Role of SWI/SNF in regulating pre-mRNA processing in *Drosophila melanogaster*

Funktion von SWI/SNF in der Regulation der prä-mRNA-Prozessierung in *Drosophila melanogaster*

**Doctoral thesis for a doctoral degree
at the Graduate School of Life Sciences,
Julius-Maximilians-Universität Würzburg,
Section Biology
Submitted by**

**Anu Tyagi
From
Ghaziabad, India**

Würzburg, 2012

Submitted on:

Members of the *Promotionskomitee*:

Chairperson: **Prof. Dr. Peter Jakob**

Primary Supervisor: **Prof. Dr. Utz Fischer**

Supervisor (Second):   **Prof. Dr. Caroline Kisker**

Supervisor (Third):   **Prof. Dr. Alexander Buchberger**

Date of Public Defense:

Date of Receipt of Certificate:

Table of Contents

1.Summary

ATP dependent chromatin remodeling complexes are multifactorial complexes that utilize the energy of ATP to rearrange the chromatin structure. The changes in chromatin structure lead to either increased or decreased DNA accessibility. SWI/SNF is one of such complex. The SWI/SNF complex is involved in both transcription activation and transcription repression. The ATPase subunit of SWI/SNF is called SWI2/SNF2 in yeast and Brahma, Brm, in *Drosophila melanogaster.* In mammals there are two paralogs of the ATPase subunit, Brm and Brg1. Recent studies have shown that the human Brm is involved in the regulation of alternative splicing.

The aim of this study was to investigate the role of Brm in pre-mRNA processing. The model systems used were *Chironomus tentans*, well suited for *in situ* studies and *D. melanogaster*, known for its full genome information.

Immunofluorescent staining of the polytene chromosome indicated that Brm protein of *C. tentans*, ctBrm, is associated with several gene loci including the Balbiani ring (BR) puffs. Mapping the distribution of ctBrm along the BR genes by both immuno-electron microscopy and chromatin immunoprecipitation showed that ctBrm is widely distributed along the BR genes. The results also show that a fraction of ctBrm is associated with the nascent BR pre-mRNP. Biochemical fractionation experiments confirmed the association of Brm with the RNP fractions, not only in *C. tentans* but also in *D. melanogaster* and in HeLa cells. Microarray hybridization experiments performed on S2 cells depleted of either dBrm or other SWI/SNF subunits show that Brm affects alternative splicing and 3′ end formation. These results indicated that BRM affects pre-mRNA processing as a component of SWI/SNF complexes.

2. Zusammenfassung

ATP abhängige Chromatin Remodelling Komplexe bestehen aus diversen Faktoren, welche die bei der Umsetzung von ATP freiwerdende Energie dazu nutzen, die Chromatinstruktur neu zu ordnen. Diese Veränderungen führen zu einer Zu- bzw. Abnahme in der Zugänglichkeit der DNA. Ein Beispiel dafür ist der SWI/SNF-Komplex, der sowohl in die Aktivierung als auch die Inhibierung der Transkription involviert ist. Die ATPase-Untereinheit von SWI/SNF heißt in Hefe SWI2/SNF2 und in *Drosophila melanogaster* Brahma (Brm). Im Gegensatz dazu besitzen Säuger zwei Paraloge der ATPase-Einheit, nämlich Brm und Brg1. Neueste Studien haben gezeigt, dass das humane Brm in der Regulation des Alternativen Spleißen beteiligt ist.

Ziel dieser Arbeit ist es, die Rolle von Brm in der prä-mRNA-Prozessierung zu untersuchen. Als Versuchssysteme wurden *Chironomus tentans* und *D. melanogaster* herangezogen. Dabei eignete sich *C. tentans* vor allem für die *in situ* Studien während bei *D. melanogaster* das vollständig sequenzierte Genom von Vorteil war.

Immunfluoreszenzfärbungen von Polytän-Chromosomen zeigen eine Assoziation von Brm von *C. tentans*, ctBrm; mit unterschiedlichen Genloci, einschließlich der Balbiani-Ringe (BR). Mit Hilfe von Immun-Elektronenmikroskopie und Chromatin-Immunpräzipitation (ChIP) wird die Verteilung von ctBrm entlang der BR-Gene untersucht. Dabei zeigt ctBrm eine weite Streuung. Die Ergebnisse lassen außerdem darauf schließen, dass ein Teil des ctBrm-Proteins mit naszierenden BR-prä-mRNPs interagiert. Biochemische Fraktionierungs-experimente bestätigen die Assoziation von Brm mit RNP-Fraktionen nicht nur in *C. tentans*, sondern auch in *D. melanogaster* und in HeLa-Zellen. Microarray-Untersuchungen in S2-Zellen, in denen entweder dBrm oder eine andere Untereinheit von SWI/SNF depletiert war, zeigen, dass BRM als eine Komponente des SWI/SNF-Komplexes sowohl Alternatives Spleißen und die Formierung des 3´ Endes, als auch die prä-mRNA-Prozessierung beeinflusst.

3. List of Publications:

Publications included in the thesis:

1. Botelho SC, **Tyagi A**, Hessle V, Farrants AK, and Visa N (2008). The association of Brahma with the Balbiani ring 1 gene of Chironomus tentans studied by immunoelectron microscopy and chromatin immunoprecipitation. *Insect Mol Biol*. 17 :505-513.

2. **Tyagi A**, Ryme J, Brodin D, Östlund Farrants AK and Visa N (2009). SWI/SNF Associates with Nascent Pre-mRNPs and Regulates Alternative Pre-mRNA Processing. *PLoS Genet*. 5:e1000470.

Other publications:

3. Waldholm J, Wang Z, Brodin D, **Tyagi A**, Yu S, Theopold U, Farrants AK and Visa N (2011). SWI/SNF regulates the alternative processing of a specific subset of pre-mRNAs in Drosophila melanogaster. *BMC Mol Biol*. 12: 46.

4. Hessle V, Björk P, Sokolowski M, González de Valdivia E, Silverstein R, Artemenko K, **Tyagi A**, Maddalo G, Ilag L, Helbig R, Zubarev RA and Visa N. (2009). The exosome associates cotranscriptionally with the nascent pre-mRNP through interactions with heterogeneous nuclear ribonucleoproteins. *Mol Biol Cell.* 15: 3459-3470.

Project I

4. Introduction

Gene expression is a complex process in which the information stored in the gene is used to make a functional gene product. The gene product can be a protein or a non-protein product such as ribosomal RNA (rRNA), transfer RNA (tRNA) etc. (reviewed by Jaenisch and Bird, 2003). The first step in gene expression is the transcription of a gene into RNA. To initiate transcription of the protein coding genes RNA polymerase binds to the promoter of the gene and transcribes into mRNA. Once few nucleotides of mRNA are synthesised a methylated guanosine (m^7G) 'cap' is added to the 5´ end of the transcribed RNA. Splicing of the pre-mRNA occurs by stepwise cleavage and ligation events to remove intron sequences and bring exons together. Finally, after splicing, the 3´ end of the mRNA is cleaved and a series of adenosine residues, known as a poly(A) tail, is added. After polyadenylation the mRNA is ready for transport from the nucleus to the cytoplasm (reviewed by Jaenisch and Bird 2003). Once in the cytoplasm the mRNA is ready for translation. During translation polypeptides are synthesized in three steps: initiation, elongation and termination. After initiation, the genetic codes specified by the mRNA are read and the encoded amino acids are incorporated into a growing polypeptide chain during the elongation process. Termination of translation occurs when one of the termination codons (UAG, UAA and UGA) signals the release of a completed polypeptide chain. Ribosome releases the mRNA and the ribosomal subunits dissociate for the next cycle. The newly synthesized protein undergoes posttranslational modifications so that it can be used for its dedicated role. From the start of gene expression to the end there are several levels of regulation of gene expression (reviewed by Gibney and Nolan, 2010)

All cells in an organism contain the same DNA. Cell types and functions differ because of differences in their gene expression patterns. This is why the differentiated cells in multi-cellular organism express only a subset of genes that are necessary for their activity. The regulation of gene activity is called gene expression regulation. The control of gene expression is the root of differentiation and development. The patterns of gene expression are established during development

and are maintained as the cells divide by mitosis. The environment can also influence the patterns of gene expression. (Woodcock and Ghosh, 2010).

4.1 Gene expression regulation

Gene expression regulation is a cellular process that controls the rate and manner of the gene expression. A complex set of interactions between genes, RNA molecules, proteins (including transcription factors), other components of the expression system and the gene product itself determine the activation or repression of the genes (reviewed by Mellor *et al*., 2008 and Muchardt, 2010). Some genes are expressed continuously as they produce proteins involved in basic metabolic functions, others are expressed as part of the process of cell differentiation and few are expressed as a result of cell differentiation. The gene expression can be regulated at multiple levels. First, the transcriptional regulation, which includes regulation of gene expression by structure of chromatin, regulation of transcription initiation and regulation by transcription factors. Second, by regulating the processing of mRNA molecules, including alternative splicing. In some instances alternative splicing results in production of more than one protein forms from a single gene. In other cases alternative splicing produces multiple mRNAs with different 3′ untranslated regions (UTR). The second regulation mechanism also includes regulating the stability of mRNA molecules. The third level includes regulation of translation, translation initiation and posttranslational modifications of the proteins.

Lately, the understanding of the influence of chromatin on gene regulation suggests that DNA methylation and histone post-translational modifications lead to the recruitment of protein complexes that regulate transcription. Recent studies suggest that transcription occurs when favored by mixtures of complex modifications, which probably have several roles (reviewed by Berger, 2007). In most of the cases chromatin remodeling has major influence on transcription regulation.

4.2 Chromatin and Nucleosomes

In a eukaryotic cell, chromatin is packed into the nucleus, which is only few μm in diameter. As the chromatin contains several hundred centimeters of DNA, efficient and compact packing of DNA is crucial. Chromatin is a complex of DNA and DNA associated proteins. The major group of proteins associated with the DNA is histones.

The histones together with DNA form the smallest subunit of chromatin called the nucleosome. The nucleosomes contain a core of eight histones with the DNA wrapped around them. This arrangement gives the appearance as the beads on a string (see Figure1). The DNA in between the nucleosomes is called the linker DNA and the DNA in tight association with nucleosomes is called the core DNA. About 147 base pairs of DNA are wrapped around the histone core of the nucleosome (reviewed by Maeshima *et al*., 2010).

There are five types of histones in a eukaryotic cell: H1, H2A, H2B, H3 and H4. Histones H2A, H2B, H3 and H4 are the core histones present in the nucleosome. Histone H1 is called the linker histone as it binds to the linker DNA, the DNA between the two nucleosomes. Histones are positively charged proteins. More than 20% of the amino acid residues are either lysine or arginine. The histones are small proteins with the molecular weight ranging between 11-20 kDa. The assembly of nucleosomes involves the ordered association of histones with DNA. First the H3:H4 tetramer binds to the DNA and then two dimers of H2A and H2B join to form the final nucleosome. All the core histones have N-terminal tails protruding out of the nucleosomes. These histone tails do not have any defined structure and are highly accessible for interacting proteins. Nucleosomes appear to be the primary determinants of chromatin structure in eukaryotes. This is due to nucleosome positioning on the DNA strand and also because the chemical nature of the histone proteins is the major factor that determines the degree of DNA packaging (reviewed by Maeshima *et al*., 2010).

After the formation of nucleosomes, the DNA is packed into a more compact structure and bound to the linker histone H1. The binding of histone H1 stabilizes high order chromatin structure. The N-terminal tails of the core histones also play an important role in the stabilization of chromatin structure. Chromatin structure is hierarchic, ranging from the two lowest levels of DNA packaging the nucleosome and the 30 nm chromatin fiber to metaphase chromosomes, which represent the most compact form of chromatin in eukaryotes. Light and dark areas of chromatin are observed within the nucleus when non-dividing cells are examined by phase contrast microscopy. The dark areas are called heterochromatin and contain DNA that is in compact organization, whereas the light areas are transcriptionally active regions called euchromatin. Two types of heterochromatin are recognized: constitutive and

facultative heterochromatin. Constitutive heterochromatin represents DNA that does not contain any genes and can always be maintained in a compact organization. This includes centromeric and telomeric DNA as well as certain regions of some other chromosomes. For example, most of the human Y chromosome is made of constitutive heterochromatin. Facultative heterochromatin is not a permanent state of chromatin. Facultative heterochromatin contains genes that are inactive in some cells or at some periods in cell cycle. When these genes are inactive, their DNA regions are compacted into heterochromatin (reviewed by Alkhatib and Landry, 2011).
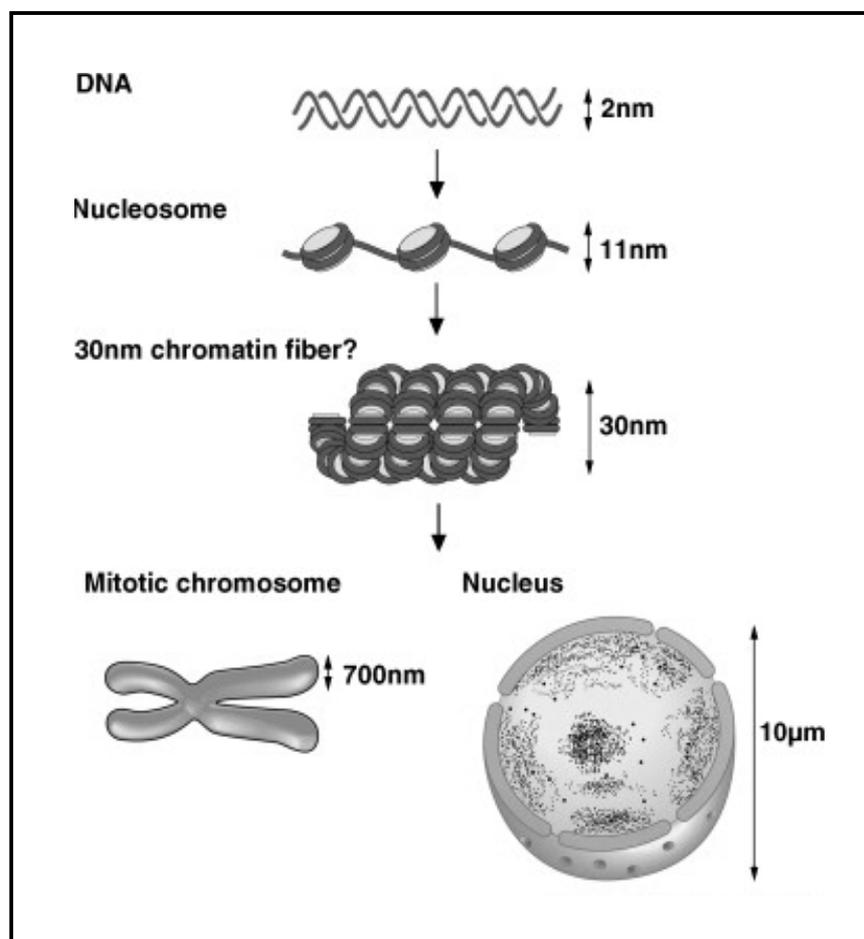
Figure 1. Different stages of chromatin packaging. A long stretch of DNA with 2nm diameter is wrapped around histone octamer. The histone octamer consists of histone H2A, H2B, H3 and H4. The DNA wrapped around histone octamer makes chromatin structure with 11nm diameter, which again folds to make 30nm chromatin fiber. The 30nm chromatin fiber further condenses and finally organizes to make mitotic chromosome. (Modified from Maeshima *et al*., 2010)

The chromatin structure is dynamic. The euchromatic and heterochromatic states of chromatin are interchangeable to some extent. The compact packaging of DNA presents a big obstacle for transcription, replication, recombination and repair processes. To facilitate these events the chromatin structure should be loose enough to be accessible to the factors responsible for each process and after the event the nucleosomal organization should be maintained (reviewed by Mellor, 2006 and Woodcock and Ghosh 2010). The interconversion between the accessible and the non-accessible state of chromatin is called chromatin remodeling.

4.3 Chromatin remodeling

The term chromatin remodeling is used to describe a wide variety of changes in chromatin structure and is generally defined as any activity that generates detectable changes in the histone-DNA interaction. Chromatin remodeling is separated into two main categories: Covalent chromatin remodeling and ATP dependent chromatin remodeling.

4.3.1 Covalent histone modifications

Histone modifications include addition or removal of various molecular moieties on histone proteins. Acetylation (Carrozza *et al*., 2003 and Thiagalingam *et al*., 2003), methylation (Lachner and Jenuwein 2002 and Sims *et al*., 2003), phosphorylation (Iizuka and Smith, 2003) and ubiquitination (Spencer and Davie 1999) are the covalent modifications that have been well studied. Other modifications have also been identified such as sumoylation, ADP ribosylation (Hassa *et al*., 2006), deimination (Cuthbert *et al*., 2004) and proline isomerization (Nelson *et al*., 2006). These modifications are now under extensive study to understand their biological function. The effect of modified histones can be very local, acting on nucleosomes at individual promoters, or can affect the organization of the whole chromosome (Gregory *et al*., 2001 and Zhang and Reinberg, 2001).

Histones can be modified at several sites. Up to date more than 60 different residues on histones that can be modified have been identified, either by specific antibodies or by mass spectrometry (reviewed by Kouzarides, 2007). There are two well characterized mechanisms for the functions of histone modifications. The first

mechanism is to break the histone-DNA interaction revealing hidden DNA sequences to make the DNA more accessible. The second mechanism is the recruitment of non-histone proteins, for instance ATP dependant chromatin remodeling factors. A specific set of proteins is favored to bind or tois removed from chromatin depending on the modifications on a given histone (reviewed by Kouzarides, 2007). The modifications affect the higher order structure of chromatin.

4.3.2 Histone modifying enzymes

Enzymes that direct the modifications of histone residues have been focus of study for many years. There are several enzymes identified for acetylation (Sterner and Berger, 2000), methylation (Zhang and Reinberg, 2001), phosporylation (Nowak and Corces, 2004), ubiquitination (Shilatifard, 2006), sumoylation (Shilatifard, 2006), ADP-ribosylation (Hassa *et al*., 2006), deimination (Cuthbert *et al*., 2004) and proline isomerization (Nelson *et al*., 2006). Most of the modifications are reversible, and enzymes that remove the modification have been identified (Bannister and Kouzarides 2011). Among all the histone modifying enzymes, methyltransferases and kinases are the most specific enzymes. In some cases, the specificity of the enzymes that modify histones can be influenced by other factors. The complexes in which the enzymes are found may specify a preference for nucleosomal or free histones (Lee *et al*., 2005). Proteins associated with the enzyme affect the selection of the residue to be modified (Metzger et al., 2005).

