRAKE EAKKKKEEEK ELKEKERREK REKDEKEKAE
KQRLKEERRK ERQEALEAKL EEKRKKEEEK RLREEEKRIK AEKAEITRFF QKPKTPQAPK
TLAGSCGKFA PFEIKEHMVL APRRRTAFHP DLCSQLDQLL QQQSGEFSFL KDLKGRQPLR
SGPTHVSTRN ADIFNSDVVI VERGKGDGVP ERRKFGRMKL LQFCENHRPA YWGTWNKKTA
LIRARDPWAQ DTKLLDYEVD SDEEWEEEEP GESLSHSEGD DDDDMGEDED EDDGFFVPHG
YLSEDEGVTE ECADPENHKV RQKLKAKEWD EFLAKGKRFR VLQPVKIGCV WAADRDCAGD
DLKVLQQFAA CFLETLPAQE EQTPKASKRE RRDEQILAQL LPLLHGNVNG SKVIIREFQE
HCRRGLLSNH TGSPRTPSTT YLHTPTPSED AAIPSKSRLK RLISENSVYE KRPDFRMCWY
VHPQVLQSFQ QEHLPVPCQW SYVTSVPSAP KEDSGSVPST GPSQGTPISL KRKSAGSMCI
TQFMKKRRHD GQIGAEDMDG FQADTEEEEE EEGDCMIVDV PDAVEVQAPC GAASGAGGGV
GVDTGKATLT ASPLGASENL YFQ

p60

GKVITCEIAW HNKEPVYSLD FQHGTAGRIH RLASAGVDTN VRIWKVEKGP DGKAIVEFLS
NLARHTKAVN VVRFSPTGEI LASGGDDAVI LLWKVNDNKE PEQIAFQDED EAQLNKENWT

VVKTLRGHLE DVYDICWATD GNLMASASVD NTAIIWDVSK GQKISIFNEH KSYVQGVTWD
PLGQYVATLS CDRVLRVYSI QKKRVAFNVS KMLSGIGAEG EARSYRMFHD DSMKSFFRRL
SFTPDGSLLL TPAGCVESGE NVMNTTYVFS RKNLKRPIAH LPCPGKATLA VRCCPVYFEL
RPVVETGVEL MSLPYRLVFA VASEDSVLLY DTQQSFPFGY VSNIHYHTLS DISWSSDGAF
LAISSTDGYC SFVTFEKDEL GIPLKEKPVL NMRTPDTAKK TKSQTHRGSS PGPRPVEGTP
ASRTQDPSSP GTTPPQARQA PAPTVIRDPP SITPAVKSPL PGPSEEKTLQ PSSQNTKAHP
SRRVTLNTLQ AWSKTTPRRI NLTPLKTDTP PSSVPTSVIS TPSTEEIQSE TPGDAQGSPP
ELKRPRLDEN KGGTESLDPE NLYFQ

p48

GMKRRWKKNF IAVSAANRFK KISSSGLVPR GSADKEAAFD DAVEERVINE EYKIWKKNTP
FLYDLVMTHA LEWPSLTAQW LPDVTRPEGK DFSIHRLVLG THTSDEQNHL VIASVQLPND
DAQFDASHYD SEKGEFGGFG SVSGKIEIEI KINHEGEVNR ARYMPQNPCI IATKTPSSDV
LVFDYTKHPS KPDPSGECNP DLRLRGHQKE GYGLSWNPNL SGHLLSASDD HTICLWDISA
VPKEGKVVDA KTIFTGHTAV VEDVSWHLLH ESLFGSVADD QKLMIWDTRS NNTSKPSHSV
DAHTAEVNCL SFNPYSEFIL ATGSADKTVA LWDLRNLKLK LHSFESHKDE IFQVQWSPHN
ETILASSGTD RRLNVWDLSK IGEEQSPEDA EDGPPELLFI HGGHTAKISD FSWNPNEPWV
ICSVSEDNIM QVWQMAENIY NDEDPEGSVD PEGQGSRSGG ENLYFQG

CFP

GVSKGEELFT GVVPILVELD GDVNGHKFSV SGEGEGDATY GKLTLKFICT TGKLPVPWPT
LVTTLTWGVQ CFSRYPDHMK QHDFFKSAMP EGYVQERTIF FKDDGNYKTR AEVKFEGDTL
VNRIELKGID FKEDGNILGH KLEYNYISHN VYITADKQKN GIKANFKIRH NIEDGSVQLA
DHYQQNTPIG DGPVLLPDNH YLSTQSALSK DPNEKRDHMV LLEFVTAAGI TLGMDELYKS
GLRSRRAH