4.3.3 Biological functions of histone modifications

The function of histone modifications is to establish a euchromatic or heterochromatic environment and to direct the DNA-based biological events. The heterochromatin and euchromatin environments are associated with distinct set of histone modifications. In mammals, the heterochromatin is associated with low levels of acetylation and high levels of certain methylated residues such as H3K9, H3K27 and H4K20. Whereas, euchromatin regions are hyperacetylated and contain several combinations of histone modifications. The individual role of each modification is under study. Acetylation is the type of histone modification that is associated with activation of gene transcription whereas, deacetylation is associated with

transcriptional repression. Phosphorylation of histone residues is also an important modification. It has been observed that extensive cross talk exists between phosphorylation and other posttranslational modifications of histones, which together regulate a specific biological process, including gene transcription, DNA repair, and cell cycle progression. Recently it has been demonstrated that nearly all histone types are phosphorylated at specific residues. These modifications act as a critical intermediate step in chromosome condensation during cell division, transcriptional regulation, and DNA repair (reviewed by Banerjee and Chakravarti 2011). Evidences suggest that phosphorylation of histone does not act as general code, instead act in combination with other histone modifications as a platform for recruitment or release of effector proteins, leading to downstream sequence of events. Methylation of lysine residues has variable affect on gene expression (Bannister and Kouzarides, 2005). For several years it was believed that methylation was a stable modification but the discovery of the histone demethylase LSD1 (Shi *et al.*, 2004) and other demethylases revealed that this modification is also dynamic (reviewed by Kouzarides 2007). The presence of mono-, di- and tri-methylation presents a large functional diversity of lysine methylation (reviewed by Kouzarides, 2007).

Like lysine methylation, arginine methylation is either associated with activation or with repression of transcription. Deimination is the conversion of arginine to citrulline. It has the capacity to counteract the activating effect of arginine methylation since citrulline prevents arginines from being methylated (Cuthbert et al., 2004). Ubiquitination is a large modification that has been found on H2A and H2B. The function of ubiquitylation is unclear but it has been indicated that both addition and removal of ubiquitin is important for transcription activation (Emre et al., 2005 and Gardner et al., 2005). Sumoylation is also a large modification found on all core histones. This modification is thought to contribute in making the chromatin environment repressive for transcription (Nathan et al., 2006). ADP ribosylation is a less studied modification with respect to its function. Histone proteins can be mono- and/or poly-ADP-ribosylated (Hassa *et al.*, 2006). ADP-ribosylation of nucleosomes is involved in the regulation of higher order chromatin structure. ADP-ribosylated histone H1 has been shown to cause chromatin unwinding (reviewed by Messner and Hottiger 2011). Recent research in the field of alternative splicing regulation has revealed that histone modifications also play an important role in the selection of

alternatively spliced exons in alternatively spliced genes (reviewed by Luco and Mistelli, 2011) (for details see section 4.7 Coupling between mRNA processing steps).

4.4 ATP dependent chromatin remodeling

Covalent chromatin remodeling presents a major contribution to the state of chromatin in the nucleus but it is just not sufficient enough to displace nucleosomes. Therefore, the energy of chromatin remodeling complexes is required to displace single or multiple core histones (reviewed by Johnson *et al.*, 2005). The chromatin remodeling complexes utilize energy of ATP hydrolysis and use chromatin as a substrate for processes involving DNA, such as the expression of genes, the replication of the genome, repair of DNA damage and the recombination of chromosomes. Controlled energy dependent nucleosome remodeling is important for cell proliferation and differentiation. Most importantly, 'remodeling' of nucleosomes increases the accessibility of DNA sequences to regulatory proteins that scan the genome for target sites. *In vivo*, the action of ATP dependent chromatin remodeling complexes leads to changes in chromatin structure, ranging from the complete absence of nucleosomes at regulatory sites (Reinke and Horz, 2003) to shifting nucleosome positions (Belikov *et al.*, 2001), increasing the access of DNA on the surface of positioned nucleosomes (Truss *et al.*, 1995) and exchange of H2A variants (Krogan *et al.*, 2003).

The chromatin remodeling complexes are multi-subunit protein complexes. All nucleosome remodeling ATPases belong to the SNF2 family of ATPases (Eisen *et al.*, 1995 and Lusser and Kadonaga, 2003). The ATPases can be divided into seven subfamilies according to characteristic domain features. Much is known about function of SWI2/SNF2, ISWI, CHD and Ino80 subfamilies (reviewed by Eberharter and Becker 2004). The other members are the Cockayne Syndrome protein B (CSB), Rad54 and DDM1 subfamilies. Very little is known about these subfamilies (reviewed by Eberharter and Becker, 2004)

4.4.1 The SWI/SNF Family

The SWI/SNF ATPases are the central subunits of a large complex. These ATPases are characterized by a bromodomain (Horn and Peterson, 2001 and Martens and Winston, 2003). The SWI/SNF complex from yeast is involved in transcriptional activation as well as in repression of several yeast genes (Sudarsanam and Winston, 2000). The ATPase subunit of the yeast SWI/SNF complex is SWI2/SNF2. In human cells there are two ATPases hBrm and Brg1. The *Drosophila* homolog of Swi2p/Snf2p is Brahma (Brm). In *Drosophila,* Brm resides in two closely related complexes that share many subunits: the BAP (Brahma-associated proteins) complex, (Wang, 2003) and the PBAP (polybromo-associated BAP) complex (Mohrmann *et al*., 2004) (reviewed by Eberharter and Becker, 2004). The SWI/SNF complexes are able to slide histone octamers and modify the path of DNA within the nucleosome without nucleosome relocation. The SWI/SNF ATPases are required to relieve the constraints of nucleosomal DNA to regulate the accessibility of genomic DNA to transcription factors (reviewed by Kwon and Wagner, 2007).

4.4.2 The ISWI Family

The Imitation switch (ISWI) is largest subfamily of the SWI2/SNF2 superfamily. The ISWI family of chromatin remodelers has a characteristic C-terminal SANT-like domain (Grüne *et al*., 2003). There are several complexes described in a variety of species. The ISWI family was first identified in *Drosophila*. There are three different ISWI complexes known in *Drosophila*: NURF (nucleosome remodeling factor), ACF (ATP dependent chromatin assembly and remodeling factor) and CHRAC (chromatin accessibility complex) (Längst and Becker, 2001). In mammals two ISWI-homologs exist, SNF2H and SNF2L (Barak *et al*., 2004).

4.4.3 The CHD Family

The most prominent member of CHD family of chromatin remodeling complexes is the ATPase Mi-2. Mi-2 is found in nucleosome remodeling histone deacetylase complexes that have histone deacetylase activity (Feng and Zhang, 2001). Members of the CHD family have the characteristic chromodomain (Kelley *et al*., 1999). A

multi-subunit Mi-2 complex has been isolated from *D. melanogaster* (reviewed by Bouazoune and Brehm, 2006). There are four members of *Drosophila* CHD family, dCHD1, dCHD3, dMi-2 and Kismet. Most studies imply that dMi-2 has a role in transcription repression (reviewed by Bouazoune and Brehm 2006 and Brehm *et al.,* 2000). Very little is known about the other members of the CHD family.

4.4.4 The Ino80/SWR Family

This family of ATP dependent chromatin remodeling complexes is less well characterized. They are also called as 'Split ATPase domain enzyme' as the ATPase domain contains an insert, which splits the region in two segments. Not much is known about these enzymes. Until now the multi subunit complexes of this family are only known in yeast (Krogan *et al*., 2003) (reviewed by Eberharter and Becker 2004).

4.5 mRNA Processing

Expression of the eukaryotic genes is a complex process involving several steps. The processing of pre-mRNA includes the capping of the 5´ end of the nascent pre-mRNA, the removal of introns and the cleavage of the 3´ end for the production of a polyadenylated mature mRNA. Each step requires several factors, co-factors and multi-subunit complexes (reviewed by Maniatis and Reed, 2002 and Bentley, 2002). Prior to the formation of fullength pre-mRNA, the nascent mRNA should be modified and loaded with several RNA binding proteins and complexes. For the mature mRNP particle, which is ready to be exported into the cytoplasm, the processing becomes mandatory. Most of the current research in the field implies that the processing events occur co-transcriptionally and are interconnected to each other (reviewed by Maniatis and Reed, 2002; Bentley, 2002 and Sims *et al*., 2004).

The RNA Polymerase II (RNAPII) plays an important role in mRNA processing during transcription. RNAPII is a multimeric enzyme. The C-terminal domain (CTD) of the largest subunit of RNAPII is composed of a heptad repeat sequence YSPTSPS (single-letter amino acid code). This consensus sequence contains serine residues that can be phosphorylated. In particular the phosphorylation of serines 2 and 5 have an important functional significance in the process of transcription (reviewed by Cramer *et al*., 2001). The phosphorylated CTD acts as a signal for various factors and complexes to interact with RNAPII and facilitate the upcoming processing event (reviewed by Bentely, 2002).

4.5.1 5´ end processing of mRNA

All RNAPII transcripts are modified with a 7-methyl G5´ppp5´N cap. The cap is added when RNA is 20-30 nucleotides long (reviewed by Proudfoot *et al*., 2002). The capping of 5´ end occurs in three steps involving three enzymes: RNA phosphatase, guanylyltransferase (GT) and 7-methyltransferase (MT) (reviewed by Bentley 2002 and Proudfoot *et al*., 2002). The first step involves hydrolysis of the first nucleotide of the pre-mRNA to a diphosphate by the enzyme 5´ RNA phosphatase. During the second step, a GMP moiety is added to the first nucleotide of pre-mRNA by guanylytransferase. In the final step the methyltransferase methylates the N7 position of the added GMP moiety. This cap structure is further recognized by the cap binding

complex (CBC). The CBC comprises of two cap binding proteins CBP20 and CBP80 (reviewed by Shatkin and Manley 2000). Binding of CBC to the cap structure stabilizes mRNA by preventing its degradation by 5´-3´ exonucleases (reviewed by Proudfoot *et al*., 2002). The CTD of RNAPII plays an important role in capping the 5´ end of the mRNA molecule. It has been shown that the enzymes guanylyltransferase and methyltransferase bind to the phosphorylated CTD (reviewed by Bentley, 2002). The major function of the RNAPII CTD is to act as a platform to bind directly to factors involved in pre-mRNA capping, 3´ end processing, transcription elongation, termination, and chromatin modification thereby coupling these reactions to each other (reviewed by Perales and Bentley, 2009).

4.5.2 Splicing

In most eukaryotic protein coding genes, introns are found interrupting the coding sequences. These interrupting introns should be removed in order to get the functional transcripts from the DNA template. The pre-mRNA contains several consensus sequences that play an important role in the splicing reaction. The pre-mRNA contains two kinds of splice sites, the 5´ and 3´ splice sites. The 5´ splice site is marked by the presence of a conserved sequence AG/GURAGU (R, purine). This sequence is present at the exon-intron junction. The 3´ splice site is recognized by the sequence YAG/RNNN (R, purine; Y, pyrimidine) present at the end of intron. The 3´ splice site follows a highly conserved sequence called the branch point. The branch point sequence consists of a conserved adenosine followed by a polypyrimidine stretch. The splicing reaction involves two steps in chemical terms. The first reaction involves freeing of 5´ exon producing a lariat structure containing the intron and the 3´ exon. The second step involves the ligation of the two exonic sequences and releasing the lariat intron. These reactions are facilitated by assembled spliceosomes. The spliceosome contains five small nuclear RNAs (snRNAs), U1, U2, U4, U5 and U6 complexed with proteins forming snRNP (snRNP). The snRNPs interact with SR proteins eg. ASF/SF2 and other splicing regulators. The interaction of U1 snRNA with the 5´ splice site initiates the assembly of spliceosome by recruiting other factors at the site of action. The assembly of spliceosome is regulated by several factors including SR proteins (reviewed by Krämer, 1996, Proudfoot *et al*., 2002) and the CTD of RNAPII. It has been shown

that phosphorylated CTD stimulates the assembly of the spliceosome *in vitro* and is required by splicing enhancers *in vivo* (reviewed by Bentley, 2002).

One of the major regulatory steps in the gene expression pathway is the regulation of pre-mRNA splicing. The splicing patterns of several pre-mRNAs are very variable. The use of different splice sites gives rise to multiple transcripts. This phenomenon is called alternative splicing (reviewed by Holste and Ohler, 2008). Recent studies suggest that almost 70% of the genes in the human genome are subjected to alternative splicing (reviewed by Holste and Ohler, 2008). One of the most impressive examples of alternative splicing comes from the *Dscam* gene of *D. melanogater. Dscam* stands for Down Syndrome Cell Adhesion Molecule. This gene encodes for an axon guidance receptor. It is now known that *Dscam* is expressed into several hundreds to thousands of alternative mRNA splice variants (reviewed by Graveley, 2001). The diversity in the splicing pattern shows the flexibility of the spliceosome in recognizing and processing exons in a given pre-mRNA. Therefore, the splicing process should be tightly regulated for the efficient and correct splicing of the transcripts. The regulation of the splicing process is also important because it determines when a particular protein isoform should be expressed (reviewed by House and Lynch, 2008).

The process of splicing regulation is governed by several parameters such as splice site strength and the secondary structure of RNA, the exon/intron architecture and the process of pre-mRNA synthesis (reviewed by Hertel, 2008). The recognition of splice sites holds an important position in the regulation of splicing. The recognition of splice sites is regulated by splicing enhancers and silencers and the local concentration of different splicing factors. The splicing enhancers and silencers are cis-acting RNA sequence elements. The exonic splicing enhancers (ESEs) serve as binding site for splicing enhancer complexes. The sequences that repress splicing or block the recruitment of heterogeneous nuclear RNPs (hnRNPs) are called splicing silencers. They can be either exonic splicing silencers (ESS) or intronic splicing silencers (ISS) (reviewed by Hertel, 2008).

The secondary structure of the pre-mRNA also plays an important role in splicing efficiency. It has been proposed that the local secondary structure of the pre-mRNA can facilitate or inhibit the assembly of spliceosome (Hertel and Graveley, 2005 and Hertel 2008). Not only the sequence elements but several transcription

regulators and coregulators are known to influence alternative splicing. It has been reported that coregulators of transcription initiation are also involved in further pre-mRNA processing events, specifically splicing (Auboeuf *et al*., 2007). It has also been found that steroid hormones that act as transcription stimuli are involved in splicing and other post-transcriptional events (Lou *et al*., 2001). The steroid hormones are known to regulate the exon composition of their target gene transcript (Auboeuf *et al*., 2002).

The serine/arginine rich (SR) proteins have multiple functions in pre-mRNA splicing. The SR proteins are known to bind to the pre-mRNA and recruit number of splicing factors to the pre-mRNA during spliceosomal assembly (reviewed by Graveley, 2000). The SR proteins also regulate alternative splicing by mediating the interaction between the splicing factors bound to the 5´ and 3´ splice sites (Wu and Maniatis, 1993 and Graveley, 2000).
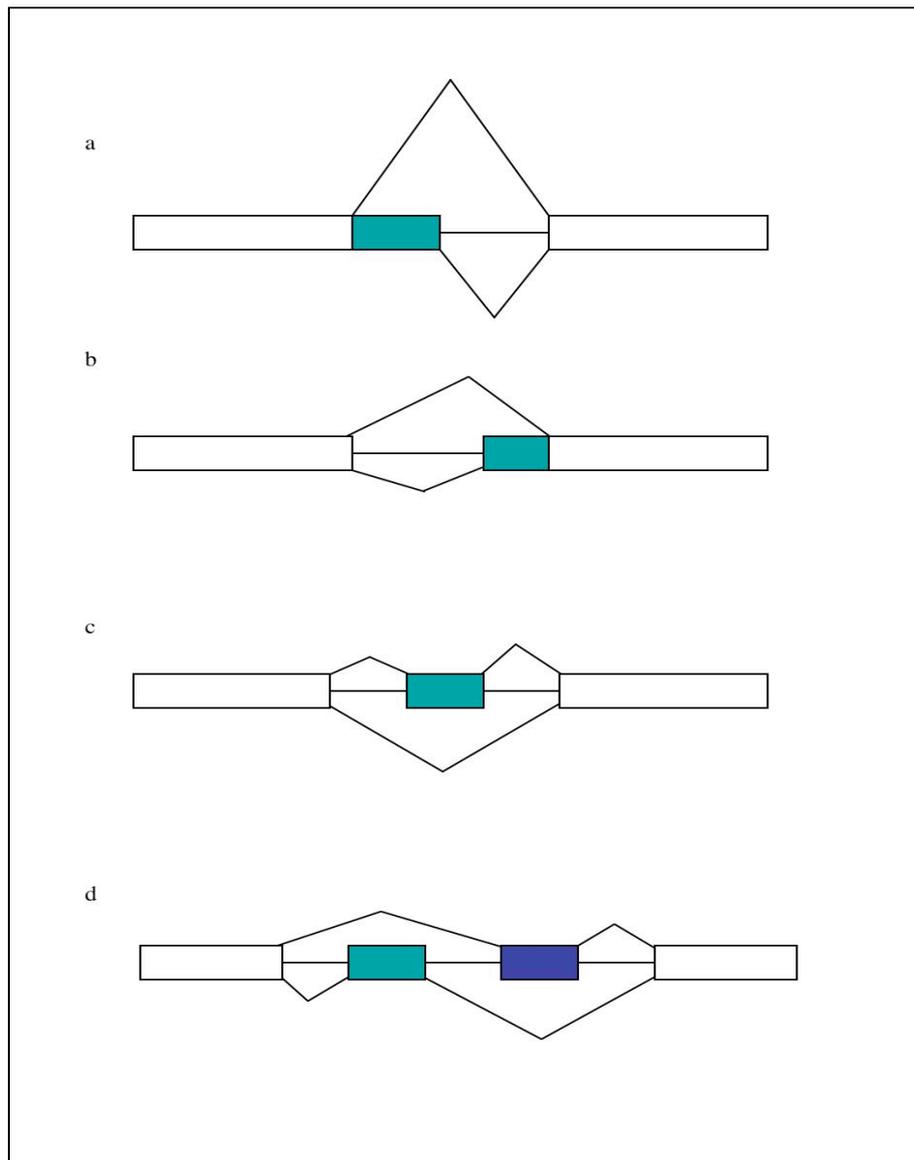
Figure 2: Different ways of alternative splicing. Exons can have alternatively spliced 5´ or 3´ (Figure 2a and b). An alternative exon can reside between two constitutive exons so that it can be either included or excluded (2c). There can be multiple alternative exons between two constitutive exons. The presence of mutiple alternative exons presents a choice to the splicing machinery.

4.5.3 3´ end processing of mRNA

Almost all eukaryotic mRNAs contain a poly(A) tail. The 3´ end processing is a two-step reaction involving endonucleolytic cleavage followed by addition of poly(A) tail (reviewed by Proudfoot, 2004). There are two sequence elements that decide the specificity and efficiency of the 3´ end processing event. The first element is the conserved sequence AAUAAA. This sequence is defined as polyadenylation signal.

The polyadenylation signal is present upstream of a dinucleotide cleavage site CA. The polyadenylation site is recognized by cleavage and polyadenylation specificity factor, (CPSF). This RNA-protein interaction decides the site of cleavage immediate down stream of the dinucleotide CA. The second elements that determine the specificity is the presence of a GU-rich sequence about 20-30 nucleotides downstream of the site of cleavage. This downstream sequence element, DSE, interacts with the cleavage stimulating factor (CstF), increasing the efficiency of the 3′ end processing (reviewed by Danckwardt *et al*., 2008). Together CPSF and CstF form the basal 3′ end processing machinery (Ryan *et al*., 2004 and Mandel *et al*., 2006). The assembly of a basal machinery leads to cleavage of the transcript and subsequently addition of a poly(A) stretch to the transcript by a nuclear poly(A) polymerase (reviewed by Danckward *et al*., 2008). This event leads to the production of a mature mRNA stabilized by the added poly(A) tail and a downstream RNA product that is degraded by exonucleases (reviewed by Proudfoot, 2004).

Like in other processing events, the CTD plays an important role in proper 3′ end processing. The complexes CPSF and CstF of the basal 3′ end processing machinery are shown to interact directly with the CTD (reviewed by Proudfoot, 2004). The research in the field suggests that a large number of genes are subjected to regulation of 3′ end processing. Studies of genome and EST data suggest that most of the genes coding for multiple transcripts have multiple polyadenylation sites (reviewed by Danckwardt *et al*., 2008). It has been observed that about half of the human mRNAs contain more than one polyadenylation site (Tian *et al*, 2005 and Yan and Marr, 2005). Some of the determinants of alternative poly(A) site selection are (1) the strength of the sequence elements, (2) the concentration or activity of polyadenylation factors and/or (3) the presence of tissue- or stage-specific regulatory factors (Barabino and Keller, 1999). A well known example of alternative polyadenylation site usage is that of immunoglobulin class switch during the differentiation of B lymphocytes (Takagaki *et al*., 1996 and Zhang *et al*., 2005). The B cells and their precursor cells produce a large amount of mRNAs that encode for membrane bound form of IgM H (heavy) chain (μm). The plasma cells produce mRNAs that encode for a secretory form (μs). The switch from μm to μs form is due to the use of μm specific polyadenylation signal in B cells and the use of μs specific polyadenylation signal in plasma cells. This switch is due to the change in the

splicing pattern of the pre-mRNA and subsequently the usage of an alternative poly(A) signal. The *μ*m-specific polyadenylation site is used in B cells due to an upstream splicing event that removes the *μs* polyadenylation signal (Takagaki *et al*., 1996). The selection of alternative poly(A) site in this case is known to be regulated by the concentration of the CstF 64-kDa subunit (reviewed by Danckwardt *et al*., 2008).

Alternative polyadenylation produces mRNA isoforms with 3′ UTR that either contains cis-regulatory elements for example, AU rich elements (AREs) or miRNA binding sites, depending on the proximal or distal poly(A) site used. Legendre *et al.* have shown by analysis of 3′ UTRs produced due to alternative polyadenylation that 52% of miRNA target sites are located downstream of the first poly(A) site (Legendre *et al*., 2006) Sandberg *et al.* also found that longer 3′ UTRs have a 2.1-fold higher number of miRNA target sites than those with shorter 3′ UTRs in T cells mRNAs (Sandberg *et al*., 2008). AREs were found to be present in approximately 10–15% of all transcripts (Halees *et al.*, 2008) and were shown to interact with several proteins, some of which contribute to mRNA stability (reviewed by Barreau *et al*., 2005) and others control translation (reviewed by Espel, 2005).

4.6 Coupling between mRNA processing steps

Eukaryotic gene expression is a complex process and each step is tightly regulated. Recent research has shown that most of the processing steps occur cotranscriptionally, coupling mRNA processing to transcription. One of the well studied events in gene expression is splicing. Two mechanisms have been proposed to explain the effect of transcription on splicing. One model is the "kinetic model" and the other is based on the recruitment of splicing factors to the RNAPII CTD (reviewed by Auboeuf *et al*., 2007). The kinetic coupling suggests that the rate of transcription elongation governs the alternative splicing events of cassette exons that are not constitutively included in the mature mRNA. According to this model, when RNAPII elongates with a faster rate then both strong and weak 3´ splice sites are presented to the splicing machinery almost at the same time. Under this condition the strong splice site outcompetes the weak splice site, resulting in skipping of alternative exons (reviewed by Kornblihtt 2006). On the other hand, if the RNAPII slows down or pauses between the weak and the strong 3´ splice site, then only weak splice site is presented or available to the spliceosome. So splicing occurs at the weak 3´ splice site and this results in the inclusion of alternative exons (reviewed by Kornblihtt 2006). Different factors can affect the elongation rate of RNAPII. It has been known that changes in chromatin structure affect splicing. For example a study shows that trichostatin A (inhibitor of histone deacetylation) favors exclusion of alternative exons. It is believed that this is due to hyperacetylation of histones, which makes the passage of transcribing RNAPII easier (Nogues *et al*., 2002). Another study has also revealed the new role of chromatin remodeling factor, Brm, the ATPase subunit of chromatin remodeling factor SWI/SNF. In this study CD44 is used as the model gene. This gene has a stretch of 10 variable exons in between constitutive exon 5 and 16. Batsché *et al.* propose that when RNAPII reaches the variable exons, a macromolecular complex is formed that includes Brm and the RNA-binding protein Sam68, which belongs to the STAR (Signal transduction and activation of RNA) family. This complex somehow stalls the polymerase thus favoring the inclusion of the variable exons (Batsché *et al*., 2006 and reviewed by Kornblihtt 2006).

The second model explains transcription and splicing coupling by relying on the fact that different splicing factors are recruited to the CTD of RNAPII. The CTD contains the heptad repeats that contain serines at 2nd and 5th positions. During

transcription initiation the CTD is phosphorylated at serine-5 position and later serine-2 position becomes highly phosphorylated. The RNAPII phosphorylated at serine 2 position is believed to be processive and to interacts with several splicing factors. This interaction of the splicing factors with the CTD influences the concentration of specific splicing factors in the vicinity of the transcript and thereby affects the splicing decision (reviewed by Auboeuf *et al*., 2007). Also it has been known that some of the eukaryotic genes have multiple promoters. Each promoter defines a different start site, giving rise to different transcript. Each alternative promoter is regulated by a different set of transcription factors and these transcription factors can recruit specific splicing factors (reviewed by Kornblihtt, 2005). It has been suggested that there is a strong link between alternative promoter usage and alternative splicing of the variable internal exons (Kornblihtt *et al*., 2005).

Recent studies in the field of alternative splicing have established that chromatin acts a regulator of splicing (reviewed by Luco and Mistelli, 2011). Contribution of RNAPII elongation rate in alternative splicing leads to the assumption that chromatin structure is important for RNAPII elongation. Transcription through nucleosome-dense regions, such as compacted chromatin, is slower than through nucleosome-poor regions (reviewed by Luco *et al.,* 2011). The sequence-dependent nucleosome positioning in a transcription unit leads to RNAP II accumulation upstream of the positioned nucleosome suggesting that nucleosome density and chromatin structure may regulate RNAP II kinetics (Subtil-Rodriguez and Reyes, 2010). A recent study by Spies *et al*. shows that isolated exons flanked by long introns are enriched in nucleosomes (Spies *et al.,* 2009). Another study by Schwartz *et al*. shows that alternatively spliced exons have less enrichment than constitutive exons suggesting that nucleosome positioning and alternative splicing regulation are connected. Also, genome-wide mapping of RNAPII reveals a non-random enrichment at exons (Schwartz *et al*., 2009 and Chodavarapu *et al*., 2010). In addition to nucleosomes, histone modifications and DNA methylation have also been shown non-randomly distributed over exons (Schwartz *et al*., 2009 and Chodavarapu *et al*., 2010). High-throughput mapping of histone marks by Kolasinska-Zwierz *et al*. in human T cells showed that H3K36me3 is more enriched in constitutive exons than in alternatively spliced exons (Kolasinska-Zwierz *et al*., 2009). Distinct subsets of histone modifications have been correlated with particular splicing patterns in several

genes. Modulation of splicing-specific chromatin signatures by inhibitors, overexpression of histone modifiers, or knockdown experiments show changes in alternative splicing of several genes suggesting that some histone modifications play a direct role in regulating alternative splicing (Allo *et al*., 2009; Schor *et al*., 2009; Luco *et al*., 2010 and Saint-Andre *et al*., 2011). The function of histone modifications in alternatively spliced genes may be to establish and maintain nucleosome density in the region, which in turn might regulate RNAP II elongation rate (Luco *et al*., 2011). In addition, it is unclear whether kinetic regulation alone can account for competition between splice sites of equal strength and situations where more than two splice sites are involved (reviewed by Luco and Mistelli, 2011).

Recent research also suggests that 3′ end processing is regulated by chromatin and histone modifications. The nucleosome organization around the promoter regions has been studied a lot but little is known about their organization at the 3′ end of genes. Recent ChIP-seq analysis by Spies *et al.* shows a connection between polyadenylation and histone positioning *S. cerevisiae*. The 3′ region near the poly(A) site was shown to be depleted of nucleosomes (Mavrich *et al.*, 2008; Shivaswamy *et al*., 2008 and Spies *et al.,* 2009). The depletion of nucelosomes in this area could be because of lower intrinsic affinity for nucleosomes of the nucleotide sequence, so it could be that the nucleosome free regions are related to transcription and not 3′ processing (reviewed by Richard and Manley, 2009). Recently, it was confirmed that nucleosome depletion around human poly(A) sites is indeed connected to 3′ processing. Spies *et al*. analyzed published ChIP-Seq data sets from human T cells (Barski *et al*., 2007; Schones *et al.,* 2008 and Spies *et al.,* 2009) and found low nucleosome density observed at the AATAAA sequence was quite prominent around actively used poly(A) sites (in genes with multiple poly(A) sites), suggesting either that sequences around the poly(A) signal may play a role in nucleosome positioning or that a yet unknown nucleosome-excluding DNA binding protein maybe be commonly bound near the poly(A) sequence. It has been suggested that nucleosome positioning may affect alternative polyadenylation, either by influencing the rate of polymerase elongation or via 3′ complex-dependent recruitment of a chromatin remodeling factor (reviewed by Di Giammartino *et al*., 2011).

5. Model systems

5.1 *Chironomus tentans* as a model system

The life cycle of *C. tentans* includes four stages: egg, larvae, pupa and adult midge. The larval stage has four instars. At fourth instar the larvae are about 1.5 cm long and contain well developed salivary glands. Each salivary gland contains 35-40 giant cells around the central lumen. In each cell, the nucleus contains four polytene chromosomes each with a specific striped pattern of bands and interbands. This pattern reflects the uneven but specific compaction level of chromatin along the chromosomes. Each chromosome is made of 8000-16000 chromatids that are longitudinally aligned (Daneholt and Edström, 1967). An Italian researcher Balbiani initially observed the polytene chromosomes of C. tentans in 1881 (Balbiani, 1881). Among all the other genes, the polytene chromosomes contain Balbiani ring (BR) genes. There are five BR genes. Four genes among them BR1, BR2.1, BR2.2 and BR6 are exceptionally large, about 40 kb, whereas BR3 is about 11 kb in size.

The BR genes are highly transcribed in the larval salivary glands. There are four BRs (BR1, BR2.1, BR2.2 and BR3) that are present on chromosome IV. The fifth BR, BR6, appears on chromosome III under special conditions. The BRs are the unfolded chromatin loops that are active sites of transcription (Lamb and Daneholt, 1979). The BR1 and BR2 genes contain four introns, three located close to the 5´ end and one near the 3´ end. BR3 contains 38 short intron distributed along the whole transcript (Case and Daneholt 1978 and Paulsson *et al.,* 1990). After processing, the BR1 and BR2 transcripts produce 35 kb long mRNAs, and BR3 mRNA is about 5.5 kb long. The BR mRNAs are like any other mRNA but exceptionally long. They are capped, spliced, cleaved and polyadenylated in the normal way (Bauren *et al*., 1998). They are also associated with different nuclear proteins and protein complexes like SR proteins, hnRNPs, snRNPs (Kiseleva *et al*., 1994) and various export factors (Zhao *et al*., 2002 and Kiesler *et al*., 2002). The BR1 and BR2 mRNAs are released into the nucleoplasm and are transported to the cytoplasm (Daneholt and Hosick, 1973). These long mRNAs are translated into giant secretory polypeptide of about 10^3 kDa. The secretory proteins are used by the larva to build the tubes in which they reside (Daneholt *et al*., 1976).

Upon transcription, the BR pre-mRNA is assembled into a large ribonucleoprotein complex called BR RNP particle. The mature BR particles are about 50 nm in diameter and can be morphologically distinguished at the site of synthesis in the interchromatin region (Skoglund *et al*., 1986) and at the nuclear pore complex during translocation (Mehlin *et al*., 1992).

The BR genes display a distinct polarity defined by the size and the structure of the nascent BR RNP particle along the gene (Skoglund *et al*., 1986). Based on this polarity, the BR genes can be divided into proximal, middle and distal regions. In the proximal region, close to the promoter, the growing pre-mRNA is packed into a RNP fiber. In the middle region, as the RNP fiber grows in length, it acquires a globular shape. In the distal part of the gene, the pre-mRNP appears as a globular structure with a diameter of approximately 50 nm. Thanks to this polarity, it is possible to map the binding of proteins to the different parts of the BR gene using immuno-electron microscopy (immuno-EM). This approach has been successfully applied to analyze the location of some proteins like, histones (e.g. Ericsson *et al*., 1990), splicing factors (Kiseleva *et al*., 1997) and a chromatin remodeling factor (Sjölinder *et al*., 2005) using specific antibodies.

Not only the distribution of proteins along the gene can be determined, but also if the protein is associated with the nascent pre-mRNP or in close association with the chromatin. The diameter of the BR particle together with the length of the stalk, with which the BR pre-mRNP is attached to the template, is large enough to allow distinguishing between the labeling on RNP particle and/or the labeling in close contact with chromatin (see Figure 3).

5.2 Other model systems

Chironomus provides unique features for the study of gene expression *in situ* but there are some limitations. Firstly, the genome of *C. tentans* is not fully sequenced. Secondly, there are no tools available for genetic manipulation. And finally there are no cell lines available that can be used for transfection assays. Therefore for RNAi assays and for genome wide analysis we used *Drosophila* S2 cells. The S2 cell line was derived from a primary culture of late stage (20-24 hours

old) *Drosophila melanogaster* embryos. This cell line grows rapidly at room temperature in suspension.
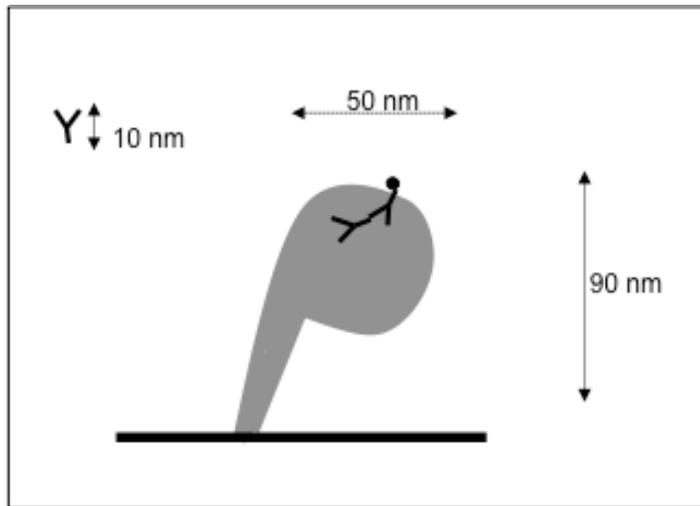
Figure 3. The shape and size of the BR RNP particle in the distal part of the BR gene. The figure shows the labeling on the BR particle and the dimensions of the immunoglobulins relative to the RNP.

The third model system that we used is HeLa cells. The HeLa cells (also Hela or hela cell) are immortal cell lines used in medical research. The cell line was derived from cervical cancer cells taken from Henrietta Lacks, who died from her cancer in 1951. We used S2 cells and Hela cells for the comparative studies of our biochemical fractionation.

6. Aim of the Study

The aim of this work is to study the role of Brahma (Brm) in gene expression using the insect model systems *C. tentans* and *D. melanogaster*. Brm is the ATPase subunit of the SWI/SNF complex and has the known function of chromatin remodeling. The complex has a DNA stimulated ATPase activity and can destabilize histone-DNA interactions in reconstituted nucleosomes in an ATP dependent manner. A lot of work has been done regarding the role of Brm as a chromatin remodeler. Recently it was found that Brm also acts as a regulator of alternative splicing in humans (Batché *et al*., 2006).

The ability of BRM to remodel nucleosomes as the ATPase subunit of SWI/SNF has already been established. Using the possibility to perform electron microscopy on the isolated polytene chromosomes of *C. tentans* and to locate the protein, the aim of this study is to map the presence of Brm in the BR system. Since depletion or over expression of proteins in the *C. tentans* cells cannot be attained, using *Drosophila* S2 cells this study aims to investigate the role of Brm in gene expression and, more precisely, in pre-mRNA processing.

7. RESULTS

7.1 The association of Brahma with the Balbiani ring 1 gene of Chironomus tentans studied by immunoelectron microscopy and chromatin immunoprecipitation. (2008)

Botelho SC, Tyagi A, Hessle V, Farrants AK, Visa N.

The association of Brahma with the Balbiani ring 1 gene of *Chironomus tentans* studied by immunoelectron microscopy and chromatin immunoprecipitation

S. Calado Botelho*‡, A. Tyagi*, V. Hessle*,
A. K. Östlund Farrants† and N. Visa*

*Department of Molecular Biology & Functional Genomics
and †Department of Cell Biology at the Wenner-Gren
Institute, Stockholm University, SE-10691 Stockholm,
Sweden*

Abstract

**Many steps of gene expression take place during
transcription, and important functional information
can thus be obtained by determining the distribution of
specific factors along a transcribed gene. The Balbiani
ring (BR) genes of the dipteran *Chironomus tentans*
constitute a unique system for mapping the associa-
tion of specific factors along a eukaryotic gene using
immuno-electron microscopy (immuno-EM). The
chromatin immunoprecipitation (ChIP) technique has
provided an alternative, more general method for stud-
ying the association of proteins with specific genomic
sequences. The immuno-EM and the ChIP methods
suffer from different limitations, and thus a combination
of both is advantageous. We have established optimal
conditions for ChIP on chromatin extracted from the
salivary glands of *C. tentans*, and we have analyzed
the association of the SWI/SNF chromatin remodelling
factor Brahma (Brm) with the *BR1* gene by combined
immuno-EM and ChIP. We show that Brm is not restricted
to the promoter of the *BR1* gene but is also associated
with sequences in the middle and distal portions of the
gene, which suggests that Brm has additional roles
apart from regulating transcription initiation.**

Introduction

Gene expression in eukaryotes is a multi-step process that
requires the co-ordinated activity of many factors. Many of
the gene expression steps, including modifications of
chromatin structure, transcription itself, pre-mRNA proces-
sing and surveillance, take place at the transcription site
(reviewed by Maniatis & Reed, 2002). The Balbiani ring
(BR) genes of the dipteran *Chironomus tentans* have been
used as a model system for mapping the association of
specific factors with transcribed genes by immuno-electron
microscopy (immuno-EM). Two exceptionally large puffs,
called *BR1* and *BR2*, are formed in the polytene chromo-
somes of the larval salivary glands of *C. tentans* as a con-
sequence of the expression of the *BR* genes. The *BR* genes
code for large secretory proteins of the salivary glands
(reviewed by Wieslander, 1994; Daneholt, 2001). In EM
preparations, the transcribed *BR* genes appear as chromatin
loops with a central axis: this is the transcribed chromatin
template, decorated with nascent pre-mRNPs (Skoglund
et al., 1983). The *BR* loops are relatively extended, and
they start and end in tightly packed chromatin (Ericsson
et al., 1989).

The possibility of mapping proteins along the transcribed
BR genes by immuno-EM is based on two facts. Firstly, the
BR puffs are the largest puffs in the polytene chromosomes
of *C. tentans* and can be easily identified using microscopy
methods. Second, each region of the transcribed *BR* genes
– proximal, middle or distal relative to the promoter – is
characterized by a specific structure that makes it possible
to identify the different parts of the gene by microscopic
examination (reviewed by Daneholt, 2001). In addition,
the *BR* genes appear to have all the features of a typical
eukaryotic gene, their transcription is regulated in a tissue-
specific manner (reviewed by Wieslander, 1994), and their
primary transcripts undergo normal processing (Baurén &

505

Wieslander, 1994; Visa *et al.*, 1996; Baurén *et al.*, 1998). For these reasons, the *BR* genes of *C. tentans* have provided a suitable system for the direct mapping of proteins along transcribed eukaryotic genes *in situ*. Two types of proteins have been mapped along the *BR* genes: chromatin components such as histones (see, for example, Ericsson *et al.*, 1990), and mRNA-binding factors (see, for example, Sjölinder *et al.*, 2005). However, the spatial resolution achieved by the immuno-EM analysis of *BR* genes is low, of the order of 5–10 kb, due to intrinsic limitations of the morphological criteria used for the identification of the different parts of the gene.

Development of the chromatin immunoprecipitation (ChIP) technique in recent years has provided a general method for studying the association of proteins with specific genomic sequences (Solomon *et al.*, 1988; Orlando *et al.*, 1997). ChIP has been applied to different types of samples, including cultured cells and tissues (e.g. Legube *et al.*, 2006). ChIP offers several advantages over immuno-EM mapping. ChIP can be used to study the association of the protein of interest with virtually any genomic sequence at a resolution that is much higher than that achieved by immuno-EM mapping. However, ChIP is limited by the efficiency of the crosslinking between the protein under study and the DNA.

Immuno-EM and ChIP thus suffer under different restraints, and a combination of the two methods may be useful, providing reliable information about the association of proteins of interest with specific gene regions. The goals of the present study were to extend the use of ChIP to the salivary glands of *C. tentans* in order to be able to carry out immuno-EM and ChIP analyses in the same system, and to analyze the association of Brahma (Brm) with the *BR* genes. Brm is a component of SWI/SNF chromatin remodeling complexes and participates in the transcriptional regulation of many genes (Khavari *et al.*, 1993; Muchardt & Yaniv, 1993). Brm is evolutionarily conserved and its importance for gene expression has been shown in different organisms. In *Drosophila melanogaster*, the Brm chromatin remodelling complexes are associated with many transcribed genes and facilitate transcription by RNA polymerase II (Armstrong *et al.*, 2006 and references therein). Brm interacts with the genome in a chromatin context and regulates transcription by remodelling nucleosome structure at the promoters (Soutoglou & Talianidis, 2002; Sif, 2004; Saha *et al.*, 2006). Furthermore, a recent study has shown that the human Brm regulates the alternative splicing of the CD44 gene (Batsché *et al.*, 2006), which suggests that Brm has additional roles apart from regulating promoter activity. We were thus interested in mapping the distribution of Brm along the *BR1* gene by immno-EM and ChIP.

In order to set up a ChIP protocol for *C. tentans* salivary glands, we have chosen to work with antibodies against Brm and against two additional proteins that associate with the transcribed *BR* genes through different types of molecular interaction. One protein, the histone H3 trimethylated at lysine 4 (meH3), is in direct contact with the DNA and is preferentially associated with the promoter proximal sequences of transcribed genes in both yeast and mammalian cells (Santos-Rosa *et al.*, 2002; Kouskouti & Talianidis, 2005). The second protein, Rrp4, is a component of the exosome, a multiprotein complex that is responsible for many aspects of RNA biology, including ribosome biogenesis, mRNA turnover and co-transcriptional quality control of mRNA biogenesis (reviewed by Houseley *et al.*, 2006). Studies in *D. melanogaster* have shown that the exosome is bound to transcribed genes through protein-protein interactions with components of the transcription machinery (Andrulis *et al.*, 2002). We have established conditions required to perform ChIP using antibodies against these three proteins, and we present results on the combined mapping of the Brm protein by immuno-EM and ChIP.

Results and discussion

Brm, Rrp4, and meH3 are associated with the BR puffs

We analyzed the specificity of the antibodies used in the present study by Western blotting against proteins prepared from nuclear extracts of *C. tentans* tissue culture cells (Fig. 1A) and from larval salivary glands (Fig. 1B). In nuclear extracts, each of the antibodies recognized one major band with the expected molecular mass: 190 kDa for Brm, 34 kDa for Rrp4, and 15 kDa for meH3. The same bands were recognized in total salivary gland preparations (Fig. 1B). The additional bands observed in salivary glands with the anti-H3 and the anti-Rrp4 antibodies are likely to correspond to proteins of cytoplasmic origin.

We next analyzed the association of Brm, Rrp4 and meH3 with the *BR* puffs by immunofluorescence. Polytene chromosomes were manually isolated from salivary glands of fourth instar larvae, and the isolated chromosomes were immunostained. Confocal sections of chromosomes stained with each of the antibodies are shown in Fig. 1C. The three proteins were associated with multiple loci in the polytene chromosomes, as has been seen in other systems (Andrulis *et al.*, 2002; Santos-Rosa *et al.*, 2002; Armstrong *et al.*, 2006). The fluorescent patterns obtained with the three antibodies were not fully overlapping, but in all cases the *BR* puffs were among the most intensely stained loci in chromosome IV (Fig. 1C).

We concluded that Brm, Rrp4 and meH3 are associated with the transcribed *BR* genes in the salivary gland polytene chromosomes of *C. tentans*.

Optimization of a ChIP protocol for C. tentans *salivary glands*

The ChIP method has been applied to cells from several types of tissue, including the larval salivary glands of
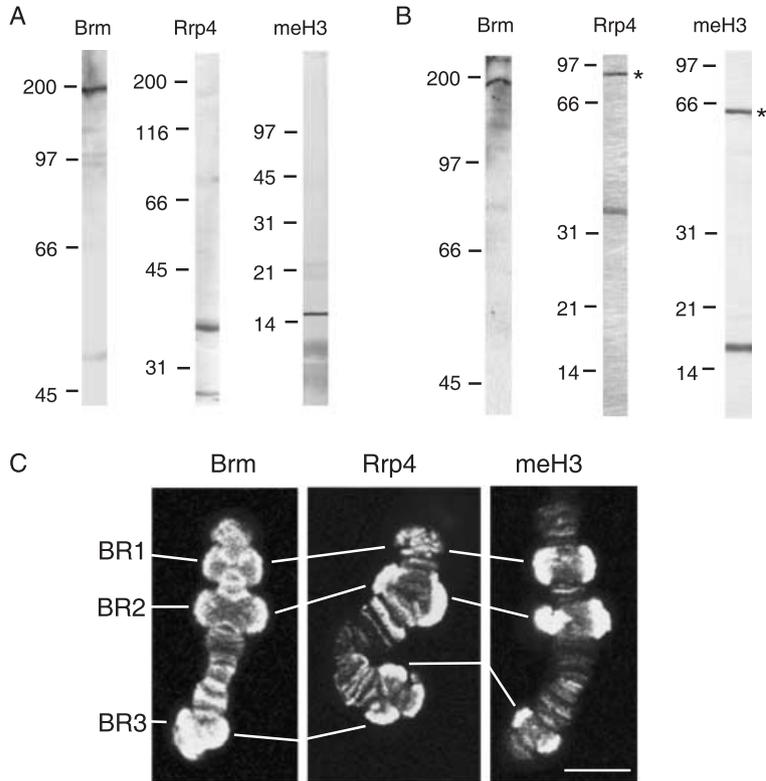
Figure 1. Immunochemical detection of Brm, Rrp4 and meH3 in *Chironomus tentans*. (A) The specificity of the antibodies against Rrp4, Brm and meH3 in *C. tentans* was analyzed by Western blotting using nuclear protein extracts prepared from *C. tentans* tissue culture cells. Molecular mass standards are shown to the left in kDa. (B) Total cell extracts prepared from larval salivary glands were probed with the same antibodies. Molecular mass standards are shown to the left in kDa. The asterisks indicate additional cross-reacting proteins presumably of cytoplasmic origin. (C) Isolated polytene chromosomes were immunostained and analyzed by confocal microscopy. The figures show examples of chromosome IV stained with antibodies against Brm, Rrp4 or meH3, as indicated. The three *BR* puffs (BR1, *BR2* and BR3) located in chromosome IV are indicated. The scale bar represents approximately 10 μm.

D. melanogaster (see, for example, Legube *et al.*, 2006). However, the available protocols failed to give detectable signals when applied to *C. tentans*. This is probably due to in the difficulty of extracting chromatin from organs composed of giant polytene cells in which the nucleus is far away from the cell surface. In a typical mammalian cell in culture, the nucleus is only 3–5 μm away from the cell surface. In a giant salivary gland cell of *D. melanogaster* the nucleus is approximately 10–20 μm away from the cell surface, and in *C. tentans* this distance is even larger, in the order of 50 μm (Fig. 2A).

In order to find suitable ChIP conditions for *C. tentans*, we analyzed the effect of the sonication time on the efficiency of chromatin extraction and on the size of the extracted DNA. We believe that sonication not only shears the DNA, it also has an important function in disrupting the tissue and facilitating the release of the chromatin. This effect may not be significant in the case of haploid or diploid cells, but is important for tissues composed of giant polytene cells. We fixed dissected salivary glands with 2% formaldehyde for 15 min, permeabilized the fixed glands with detergent-containing buffers as described by Legube *et al.* (2006) and sonicated the glands for 4, 12 or 20 min (Fig. 2B). The formaldehyde fixation was reversed after the sonication, and the DNA that had been released from the glands was purified following standard procedures. The amount and

the size of the extracted DNA were analyzed by agarose gel electrophoresis. Figure 2C shows that the amount of DNA obtained with 4 min of sonication was lower than that obtained with 12 and 20 min. The amounts of DNA extracted with 12 and 20 min of sonication were similar, but the sizes of the DNA fragments were smaller in samples that had been sonicated longer. We estimated that as much as 75% of the DNA recovered after 20 min sonication was in fragments shorter than 1 kb (lane 4 in Fig. 2C).

We also analyzed the effect of the sonication on the resolution of the PCR analysis. For this purpose, we analyzed the amplification of PCR products of increasing length, from 200 to 3000 bp, using chromatin sonicated for 4, 12 or 20 min as a template (Fig. 2D). Genomic DNA purified from non-fixed and non-sonicated cultured cells was used in parallel as a positive control (lanes 1–3 in Fig. 2D). The conditions of the PCR reaction were such that even the largest fragment, 3000 bp long, could be efficiently amplified from genomic DNA (lane 3 in Fig. 2D). All three PCR fragments could be detected when DNA extracted after 4 min of sonication was used as a template (lanes 4–6 in Fig. 2D). This observation shows that the DNA fragments obtained under these conditions are relatively long. Sonication for 12 and 20 min resulted in lower quantities of the large PCR products, and an increased yield of the 200 bp product. The best results in the PCR analysis were obtained with
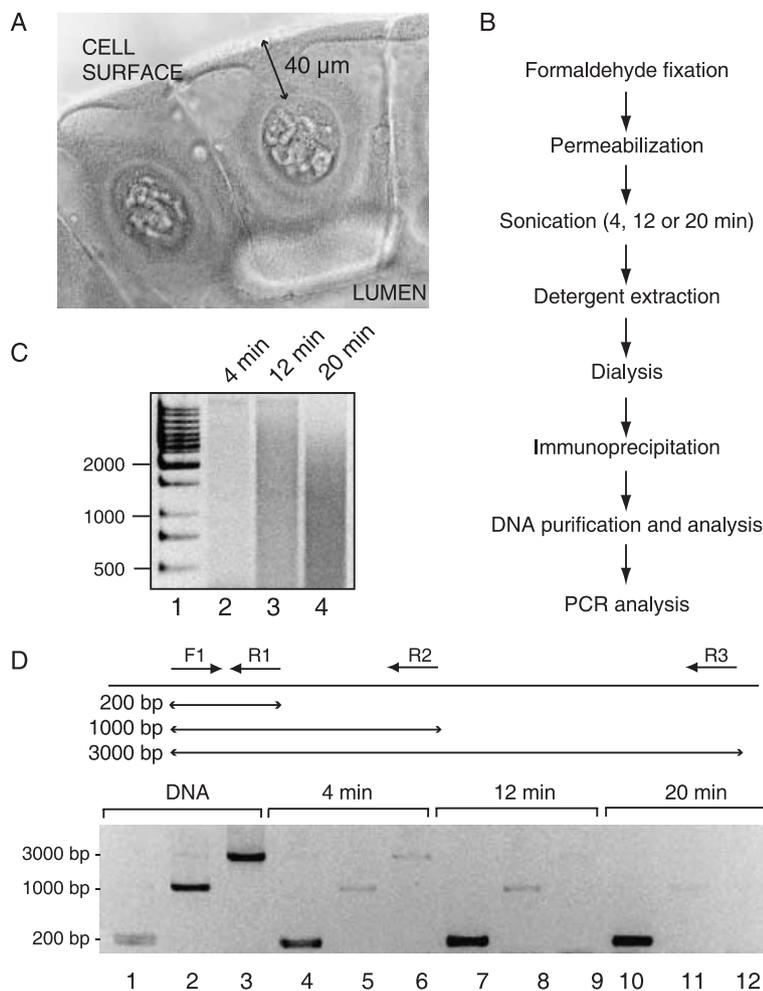
Figure 2. Optimization of sonication conditions for *Chironomus tentans* salivary glands. (A) Phase contrast micrograph of a portion of a salivary gland showing the distance between the surface of the gland and the cell nucleus. (B) Flow chart showing the basic steps of the ChIP procedure. (C) DNA extracted from salivary glands by sonication for 4, 12 or 20 min. The DNA was purified and analyzed by electrophoresis in agarose gels to determine the amount and the size of the extracted DNA. Molecular mass standards are shown to the left in bp. (D) Analysis of the resolution of the PCR amplification. PCR reactions were carried out on genomic DNA (Lanes 1–3) or on DNA purified from chromatin preparations obtained by 4, 12 or 20 min sonication, as indicated (Lanes 4–12). PCR reactions were carried out with primers designed to amplify fragments of 200, 1000 or 3000 bp. In all cases, the PCR products were separated by electrophoresis in agarose gel and stained with ethidium bromide. Molecular mass standards are shown to the left in bp. Note that fragments of 200, 1000 and 3000 bp can be amplified from genomic DNA, whereas only the 200 bp-long product can be amplified from chromatin sonicated for 20 min.

chromatin sonicated for 20 min. Under these conditions, the 200 bp fragment was strongly represented (lane 10 in Fig. 2D), whereas the larger PCR products were undetectable (lanes 11–12 in Fig. 2D).

We concluded that sonication times of up to 20 min increases the yield of extracted DNA and reduces the length of the DNA fragments. We considered that the yield and the resolution provided by 20 min of sonication were suitable for the ChIP experiments. We did not assay longer sonication times in order to avoid unnecessary heating of the chromatin samples, which could result in premature reversal of the formaldehyde crosslinking.

We used chromatin extracted after 20 min of sonication in immunoprecipitation reactions with antibodies against Rrp4, Brm and meH3. The immunoprecipitations were done in the presence of 1% DOC, a relatively high concentration of detergent, to minimize unspecific interactions. Chromatin extracted from 30 glands was used for each immunoprecipitation reaction. The precipitated DNA was analyzed by

PCR using primers designed to amplify sequences at the promoter of the *BR1* gene. The conditions of the PCR reaction were optimized to avoid saturation. Figure 3A and B show the results. The antibodies against Rrp4 and meH3 gave strong signals, which indicated that Rrp4 and meH3 are associated, directly or indirectly, with the promoter of the *BR1* gene. The presence of Rrp4 in a promoter region agrees with the results of a previous study in which the exosome, including Rrp4, was found to interact directly with the RNA polymerase II and with the transcription elongation factor Spt5 in *D. melanogaster* (17). The positive signal obtained with the anti-meH3 antibody is also consistent with the observation that trimethylation of lysine 4 in histone H3 normally occurs at the 5′ ends of highly transcribed genes (Santos-Rosa *et al.*, 2002).

We next asked whether the long sonication times used in this experiment (20 min) could cause loss of signal due to premature reversal of the crosslinking, and we repeated the ChIP experiment with the anti-Rrp4 using chromatin extracted

A

B

C

D
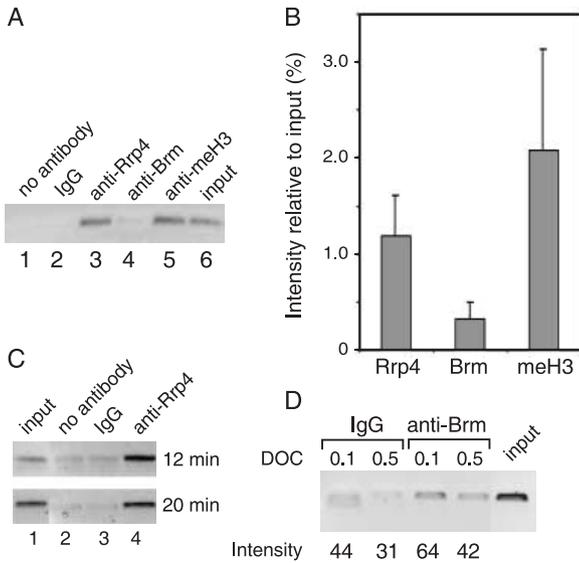
Figure 3. ChIP with antibodies against Rrp4, Brm and meH3. (A) ChIP reactions were carried out on chromatin prepared from *Chironomus tentans* salivary glands using the antibodies against Rrp4 (Lane 3), Brm (Lane 4) and meH3 (Lane 5), in the presence of 1% DOC. Two types of negative controls were processed in parallel: a sample without primary antibody (Lane 1) and a sample with an unrelated IgG preparation (Lane 2). The immunoprecipitated DNA was analyzed by PCR with primers that were specific for the promoter region of the *BR1* gene (*region 2* in Fig. 4). (B) Quantitative analysis of ChIP using the same conditions as above. For quantification purposes, the intensity of the PCR bands was measured, the values of the negative control (unrelated IgG) were subtracted, and the results were expressed as average related to the input. Error bars indicate standard deviations. $n = 3$. (C) ChIP with anti-Rrp4 antibody using chromatin extracted for 12 and 20 min, as indicated. The immunoprecipitated DNA was analyzed by PCR with primers specific for the *BR1* promoter, as above. (D) ChIP with anti-Brm antibody or with negative control antibody (IgG) using either 0.1% or 0.5% DOC, as indicated above the lanes. The figures below each lane are measurements of the intensity of the bands in arbitrary units.

after different sonication times. Sonication for 4 min failed to give detectable ChIP signals (data not shown) probably due to inefficient extraction of chromatin, whereas the intensities of the ChIP signals at 12 and 20 min were similar (Fig. 3C). Based on this result and on the superior resolution achieved with chromatin sonicated for 20 min (Fig. 2D), we chose 20 min as the optimal sonication time for subsequent experiments.

The signals obtained with the anti-Brm antibody in the ChIP conditions described above, using 1% DOC, were not significant. Brm may not associate with the *BR1* gene promoter, or it is possible that the conditions were not appropriate for the anti-Brm antibody. We sought to find better conditions for Brm detection by carrying out ChIPs in the presence of lower concentrations of DOC. Figure 3D shows that the use of either 0.1% or 0.5% DOC improved the signal-to-noise ratio and gave satisfactory results. We chose to work at 0.5% DOC to keep the background levels as low as possible.

In summary, we have found conditions suitable for ChIP with antibodies against three different proteins. Two of them, anti-Rrp4 and anti-meH3, can work under highly stringent conditions, whereas the third, anti-Brm, requires a DOC concentration below 0.5%. Our results show that the optimal conditions of the ChIP experiment may have to be calibrated for each protein under study.

Combined mapping of the Brm protein by immuno-EM and ChIP

The proposed functions of Brm in nucleosome remodelling at promoter regions (e.g. Soutoglou & Talianidis, 2002) and pre-mRNA splicing (Batsché *et al.*, 2006) prompted us to study the distribution of Brm along the *BR1* gene in the salivary glands of *C. tentans*. For immuno-EM experiments, polytene chromosomes were isolated, incubated with anti-Brm antibody and with a gold-conjugated secondary antibody, and subsequently embedded and sectioned for EM analysis. The *BR* puffs are easily identified in EM images at low magnification as enlarged chromosomal regions of granular texture (Fig. 4A). The transcribed *BR* gene consists of a chromatin loop decorated with growing pre-mRNPs (Fig. 4B). Only partial *BR* gene segments were present in the EM preparations, and there were no full-length *BR* genes (Fig. 4C). This was due to the fact that the loops became truncated during the sectioning. However, each partial segment could be classified as proximal, middle or distal (regions 2, 3 and 4, respectively, in Fig. 4C) based on the morphology of the growing pre-mRNP particles.

The analysis of the distribution of the immunogold labeling (Fig. 4C and D) revealed several interesting findings. Firstly, Brm is not restricted to the proximal promoter region of the gene, being broadly distributed along the *BR* genes. Some examples of labeling in the middle and distal regions of the gene are shown in Fig. 4C and D (see immunogold markers in regions 3 and 4). Second, approximately 30% of the labeling was associated with densely packed chromatin close to the *BR* transcription loops (Fig. 4D). We could not determine by morphological examination whether the dense chromatin labeled by the anti-Brm antibody corresponded to sequences that lay upstream of the *BR* proximal promoter.

Next we carried out ChIP experiments to confirm the association of Brm with the different regions of the *BR1* gene. For this purpose, we analyzed the association of Brm with intergenic sequences located 2 kb upstream of the TATA-box (region 1 in Fig. 4E), at the proximal promoter (region 2) and downstream of the polyadenylation site (region 4). The region 3 could not be quantified by ChIP because of the highly repetitive nature of the sequences in this part of the gene. Figure 4E shows that Brm was undetectable in region 1, was enriched at the proximal promoter, and was present, although at lower levels, at the 3′ end of the gene.

We conclude that Brm does not associate with intergenic sequences upstream of the *BR1* gene promoter and that
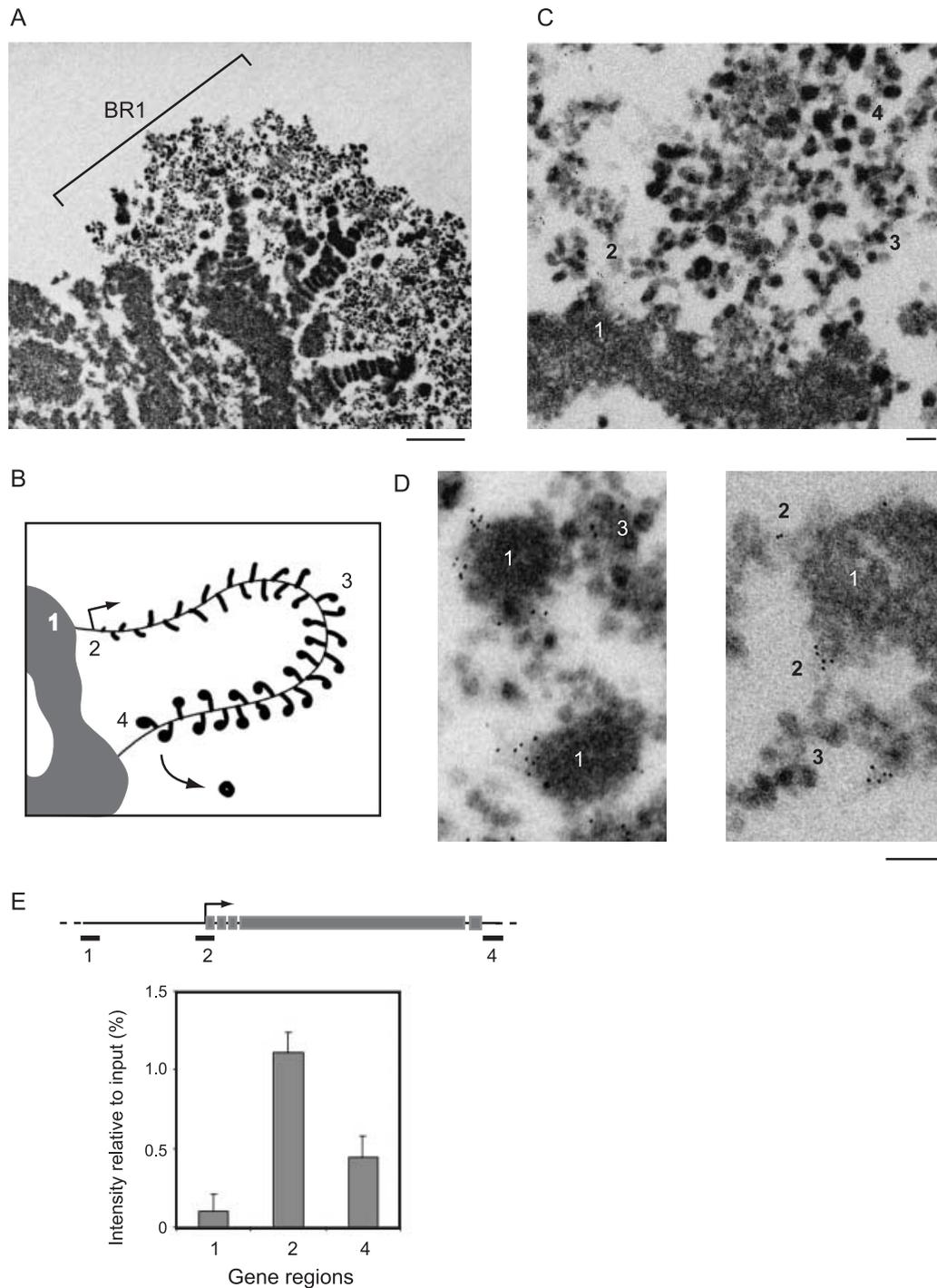
Figure 4. The association of Brm with the *BR* genes studied by combined immuno-EM and ChIP. (A) Electron micrograph of the *BR1* puff at low magnification. The scale bar represents 1 μm. (B) Schematic representation of the structure of the *BR* transcription loops. Densely packed chromatin is indicated as *region 1*. The 5′ and 3′ ends of the gene can be identified by the distinct morphology of the nascent pre-mRNPs in the different parts of the gene. In *region 2*, close to the promoter, the nascent pre-mRNPs appear as fibers of increasing length that become packed into stalked granules towards the middle (*region 3*) and distal (*region 4*) portions of the gene. (C) A section of an isolated chromosome IV immunolabeled with anti-Brm antibodies showing the different portions of a *BR* gene as visualized in the EM. The scale bar represents 100 nm. (D) Two example of immuno-EM labeling with anti-Brm antibody showing the association of immuno-gold markers with the densely packed chromatin (*region 1*) and with the proximal and middle portions of the *BR* gene (*regions 2* and *3*, respectively). The scale bar represents 100 nm. (E) The association of Brm with intergenic sequences located 2 kb upstream of the *BR1* transcription start site (region *1*), with the *BR1* proximal promoter (region *2*) and with the 3′ end of the *BR1* gene (region *4*) analyzed by ChIP. The immunoprecipitations were performed in the presence of 0.5% DOC. Quantification as in Fig. 3. The values indicate average intensity related to the input, in percentage. The error bars are standard deviations. $n = 3$ for regions *1* and *2*. $n = 2$ for region *4*.

the labeling observed on densely packed chromatin is not within 2 kb upstream of the TATA-box. This labeling is preferentially located at the surface of the dense chromatin and is likely to reflect the abundance of Brm associated with promoter sequences. Alternatively, Brm may bind to sequences located more than 2 kb apart from the *BR1* promoter. A more comprehensive mapping upstream or downstream of the *BR1* gene was not feasible due to the lack of available genomic sequences. Based on the data from immuno-EM and ChIP, we also conclude that Brm is not restricted to the promoter of the *BR1* gene but is associated with sequences in the proximal, middle and distal portions of the *BR1* gene.

Conclusions

We have established optimal conditions for ChIP on chromatin extracted from the salivary glands of *C. tentans*, and we present a combination of immuno-EM and ChIP for mapping the association of proteins of interest with the transcribed *BR* genes. The immuno-EM method is restricted to the *BR* genes, and it is limited by the spatial resolution of the morphological analysis. In contrast, the ChIP approach is virtually universal and offers a much higher resolution. Complementing immuno-EM analyses with ChIP experiments thus circumvents many of the limitations of the immuno-EM technique. Observations on the distribution of a protein mapped by immuno-EM analysis of *BR* genes can be compared to ChIP distributions along other genes of *C. tentans* that are not amenable to morphological analysis. The number of available genomic sequences for *C. tentans* is limited, but there is a *C. tentans* EST database that covers over 3000 unique transcripts (Arvestad *et al.*, 2005), providing useful sequence data for ChIP purposes.

ChIP relies on the efficiency of the crosslinking between the protein of interest and the DNA, and these crosslinking efficiencies may vary during the transcription cycle for proteins with highly dynamic interactions. The immuno-EM approach is more direct and does not depend on the proximity of the protein to the DNA. If dynamic molecular rearrangements are expected, as they are, for example, in the case of factors recruited by chromatin components and subsequently transferred to the nascent pre-mRNP or *vice versa*; immuno-EM constitutes a reliable complement to the ChIP analyses.

We have applied immuno-EM and ChIP to study the association of Brm with the *BR1* gene. We have shown that Brm is not located exclusively at the promoter but is broadly distributed along the *BR* transcription unit, which suggests that Brm may have additional roles apart from regulating transcription initiation. This observation is interesting in the light of the proposed involvement of the human Brm in regulation of pre-mRNA processing (Batsché *et al.*, 2006).

Experimental procedures

Antibodies

We used the commercially available antibody ab8580 from Abcam (Cambridge, UK) to detect histone H3 trimethylated at lysine 4. We used a peptide-specific antibody raised against amino acids 143–156 of the *C. tentans* Rrp4 protein (Acc Nr AM709652) to detect Rrp4. The anti-Rrp4 antibody was raised in rabbit and affinity purified according to standard procedures. We used an antibody raised against the rat Brg1 protein (Östlund Farrants *et al.*, 1997) to study Brm. This antibody crossreacts with the Brm proteins of *D. melanogaster* and *C. tentans*. FITC-conjugated and gold-conjugated secondary antibodies were from Jackson ImmunoResearch Laboratories (West Grove, PA, USA). Secondary antibodies conjugated to alkaline phosphatase were from DakoCytomation (Glostrup, Denmark). The negative control antibody *IgG* was a rabbit anti-mouse IgG from DakoCytomation.

Culture of C. tentans

C. tentans was cultured as described by Meyer *et al.* (1983). Salivary glands were isolated from fourth instar larvae. *C. tentans* tissue culture cells were cultivated at 24 °C as previously described (Wyss, 1982).

Chromatin immunoprecipitation

We used a modification of the method described by Legube *et al.* (2006). Thirty salivary glands were fixed for 15 min at room temperature in 1 ml of fixing solution (50 mM Hepes at pH 7.6, 100 mM NaCl, 0.1 mM EDTA, 0.5 mM EGTA, 2% formaldehyde). The glands were transferred to 1 ml of Solution 1 (PBS supplemented with 0.01% Triton X-100 and 125 mM glycine) and incubated for 10 min. The glands were subsequently incubated in 1 ml of Solution 2 (10 mM Tris-HCl at pH 8, 0.25% Triton X-100, 10 mM EDTA, 0.5 mM EGTA) and in 1 ml of Solution 3 (10 mM Tris-HCl at pH 8, 200 mM NaCl, 10 mM EDTA, 0.5 mM EGTA) for 10 min each. The glands were finally resuspended in 300 µl of Sonication Buffer (10 mM Tris-HCl at pH 8, 1 mM EDTA, 0.5 mM EGTA) and sonicated for 4, 8, 12 or 20 min (see Results) in ice-cold water using a Bioruptor (CosmoBio-Diagenode, Liege, Belgium). The sonication was done in pulses of 30 s with 30 s intervals in between. The cold water was replaced every 4 min to avoid heating of the samples. After sonication, sarcosyl was added to a final concentration of 0.5% and the samples were incubated for 10 min at room temperature. The samples were finally dialysed overnight at 4 °C against 10 mM Tris-HCl at pH 8, 1 mM EDTA, 0.5 mM EGTA, 5% glycerol, and stored at –80 °C until use. For immunoprecipitation, 300 µl of sample containing chromatin from 30 glands was precleared for 1 h at 4 °C with a 50% slurry of protein A and protein G coupled to Sepharose. After pre-clearing, the supernatant was supplemented with 1% Triton X-100 and 1% sodium deoxycholate (DOC). The primary antibody was added (final concentration 10 µg/ml) and the immunoprecipitation was allowed to proceed overnight at 4 °C. A mixture of protein A-Sepharose and protein G-Sepharose, previously blocked with salmon sperm DNA and BSA, was added to the sample and the incubation was continued for 90 min. The beads were washed 4 times for 5 min each with RIPA-0.7 buffer (50 mM Hepes at pH 7.6, 1 mM EDTA, 0.7% DOC, 1% NP-40, 0.5 M NaCl), and once with 50 mM Tris-HCl at pH 8, 0.2 mM EDTA. The beads were resuspended in TE containing

0.05% SDS, 0.1 mg/ml RNase A, 0.2 mg/ml proteinase K, and incubated for 3 h at 55 °C. The final step was to reverse the crosslinking at 65 °C overnight. The immunoprecipitated DNA was extracted twice with a mixture of phenol, chloroform and isoamyl alcohol (25:24:1), and once with chloroform. The DNA was finally precipitated with ethanol according to standard procedures (Sambrook & Maniatis, 1989) and resuspended in 100 μl distilled water.

The immunoprecipitated DNA was analyzed by semiquantitative PCR using specific primers and Taq DNA polymerase (Fermentas, Burlington, Ontario, Canada) following the manufacturer's instructions. For each primer pair, the conditions of the PCR reaction were optimized to avoid saturation. Each experiment was analyzed at least twice to confirm the results. The PCR products were resolved in agarose gels, stained with ethidium bromide and quantified using the QuantityOne software (BioRad, Hercules, CA, USA). Results presented are averages from at least three different PCRs from at least two independent immunoprecipitations except for the analysis of region *4* in Fig. 4E, which is an average of results from two PCRs from the same immunoprecipitation. The values provided in the histograms are averages percentages relative to the input after subtraction of negative control values. The following primers were used for the *BR1* gene: region 1-F 5′-ATTCAAG-CACAATTTTCGCA-3′, region 1-R 5′-GTTTGATCAAGT GCGGCTTT-3′, region 2-F 5′-CGCTCGTTTGTTCTGATGAA-3′, region 2-R 5′-TCAGGTTGTTCGGGTCCTAC-3′, region 4-A5-F 5′-AGAGGATGAAGCTCTGTT GGA-3′, region 4-A5-R 5′-GTGAG GTTGCGAATGAAAQCA-3′, region 4-A6-F 5′-TTCGCAACCTCA CAATCAAG-3′, region 4-A6-R 5′-CATTTGTGGATGTAAAGTTCG G-3′. For the experiment shown in Fig. 2D, the following primers were used to amplify fragments of the *hrp65* gene (accession number AJ404654): F1 5′-ATCAGCCAAAACCACAAAGG-3′, R1 5′-TTAAAAATGGCGGACGAAAC-3′, R2 5′-TCTTCGCACACA AGTTCTGG-3′, R3 5′-ATTTCCAGGCATAGTCGCAG-3′.

SDS-PAGE and Western blot analysis

Nuclear protein extracts were prepared from *C. tentans* tissue culture cells as described by Sjölinder *et al*. (2005). For analysis of salivary gland proteins, the salivary glands were dissected and directly boiled in sample buffer containing 8 M. The proteins were separated by SDS-PAGE according to standard procedures. For Western blotting, the proteins were transferred to polyvinylidene fluoride membranes (PVDF; Millipore, Billerica, MA, USA) in tris-glycine buffer supplemented with 0.02% SDS and 4 M urea using a semi-dry electrophoretic transfer cell (BioRad). The secondary antibodies were conjugated to alkaline phosphatase and antibody binding was detected using the NBT/BCIP system according to standard procedures.

Immunofluorescence

Salivary glands were manually dissected and pre-fixed with 2% formaldehyde in TKM buffer (100 mM KCl, 1 mM $MgCl_2$ and 10 mM triethanolamine at pH 7.6) for 2–3 min at 4 °C. The pre-fixed glands were permeabilized with 2% NP-40 in TKM. Chromosomes were isolated by repeatedly pipetting the glands through a siliconized glass micropipette. The released chromosomes were fixed in 4% formaldehyde in TKM for 30 min. The chromosomes were then washed with TKM three times for 5 min each, and then blocked with 3% bovine serum albumin (BSA) in TKM for 30 min. After blocking, the chromosomes were incubated with primary antibodies at 4 °C overnight. The working concentrations of the primary antibodies

were 20 μg/ml for anti-Rrp4, 1 μg/ml for anti-Brm, and 2.5 μg/ml for anti-meH3. The chromosomes were subsequently washed with 0.1% Tween-20 in TKM three times for 10 min each, and incubated with a FITC-conjugated antibody. The specimens were mounted with Vectashield mounting media (Vector) and examined in a LSM510 laser confocal microscope (Carl Zeiss, Oberkochen, Germany).

Immuno-electron microscopy

Polytene chromosomes were isolated and blocked using the same procedure as that used for immunofluorescence. Incubation with the anti-Brm antibody at 1 μg/ml was for 1 h. After primary antibody incubation, the chromosomes were washed and incubated with an anti-rabbit IgG conjugated to 6-nm colloidal gold. Chromosomes processed as negative controls were incubated with secondary antibody only. After the immunolabeling, the chromosomes were post-fixed with 2% glutaraldehyde, dehydrated in ethanol and embedded in Agar 100. Thin sections of plastic-embedded chromosomes were stained with 2% uranyl acetate in 50% ethanol, washed with 50% ethanol and air-dried. The preparations were examined and photographed in a Tecnai G2 transmission electron microscope (FEI Company, Hillsboro, OR, USA) at 80 kV using a US 1000P CCD camera (Gatan, Pleasanton, CA, USA).

Acknowledgements

References

Andrulis, E.D., Werner, J., Nazarian, A., Erdjument-Bromage, H., Tempst, P. and Lis, J.T. (2002) The RNA processing exosome is linked to elongating RNA polymerase II in Drosophila. *Nature* **420**: 837–841.

Armstrong, J.A., Papoulas, O., Daubresse, G., Sperling, A.S., Lis, J.T., Scott, M.P., *et al*. (2006) The Drosophila BRM complex facilitates global transcription by RNA polymerase II. *EMBO J* **21**: 5245–5254.

Arvestad, L., Visa, N., Lundeberg, J., Wieslander, L. and Savolainen, P. (2005) Expressed sequence tags from the midgut and an epithelial cell line of Chironomus tentans: annotation, bioinformatic classification of unknown transcripts and analysis of expression levels. *Insect Mol Biol* **14**: 689–695.

Batsché, E., Yaniv, M. and Muchardt, C. (2006) The human SWI/SNF subunit Brm is a regulator of alternative splicing. *Nat Struct Mol Biol* **13**: 5–7.

Baurén, G. and Wieslander, L. (1994) Splicing of Balbiani ring 1 gene pre-mRNA occurs simultaneously with transcription. *Cell* **76**: 183–192.

Baurén, G., Belikov, S. and Wieslander, L. (1998) Transcriptional termination in the Balbiani ring 1 gene is closely coupled to 3′-end formation and excision of the 3′-terminal intron. *Genes Dev* **12**: 2759–2769.

Daneholt, B. (2001) Assembly and transport of a pre-messenger RNP particle. *Proc Natl Acad Sci USA* **98**: 7012–7017.

Ericsson, C., Mehlin, H., Björkroth, B., Lamb, M.M. and Daneholt, B.

(1989) The ultrastructure of upstream and downstream regions of an active Balbiani ring gene. *Cell* **56**: 631–639.

Ericsson, C., Grossbach, U., Björkroth, B. and Daneholt, B. (1990) Presence of histone H1 on an active Balbiani ring gene. *Cell* **60**: 73–83.

Houseley, J., LaCava, J. and Tollervey, D. (2006) RNA-quality control by the exosome. *Nat Rev Mol Cell Biol* **7**: 529–539.

Khavari, P.A., Peterson, C.L., Tamkun, J.W., Mendel, D.B. and Crabtree, G.R. (1993) BRG1 contains a conserved domain of the SWI2/SNF2 family necessary for normal mitotic growth and transcription. *Nature* **366**: 170–174.

Kouskouti, A. and Talianidis, I. (2005) Histone modifications defining active genes persist after transcriptional and mitotic inactivation. *EMBO J* **24**: 347–357.

Legube, G., McWeeney, S.K., Lercher, M.J. and Akhtar, A. (2006) X-chromosome-wide profiling of MSL-1 distribution and dosage compensation in Drosophila. *Genes Dev* **20**: 871–878.

Maniatis, T. and Reed, R. (2002) An extensive network of coupling among gene expression machines. *Nature* **416**: 499–506.

Meyer, B., Mähr, R., Eppenberger, H.M. and Lezzi, M. (1983) The activity of Balbiani rings 1 and 2 in salivary glands of Chironomus tentans larvae under different modes of development and after pilocarpine treatment. *Dev Biol* **98**: 265–277.

Muchardt, C. and Yaniv, M. (1993) A human homologue of Saccharomyces cerevisiae SNF2/SWI2 and Drosophila brm genes potentiates transcriptional activation by the glucocorticoid receptor. *EMBO J* **12**: 4279–4290.

Orlando, V., Strutt, H. and Paro, R. (1997) Analysis of chromatin structure by in vivo formaldehyde cross-linking. *Methods* **11**: 205–214.

Östlund Farrants, A.K., Blomquist, P., Kwon, H. and Wrange, Ö. (1997) Glucocorticoid receptor-glucocorticoid response element binding stimulates nucleosome disruption by the SWI/SNF complex. *Mol Cell Biol* **17**: 895–905.

Saha, A., Wittmeyer, J. and Cairns, B.R. (2006) Mechanisms for nucleosome movement by ATP-dependent chromatin remodeling complexes. *Results Probl Cell Differ* **41**: 127–148.

Sambrook, J. and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, New York.

Santos-Rosa, H., Schneider, R., Bannister, A.J., Sherriff, J., Bernstein, B.E., Emre, N.C., *et al.* (2002) Active genes are tri-methylated at K4 of histone H3. *Nature* **419**: 407–411.

Sif, S. (2004) ATP-dependent nucleosome remodeling complexes: enzymes tailored to deal with chromatin. *J Cell Biochem* **91**: 1087–1098.

Sjölinder, M., Björk, P., Söderberg, E., Sabri, N., Farrants, A.K. and Visa, N. (2005) The growing pre-mRNA recruits actin and chromatin-modifying factors to transcriptionally active genes. *Genes Dev* **19**: 1871–1884.

Skoglund, U., Andersson, K., Björkroth, B., Lamb, M.M. and Daneholt, B. (1983) Visualization of the formation and transport of a specific hnRNP particle. *Cell* **34**: 847–855.

Solomon, M.J., Larsen, P.L. and Varshavsky, A. (1988) Mapping protein-DNA interactions in vivo with formaldehyde: evidence that histone H4 is retained on a highly transcribed gene. *Cell* **53**: 6470–6474.

Soutoglou, E. and Talianidis, I. (2002) Coordination of PIC assembly and chromatin remodeling during differentiation-induced gene activation. *Science* **295**: 1901–1904.

Visa, N., Izaurralde, E., Ferreira, J., Daneholt, B. and Mattaj, I.W. (1996) A nuclear cap-binding complex binds Balbiani ring pre-mRNA cotranscriptionally and accompanies the ribonucleoprotein particle during nuclear export. *J Cell Biol* **133**: 5–14.

Wieslander, L. (1994) The Balbiani ring multigene family: coding repetitive sequences and evolution of a tissue-specific cell function. *Prog Nucleic Acid Res Mol Biol* **48**: 275–313.

Wyss, C. (1982) Chironomus tentans epithelial cell lines sensitive to ecdysteroids, juvenile hormone, insulin and heat shock. *Exp Cell Res* **139**: 297–307.

7.2 SWI/SNF associates with nascent pre-mRNPs and regulates alternative pre-mRNA processing.

Tyagi A, Ryme J, Brodin D, Ostlund Farrants AK, Visa N.

SWI/SNF Associates with Nascent Pre-mRNPs and Regulates Alternative Pre-mRNA Processing

Anu Tyagi[1], Jessica Ryme[2], David Brodin[3], Ann Kristin Östlund Farrants[2], Neus Visa[1]*

1 Department of Molecular Biology and Functional Genomics, Stockholm University, Stockholm, Sweden, 2 Department of Cell Biology at the Wenner-Gren Institute, Stockholm University, Stockholm, Sweden, 3 Bioinformatics and Expression Analysis Core Facility, Department of Biosciences and Nutrition, Karolinska Institute, Huddinge, Sweden

Abstract

The SWI/SNF chromatin remodeling complexes regulate the transcription of many genes by remodeling nucleosomes at promoter regions. In *Drosophila*, SWI/SNF plays an important role in ecdysone-dependent transcription regulation. Studies in human cells suggest that Brahma (Brm), the ATPase subunit of SWI/SNF, regulates alternative pre-mRNA splicing by modulating transcription elongation rates. We describe, here, experiments that study the association of Brm with transcribed genes in *Chironomus tentans* and *Drosophila melanogaster*, the purpose of which was to further elucidate the mechanisms by which Brm regulates pre-mRNA processing. We show that Brm becomes incorporated into nascent Balbiani ring pre-mRNPs co-transcriptionally and that the human Brm and Brg1 proteins are associated with RNPs. We have analyzed the expression profiles of *D. melanogaster* S2 cells in which the levels of individual SWI/SNF subunits have been reduced by RNA interference, and we show that depletion of SWI/SNF core subunits changes the relative abundance of alternative transcripts from a subset of genes. This observation, and the fact that a fraction of Brm is not associated with chromatin but with nascent pre-mRNPs, suggest that SWI/SNF affects pre-mRNA processing by acting at the RNA level. Ontology enrichment tests indicate that the genes that are regulated post-transcriptionally by SWI/SNF are mostly enzymes and transcription factors that regulate postembryonic developmental processes. In summary, the data suggest that SWI/SNF becomes incorporated into nascent pre-mRNPs and acts post-transcriptionally to regulate not only the amount of mRNA synthesized from a given promoter but also the type of alternative transcript produced.

Introduction

Messenger RNAs (mRNAs) are synthesized in eukaryotic cells as precursor RNA molecules (pre-mRNAs), which are then assembled into ribonucleoprotein complexes (pre-mRNPs) during transcription. The newly synthesized pre-mRNAs are modified by capping, splicing and 3′-end maturation reactions that involve cleavage and polyadenylation of the 3′-end of the transcript. The existence of alternative splicing and alternative polyadenylation sites in many pre-mRNAs is a major source of protein variability, and the regulation of pre-mRNA processing is a major mode of genetic regulation [1,2]. Many observations during the last decade have indicated that some of the mechanisms that regulate the processing of the pre-mRNA are related to the transcription process itself, and that chromatin dynamics, transcription and pre-mRNA processing are functionally connected [reviewed in 3].

Transcription can influence the usage of alternative splice sites in the pre-mRNA [4] through several mechanisms. Promoter-specific coregulators can recruit splicing factors to transcribed genes or they can themselves play dual roles in the regulation of transcription initiation and alternative splicing [review by 5]. For example, the hnRNP-like protein CoAA, which coactivates the transcription of multiple genes regulated by steroid hormones [6],

is recruited to the promoters of its target genes by the coactivator TRBP/NCoA6 and regulates splice site selection [7]. The CAPER proteins, members of the U2AF65 family, and the polyC-RNA-binding protein 1, PCBP1, are further examples of transcriptional coactivators that influence alternative splicing in a promoter-dependent manner [8,9]. The RNA polymerase itself can also participate in the recruitment of pre-mRNA processing factors. A reported case is that of SRp20, a member of the SR protein family of splicing regulators. SRp20 is recruited to transcribed genes through interaction with the C-terminal domain (CTD) of the large subunit of RNA polymerase II [10]. It has been proposed that these proteins are recruited to the promoter, travel along the gene with the transcription machinery, and are eventually delivered to the nascent pre-mRNA for splicing regulation [5].

Low transcription elongation rates favor the usage of proximal splice sites by increasing the time during which the proximal sites are exposed to the splicing machinery before more distal sequences are synthesized. This provides a further mechanism by which transcription influences pre-mRNA processing [reviewed in 11].

A recent study has shown that the human Brahma (hBrm) protein, a chromatin remodeling factor, regulates the alternative splicing of several genes in human cells [12]. Overexpression and depletion experiments have shown that hBrm, together with the

Author Summary

Genetic programs in multicellular organisms often involve different levels of regulation. The expression of many genes is regulated by factors that remodel the structure of the chromatin at the promoter. SWI/SNF is one such factor, and it is highly conserved in eukaryotes. Studies in human cells suggest that Brahma, the catalytic subunit of SWI/SNF, regulates the processing of precursor mRNAs (pre-mRNAs). We have studied Brahma in two insect model systems to further elucidate the mechanisms by which SWI/SNF regulates gene expression. We show that depletion of SWI/SNF subunits changes the relative abundances of alternative transcripts from a subset of pre-mRNAs that code for proteins that regulate the postembryonic development of the flies. We also show that a fraction of Brahma is not associated with chromatin but with nascent pre-mRNPs—both in insects and mammals—which suggests that SWI/SNF acts at the RNA level to regulate pre-mRNA processing. These findings illustrate the dual role of a chromatin remodelling factor; SWI/SNF acts both at the transcriptional level and post-transcriptionally to regulate not only the amount of mRNA synthesized from a given promoter but also the type of alternative transcript produced.

mRNA-binding protein Sam68, favors the accumulation of RNA pol II at specific gene positions and decreases the elongation rate of the RNA pol II. These effects favor the inclusion of variable exons with weak splice sites [12].

The hBrm protein and its paralog hBrg1 are the ATPase subunits of the SWI/SNF chromatin remodeling complexes in human cells [reviewed in 13]. The SWI/SNF complexes regulate the transcription of many genes in mammalian cells by remodeling nucleosomes at promoter regions [reviewed in 14]. The Brm protein in *Drosophila melanogaster* (dBrm) is associated with transcribed loci in the polytene chromosomes and plays an important role in ecdysone-dependent transcription regulation [15,16]. Two types of dBrm-containing complexes, BAP and PBAP, are present in *Drosophila*. They share seven core subunits, including dBrm, but differ in the presence of additional signature subunits [17 and references therein].

We have analyzed *in situ* the association of Brm with the actively transcribed Balbiani ring (BR) genes of the dipteran *Chironomus tentans* (*C. tentans*) in order to obtain further insight into the mechanisms by which Brm influences pre-mRNA splicing. The BRs are giant puffs that form by active transcription of the BR genes in the polytene chromosomes of the salivary gland cells [18]. The BR pre-mRNAs synthesized in the BR1 and BR2 puffs have all the features of typical pre-mRNAs. They are capped at the 5′-end [19], spliced, cleaved and polyadenylated at the 3′-end [20], released from the chromosome, and finally exported to the cytoplasm [reviewed in 21]. It is possible to visualize using transmission electron microscopy (TEM) how the BR pre-mRNPs are synthesized along the BR genes [22], and it is possible to study the association of defined proteins with nascent BR pre-mRNP particles *in situ* using specific antibodies.

The Brm protein of *C. tentans* (ctBrm) is associated with the BR puffs and is widely distributed along the active BR genes, as shown by immuno-electron microscopy (immuno-EM) and chromatin immunoprecipitation (ChIP), which suggests that ctBrm has further roles in addition to that of regulating transcription initiation [23]. We have analyzed the location of ctBrm in the actively transcribed BR genes in more detail, and we have shown

that a fraction of Brm is associated with the nascent transcripts in both insect and mammalian cells. We have also determined whether Brm plays a role in pre-mRNA processing in insects and we have analyzed the expression profiles of *D. melanogaster* cells in which the levels of dBrm have been reduced by RNA interference (RNAi). We show that depletion of dBrm affects not only the splicing but also the usage of alternative polyadenylation sites.

Results

ctBrm Is Associated with the BR Genes

We used three different antibodies against ctBrm to study the association of this protein with the BR genes of *C. tentans*. The first one, Ab1, was raised against the rat Brg1 protein [24]. The second antibody, Ab2, was raised against the C-terminal part of the Ct-BRM protein (Figure S1). The third one, Ab3, was raised against dBrm and has been characterized by Zraly et al. [25]. The specificity of the antibodies was tested by Western blot against nuclear protein extracts prepared from *C. tentans* cultured cells. The three antibodies recognized one major band of approximate molecular mass 200 kDa, the expected molecular mass of ctBrm (Figure S2A). The same band was present in protein extracts prepared from larval salivary glands (Figure S2B). Moreover, the ctBrm protein immunoprecipitated by Ab1 was recognized by Ab2 and by Ab3 (Figure S2C). We concluded that the three antibodies recognized the same protein, ctBrm.

Immunofluorescent staining of isolated polytene chromosomes gave a banded pattern from all three antibodies, and the antibodies stained many chromosomal bands. The actively transcribed BR puffs were among the most intensely stained loci (Figures 1A–C and S3A–D). In some cases, the chromosomes were co-stained with a mAb against Hrp45, an hnRNP protein used as a marker to visualize the BRs in chromosome IV (Figures 1A and S3C–D).

Preparations of isolated polytene chromosomes were digested with RNase A before immunostaining to determine whether the association of Brm with the chromosomes was mediated by RNA (Figure 1D). The chromosomes were co-stained with Y12, a mAb against core snRNP proteins, to monitor the effect of the RNase treatment. Control chromosomes were incubated in parallel in the absence of RNase A. The snRNP staining (*red* in Figure 1D) was abolished by the RNase treatment, as expected. The intensity of the Brm staining (*green* in Figure 1D) was reduced. However, a part of the Brm staining was resistant to the RNase treatment. These results suggest that there are two modes of interaction of Brm with the chromosomes. One mode is independent of the presence of RNA and may be explained by a direct association of Brm with the chromatin. The other mode requires RNA.

We have previously mapped the association of ctBrm with the BR1 gene by immuno-EM using the anti-Brg1 antibody, Ab1, and we have shown that ctBrm is widely distributed along the entire BR1 gene [23]. We have here confirmed this observation by extending the immuno-EM analysis to BR1 and BR2, where we have used the *C. tentans*-specific antibody Ab2. The results obtained are summarized in Figure 2. The BR1 and BR2 genes are approximately 40 kb long and they are transcribed simultaneously by several RNA polymerases. The different regions of the gene show specific morphological features due to the progressive growth and assembly of the nascent BR pre-mRNPs (Figure 2B). Full-length genes are not available in the sections used for TEM, but partial gene segments are observed (Figure 2C–D). The gene segments can be classified into proximal, middle and distal segments, based on the morphology of the nascent pre-mRNPs. The pre-mRNPs in the proximal region appear as growing fibers
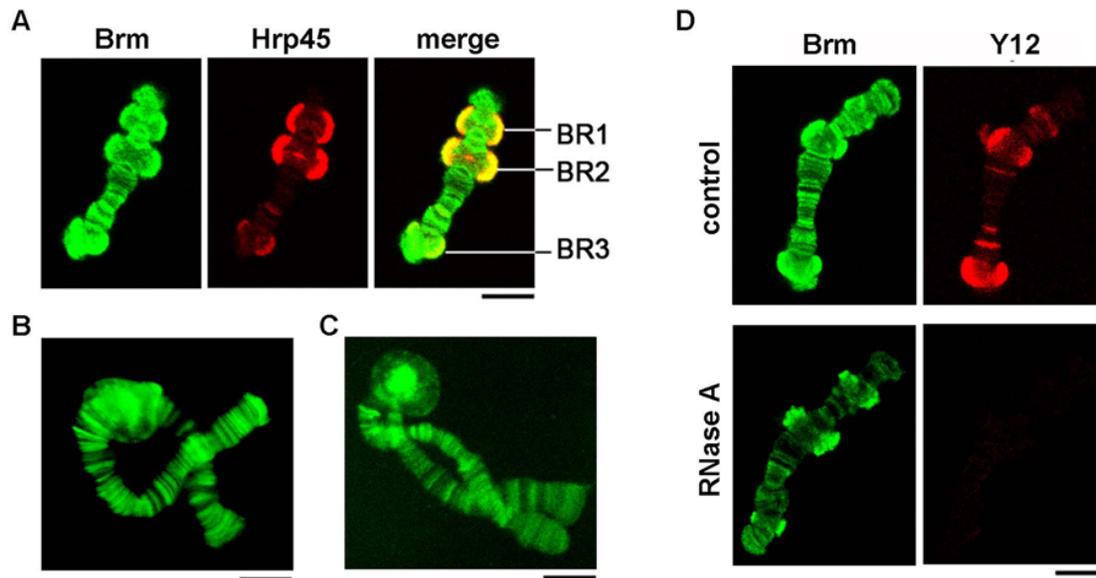
Figure 1. The association of ctBrm with the polytene chromosomes studied by immunofluorescence. (A) The distribution of ctBrm in chromosome IV. Confocal sections of isolated polytene chromosomes stained with the Ab1 antibody and with mAb 2E4 against the mRNA-binding protein Hrp45. CtBrm was associated with multiple loci, including the Balbiani ring puffs BR1, BR2 and BR3 in chromosome IV. (B–C) The association of ctBrm with other chromosomes revealed with Ab2 (B) and Ab3 (C). Both antibodies gave a banded pattern. (D) The association of ctBrm with the chromosomes is partly dependent on interactions with RNA. Polytene chromosomes were isolated and treated with RNase A before fixation and immunostaining (*RNase A*). Control chromosomes were incubated in parallel without RNase A (*control*). The chromosomes were double-stained with Ab1 (*green*) and mAb Y12 against snRNP proteins (*red*). The secondary antibodies were conjugated to FITC and Texas Red, respectively. The scale bars represent 10 μm.
doi:10.1371/journal.pgen.1000470.g001

with increasing length, whereas the pre-mRNPs in the middle and distal regions appear as stalked granules of increasing diameter.

We isolated chromosome IV from salivary glands and stained the isolated chromosomes with either Ab1 or Ab2. The antibody-binding sites were revealed using a gold-conjugated secondary antibody. Control chromosomes were processed in parallel in order to assess the specificity of the immunolabeling. The chromosomes were embedded in plastic after the immunolabeling and sectioned for TEM analysis. Photographs were taken at random positions, and each gold particle was classified according to its association with a proximal, middle or distal gene segment (Figures 2C and 2D). ctBrm was present on the proximal, middle and distal regions of the BR genes. Similar results were obtained with Ab1 and Ab2 (Figure 2E). These results confirm that ctBrm is widely distributed along the BR genes.

ctBrm Is Associated with Nascent Pre-mRNPs

We wanted to determine whether the ctBrm protein located at the BR genes was associated with the chromatin or with the nascent BR pre-mRNPs. Detailed analysis of immuno-EM data provides enough resolution to distinguish between labeling of the pre-mRNPs and labeling of the chromatin, as shown by Wetterberg et al. [26].

We selected 60 distal BR segments in which the relative positions of the chromatin axis and the nascent pre-mRNPs could be identified, and we determined for each segment whether the gold markers were close to the chromatin (within 50 nm of the axis) or distant from the chromatin (more than 50 nm from the axis). The dimensions of the antibodies mean that this latter group contains only gold markers associated with BR pre-mRNP. In contrast, the markers close to the chromatin may label ctBrm molecules bound to the stalk of the pre-mRNP, bound to the

chromatin, or bound to the transcription machinery. The gold markers were distant from the chromatin in 20 out of 60 analyzed cases (33%). A fraction of ctBrm was associated with the nascent pre-mRNPs and this association was confirmed with all three anti-Brm antibodies, as shown in Figure 3. This result agrees with the results of the RNase A digestion experiments shown in Figure 1D. We conclude that a fraction of ctBrm is associated with the nascent BR pre-mRNPs and is not in contact with the chromatin.

We used a cell fractionation assay to confirm the association of ctBrm with RNPs. We isolated nuclei from *C. tentans* tissue culture cells and prepared two types of protein extracts (Figure 4A). One of the extracts contained soluble nuclear proteins (*soluble*) and the other extract contained proteins bound to the chromosomes via RNA (*chromosomal RNP*). These proteins could be released by RNase A digestion. The proteins in each fraction were resolved by SDS/PAGE and analyzed by Western blotting. Coomassie Blue staining showed that each fraction contained a different set of proteins (Figure 4B). Antibodies against Hrp36, an abundant member of the hnRNP A family, histone H3 and the TATA-binding protein (TBP) were used as controls to assess the quality of the fractions. Hrp36 was present in both the soluble and the chromosomal RNP preparations, as expected, whereas histone H3 and TBP were found in the soluble fraction and in the pellet (which contained nuclear components that were not extracted by RNase, such as chromatin and the nuclear envelope). ctBrm was present in the soluble nuclear fraction and in the chromosomal RNP fraction (Figure 4C).

The low abundance of ctBrm in the pellet (*lane 3*, Figure 4C) was unexpected considering that a fraction of ctBrm remained associated with the chromosomes after RNase digestion (Figure 1D). The signal in the pellet was considerably increased by using a sample buffer supplemented with 8 M urea, as shown in
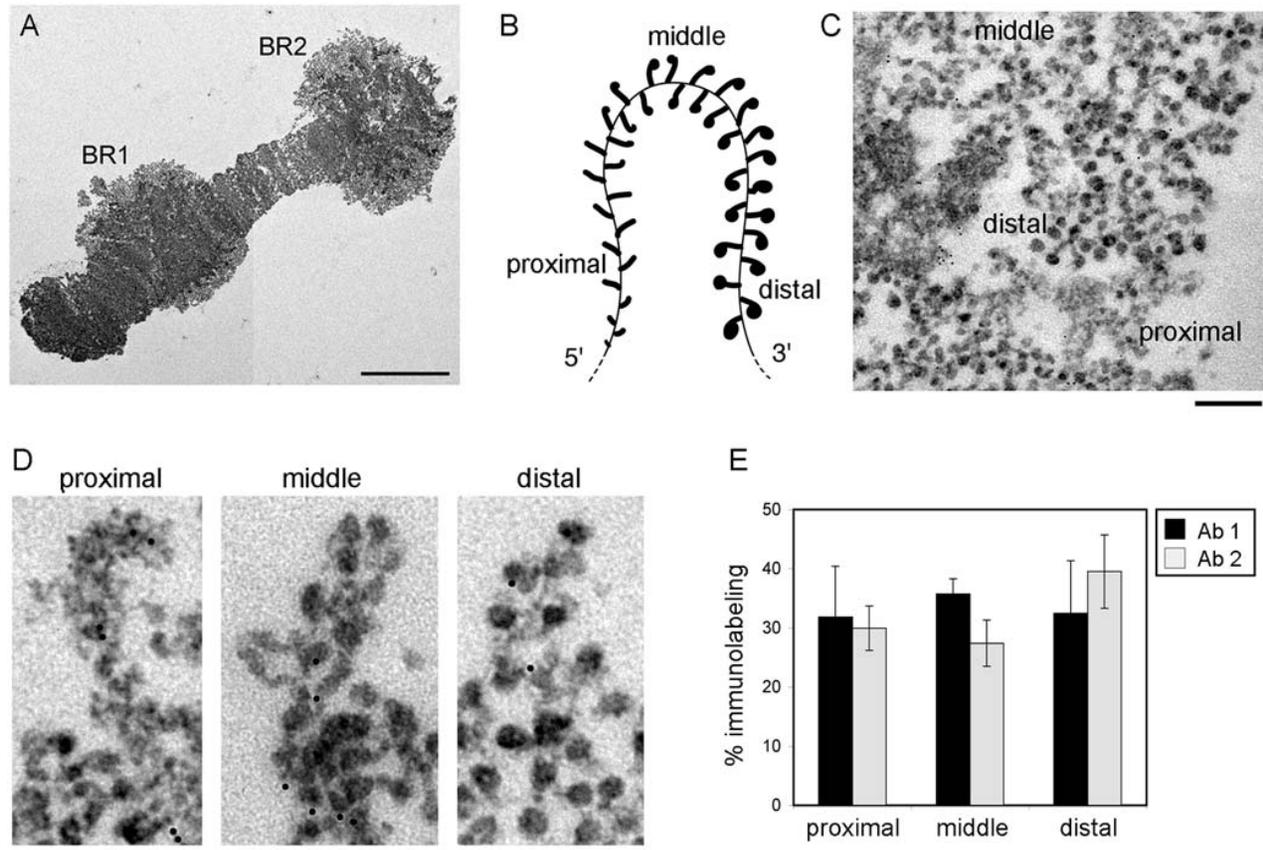
Figure 2. The distribution of ctBrm along the active BR gene analyzed by immuno-EM. (A) Micrograph of a portion of chromosome IV showing the BR1 and BR2 puffs. The scale bar represents 10 μm. (B) Scheme of the active BR transcription unit illustrating the distinct morphology of the growing pre-mRNPs in the proximal, middle, and distal portions of the BR gene. (C) Immuno-electron microscopic analysis of ctBrm distribution in the active BR transcription unit. The image shows a section of a chromosome stained with Ab2 where the proximal, middle and distal portions of the BR gene can be recognized. The bar represents 200 nm. (D) Examples of immuno-EM labeling using Ab2 in the proximal, middle and distal portions of the gene, as indicated. The bar represents 100 nm. (E) Quantitative analysis of the distribution of ctBrm along the BR gene. Comparison of the distribution of gold markers obtained with Ab1 and Ab2. The numbers of immuno-gold markers located in each portion of the gene were counted, and the distribution of labeling was expressed as the percentage of immuno-gold markers in each portion of the gene. Four chromosomes were analyzed for Ab1 and two for Ab2. The error bars are standard deviations. The total number of gold markers analyzed was 5 765 for Ab1 and 1124 for Ab2.

doi:10.1371/journal.pgen.1000470.g002

the bottom panel of Figure 4C. This observation is consistent with the immunofluorescence experiments and indicates that a fraction of ctBrm is highly insoluble.

We repeated the fractionation experiments using *Drosophila* S2 cells (Figure 4D) and human HeLa cells (Figure 4E). The mRNA-binding proteins Hrp59 and SAP155 were used as RNP controls in *Drosophila* and human extracts, respectively. In S2 cells, dBrm was clearly present in the soluble and chromosomal RNP fractions. In HeLa cells, hBrm and hBrg1 were also present in both soluble and chromosomal fractions. We analyzed the presence of other core SWI/SNF subunits in the fractions from HeLa cells. All the analyzed subunits were present in the soluble fraction, as expected, and all of them were also present in the chromosomal RNP fraction (Figure 4E).

In summary, we conclude that a fraction of the Brm protein is associated with chromosomal RNPs and that the association of Brm with the RNPs is conserved from insects to mammals. Our results also suggest that hBrm and hBrg1 are not bound to nascent RNPs as individual proteins in human cells, but as components of SWI/SNF complexes.

ctBrm Interacts with snRNP Complexes

The hBrm and hBrg1 proteins interact with snRNPs [12,27]. We showed that they interact with snRNPs in HeLa cells using immunoprecipitation experiments with the Y12 antibody against core snRNP proteins (Figure 4F). We showed also that ctBrm interacts with snRNP complexes in *C. tentans* (Figure 4G). We prepared a soluble RNP extract, immunoprecipitated snRNPs with Y12, and probed the immunoprecipitated proteins with anti-Brm antibodies. ctBrm was co-immunoprecipiated with snRNPs in the soluble fraction (Figure 4G, lane 4). To determine whether the interaction of ctBrm with snRNPs is a direct protein-protein interaction or whether it requires RNA, the immunoprecipitation was carried out as above and the bound material was treated with RNase A before elution. As shown in Figure 4G, *lane 5*, the signal intensity was significantly reduced in the RNase-treated sample. We assessed the specificity of the interaction by re-probing the blot with an antibody against TBP, a protein that is not expected to interact with snRNPs. We conclude that ctBrm is associated with snRNPs and that the association is RNA-dependent.
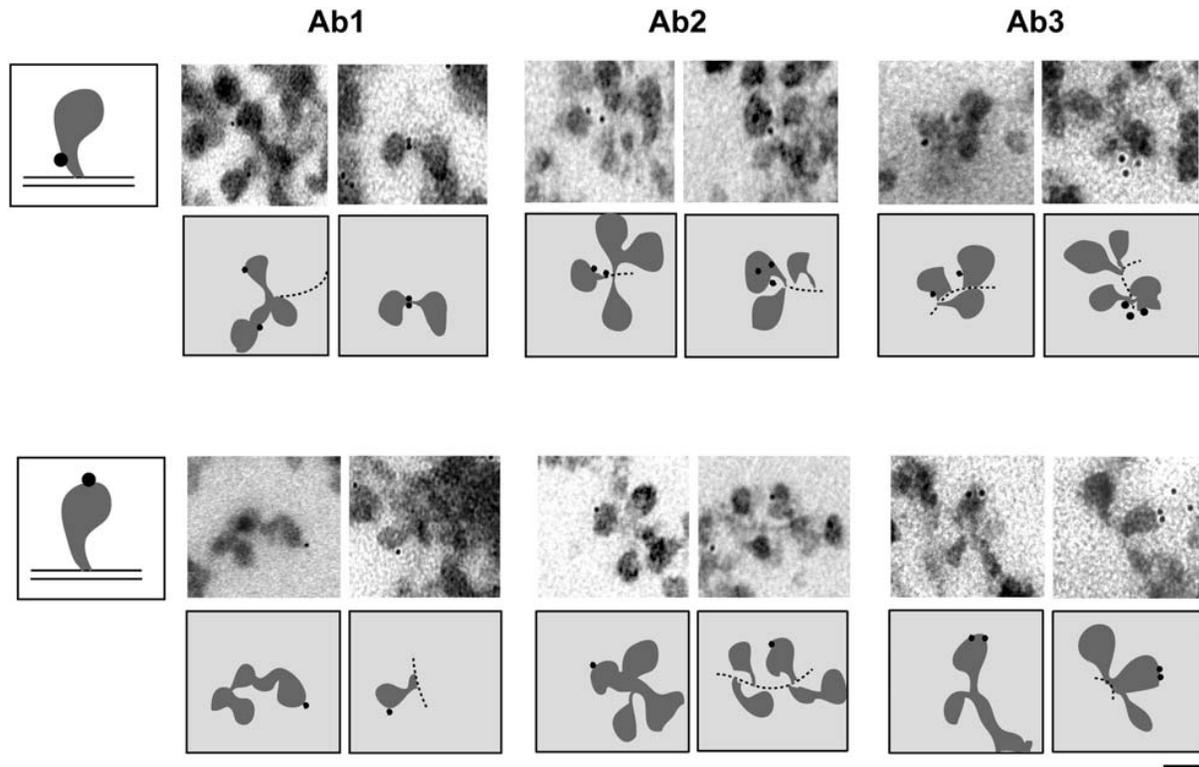
Figure 3. ctBrm is associated with nascent BR pre-mRNP particles. Polytene chromosomes were isolated and immunolabeled with either Ab1, Ab2 or Ab3 against ctBrm, as indicated in the figure. The figure shows four examples of immuno-gold labeling with each of the antibodies. Schematic interpretations of the images are provided under each micrograph. In the top part of the figure, the immuno-gold markers are located close to the chromatin axis (*dotted line*). The examples in the bottom part of the image show nascent BR pre-mRNPs with labeling in the globular domain, far from the chromatin. The single diagrams on the far left at the top and bottom illustrate BR pre-mRNPs with labeling close to the chromatin axis (*top*) and in the RNP globular domain (*bottom*). The scale bar represents 50 nm.
doi:10.1371/journal.pgen.1000470.g003

dBrm Affects Pre-mRNA Processing

We next wanted to determine whether Brm affects pre-mRNA processing in *Drosophila*. Moshkin and coworkers have determined the expression profiles of *Drosophila* S2 cells after depletion of individual SWI/SNF subunits by RNAi and microarray hybridization using the Affymetrix *Drosophila* Genome 2 arrays. Three independent RNAi experiments followed by RNA extraction and microarray hybridization were carried out for dBrm, and six independent experiments were carried out for mock-treated cells [28]. The data from these experiments is available at Array Express, E-TABM-169 (http://www.ebi.ac.uk/microarray-as/aew/). We investigated the effects of dBrm depletion on the relative abundances of alternatively spliced transcripts by mining the E-TABM-169 data and selecting those genes that were represented by more than one probe set in the *Drosophila* Genome 2 arrays (974 genes). In many cases, the multiple probe sets targeted different parts of the same transcript, pseudogenes or alternative transcripts derived from alternative promoters of the same gene. We found evidence that SWI/SNF regulates the activity of many gene promoters, as expected (not shown). We could also identify genes for which the multiple probe sets targeted transcripts that had originated by alternative splicing or alternative polyadenylation of a single pre-mRNA. We selected those genes that displayed changed expression levels specific for at least one transcript with $p<0.02$. Fifteen of these genes showed transcript-specific expression changes in the dBrm-depleted cells (Table 1). We then used the annotations available at FlyBase (http://flybase.

bio.indiana.edu/) to analyze the qualitative differences between the transcripts affected. The transcripts affected show differences in their patterns of alternative, including the use of alternative 3′ slice sites, exon skipping and intron retention (Figure S4). However, the most striking observation was that the processing of the affected transcripts also involved the alternative use of polyadenlation signals, which suggests that dBrm influences not only the splicing but also the formation of the 3′-end of the transcripts.

We validated the microarray results by silencing the expression of dBrm in S2 cells and analysing the expression of four selected genes in which the absence of dBrm affected pre-mRNA processing in different ways, according to the microarray experiments. Mock-treated cells and control cells treated with dsRNA for GFP were analyzed in parallel to assess the specificity of the depletion effects. The levels of dBrm RNA and protein were significantly reduced after 4 days of treatment with dsRNA, as shown by RT-PCR and Western blot, respectively (Figure 5A–B). We designed PCR primers for each of the selected genes in order to amplify specific transcripts and we analyzed the effects of dBrm depletion by RT-PCR. The results of the RT-PCR analyses agreed with the microarray data (Table 1, Figures 5C and S5). In summary, depletion of dBrm affects the relative abundances of alternatively spliced and/or alternatively polyadenylated transcripts.

Are the effects of SWI/SNF depletion on pre-mRNA processing direct or indirect? Pre-mRNA splicing often occurs co-transcrip-
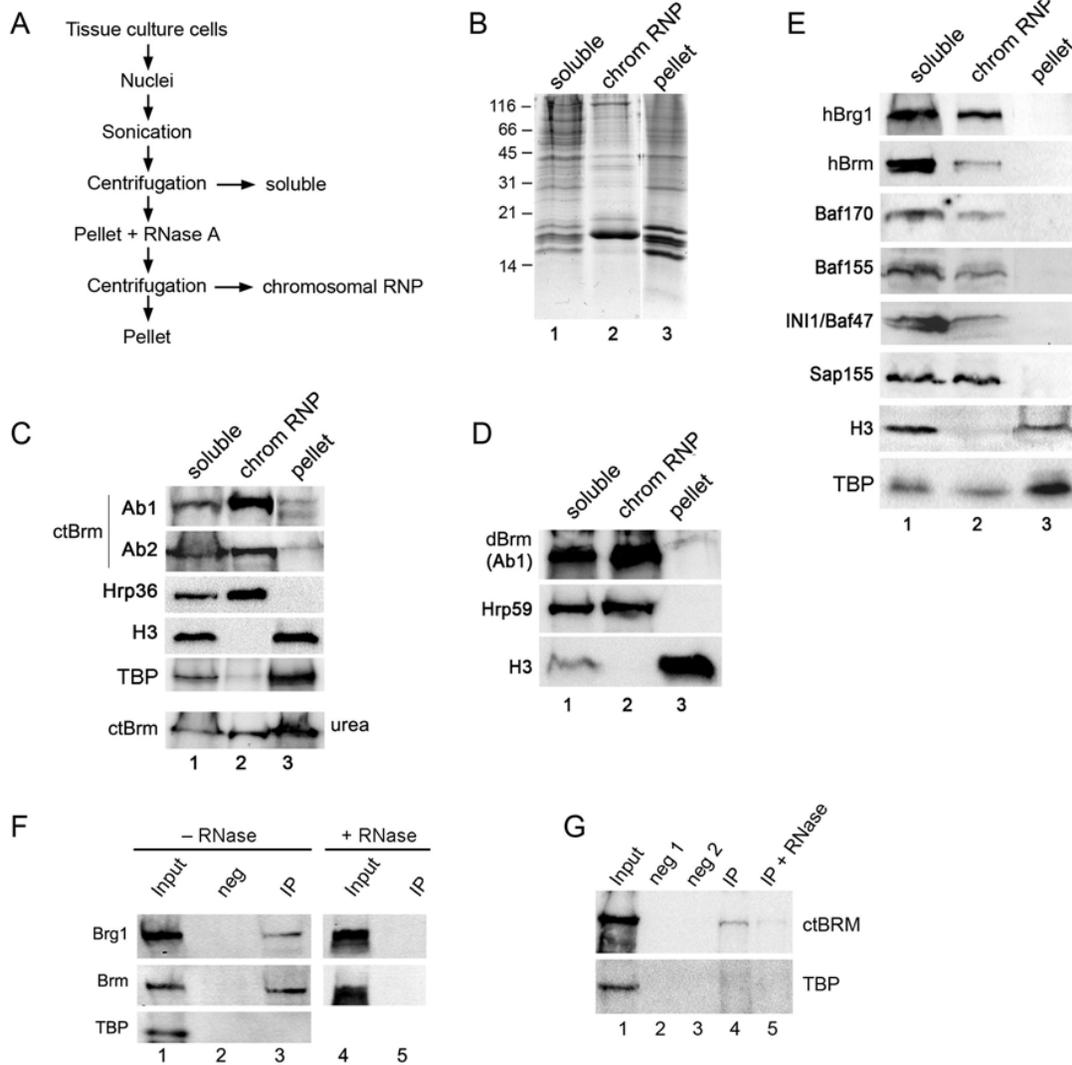
Figure 4. The association of Brm with different nuclear fractions in insect and mammalian cells. (A) Schematic description of the biochemical fractionation scheme. The experimental details are described in the Materials and Methods section. (B) The soluble fraction (*lane 1*), the chromosomal RNP fraction (*lane 2*) and the residual pellet (*lane 3*) prepared from *C. tentans* tissue culture cells were resolved by SDS-PAGE and stained with Coomassie Brilliant Blue to show the protein complexity of each fraction. (C) Nuclear fractions prepared from *C. tentans* tissue culture cells as in (B) were probed with antibodies Ab1 and Ab2 against ctBrm. Antibodies against the hnRNP protein Hrp36, histone H3 and TBP were used in parallel as controls for the composition of the fractions. A significant part of ctBrm was present in the chromosomal RNP fraction (*lane 2*). The presence of ctBRM in the pellet was best revealed in samples dissolved in sample buffer supplemented with 8 M urea (*bottom panel*). (D) Nuclear extracts were prepared from *D. melanogaster* S2 cells following the same fractionation scheme as above. In this case, an antibody against the hnRNP protein Hrp59 was used as a control for the RNP fraction. (E) HeLa cells were fractionated as above and probed with antibodies against different core subunits of SWI/SNF, as indicated. An antibody against Sap155 was used as a control for the RNP fraction. SWI/SNF factors associated with chromosomal RNPs also in HeLa cells. (F) Co-immunoprecipitation of hBrm and hBrg1 with snRNPs. Soluble RNP nuclear extracts were prepared from HeLa cells, treated with or without RNase A as indicated, and used for immunoprecipitation with the Y12 antibody against snRNP proteins (lanes 1–5). The immunoprecipitated proteins were probed by Western blot as indicated to the left. Lanes 2 (*neg*) is a negative control immunoprecipitation processed in parallel. hBrg1 and hBrm are bound to snRNPs (*lane 3*) but the interactions are lost in the RNase-digested extracts (*lane 5*). (G) Soluble RNP nuclear extracts were prepared from *C. tentans* cultured cells and used for immunoprecipitation with the Y12 antibody. An interaction between ctBrm and snRNPs was detected by Western blot with the anti-ctBrm antibody (*lanes 4*). RNase A digestion of the beads before elution drastically reduced the interaction between ctBrm and snRNPs (*lane 5*). Lanes 2 and 3 are two negative control reactions processed in parallel. The blot was reprobed with the anti-TBP antibody.
doi:10.1371/journal.pgen.1000470.g004

tionally [20,29]. We thus asked whether dBrm was associated with the gene regions involved in the alternative processing events that were affected by dBrm depletion, and we carried out chromatin immunoprecipitation (ChIP) experiments to analyze the association of dBrm with the three genes shown in Figure 5. ChIP can

detect proteins that are bound to the DNA as well as proteins associated with the nascent pre-mRNA [see for example 30]. For each gene analyzed, we used primer-pairs to detect the proximal promoter, the internal region affected by the alternative processing and the 3′-end of the gene (Figure S6). dBrm associated with all

Table 1. Transcripts affected by dBrm depletion in S2 cells[1].

Gene	Transcript Detected[2]	Ratio [3]	p-value[3]	Effect
CG11154 (ATPsyn-beta)	CG11154-RA	0.6	0.004	A down
	CG11154-RA+RB	0.8	0.089	
CG11491 (br)	CG11491-RA	0.7	0.218	
	CG11491-RF	0.1	0.001	RF down
	CG11491-RD	1.6	0.182	
CG12052 (lola)	CG12052-RL	1.5	0.004	RL up
	CG12052-RJ	0.8	0.254	
	CG12052-RF	2.2	0.133	
	CG12052-RS+RZ	1.5	0.050	
	CG12052-RJ	0.9	0.463	
	CG12052-RA	1.5	0.001	RA up
CG18251 (Msp-300)	CG18251-RB	7.3	0.001	RB up
	CG18251-RE	1.0	0.480	
CG19086	CG1906-RB	2.1	0.120	
	CG1906- all	2.2	0.152	
	CG1906-RE	2.8	0.011	RE up
CG2249	CG2249-RA+RB	0.8	0.015	RB down
	CG2249-RA	1.0	0.488	
CG32491 (mod(mdg4))	CG32491-RK	1.1	0.373	
	CG32491-RT	0.9	0.377	
	CG32491-RC	1.3	0.200	
	CG32491-RA	0.5	0.009	RA down
	CG32491-RA+RF	0.4	0.009	RA/RF down
	CG32491-RN	0.9	0.250	
	CG32491-RY	1.3	0.190	
	CG32491-RF	0.6	0.012	RF down
	CG32491-RD	0.5	0.013	RD down
CG3665 (Fas2)	CG3665-RA+RB+RC	2.6	0.166	
	CG3665-RC	2.9	0.002	RC up
	CG3665-RA	0.7	0.204	
CG4859 (Mmp1)	CG4859-RC	4.0	0.006	RC up
	CG4859-RC+RD	3.5	0.012	
	CG4859-RD	1.7	0.117	
CG5174	CG5174- all	1.4	0.174	
	CG5174-RK	2.1	0.015	RK up
	CG5174-RA+RB+RH+RI	1.2	0.333	
CG6899 (Ptp4E)	CG6899-RB	1.2	0.317	
	CG6899-RA+RB	1.6	0.109	
	CG6899-RA	1.9	0.016	RA up
CG8092[4]	CG8092-RA+RB	0.7	0.081	
	CG8092-RB	0.5	0.008	RB down
	CG8092-RA	1.5	0.200	
CG8421 (Asph) [4]	CG8421-RD+RE	1.0	0.455	
	CG8421-RA+RB+RD	2.8	0.003	RA/RB up
CG8676 (Hr39)	CG8676-all	2.3	0.001	RA up
	CG8676-RA+RB+RD	2.2	0.0005	RA up
	CG8676-RC	1.4	0.212	
CG9380[4]	CG9380-RB	17.2	0.0001	RB up
	CG9380-RC	6.4	0.005	RC up

Table 1. Cont.

regions of the three genes (lanes 2, 6 and 10 in Figure 6). We used an antibody against the C-terminus of the largest subunit of RNA pol-II as a positive control for the ChIP reactions (lanes 3, 7 and 11 in Figure 6). An unrelated anti-rabbit antibody was used as a negative control (lanes 4, 8 and 12 in Figure 6). Additional controls were carried out by analyzing the association of dBrm with the actin gene, a housekeeping gene whose expression is not regulated

by SWI/SNF. The RNA pol-II (lane 15) associated with the actin gene while the dBrm did not (lane 14 in Figure 6). An intergenic region located far from any annotated genes was devoid of both dBrm and RNA pol-II.

We next asked whether dBrm alone or the entire BAP/PBAP complex is responsible for the effects detected at the level of pre-mRNA processing. We mined the data from the E-TABM-169
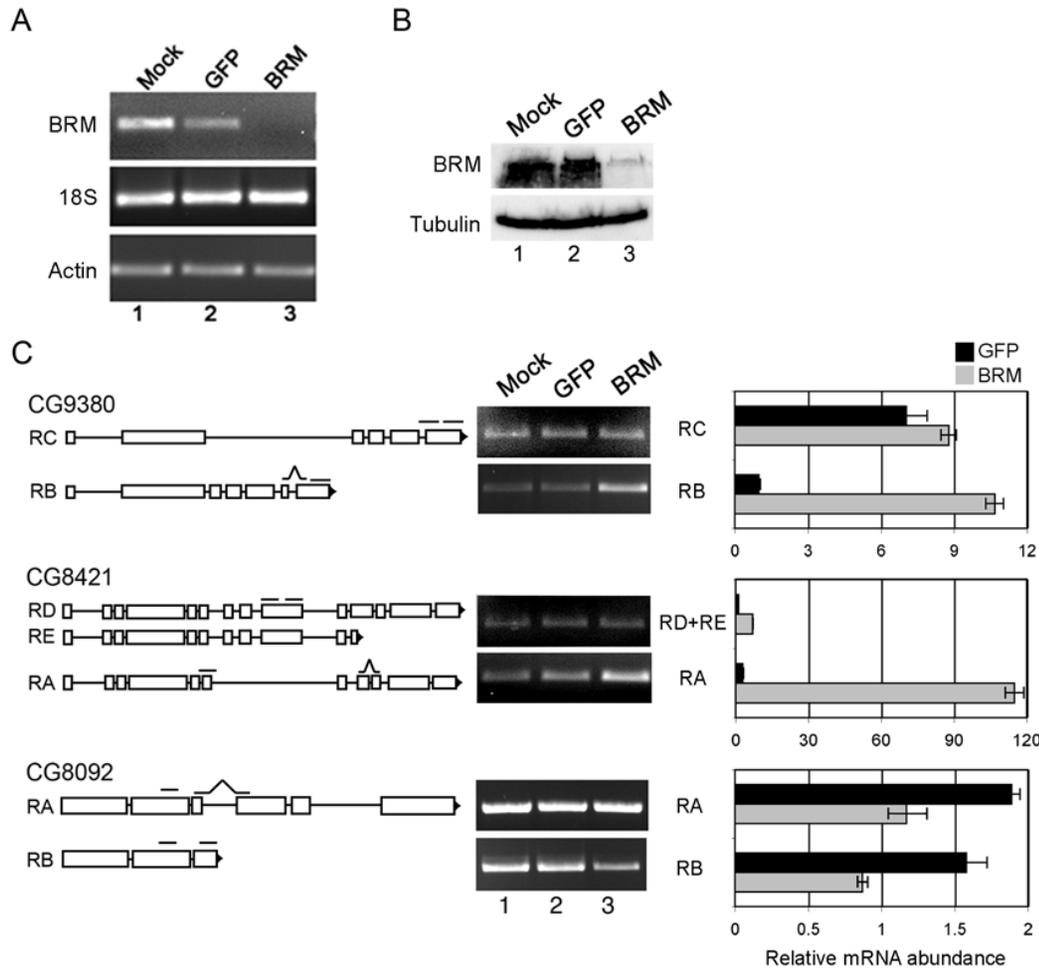
Figure 5. Depletion of dBrm changes the relative abundances of alternatively processed transcripts in S2 cells. S2 cells were treated with dsRNA for dBrm for 4 days. Control experiments were carried out with dsRNA for GFP. Mock-treated cells without dsRNA were also processed in parallel. (A) The efficiency of the depletion was analyzed at the RNA level by RT-PCR. The levels of 18S rRNA and actin 5C mRNA were analyzed in parallel to assess the specificity of the dsRNA treatments. (B) The dsRNA treatment was also effective at the protein level as shown by Western blotting. (C) The effect of dBrm depletion on the processing of pre-mRNA from three selected genes – CG9380, CG8421 and CG8092 – was validated by RT-PCR. Semi-quantitative RT-PCR reactions were analyzed in agarose gels stained with ethidium bromide. The results were also validated by qPCR and expressed as relative abundance related to actin 5C. The exon-intron organization of each transcript is shown to the left. Total RNA was purified from mock-treated cells (lane 1) and from cells treated with dsRNA for either GFP (lane 2) or dBrm (lane 3). The RNA preparations were analyzed by RT-PCR using primer-pairs designed to amplify specific transcripts, as indicated in the figure. The positions of the PCR primer pairs are indicated as short bars on top of each transcript. For each gene, dBrm depletion changed the levels of specific transcripts in agreement with the microarray data. In CG9380 and CG8421, the RB and RA transcripts, respectively, were upregulated in dBrm-depleted cells. In CG8092, the RB transcript was downregulated by dBrm depletion.
doi:10.1371/journal.pgen.1000470.g005
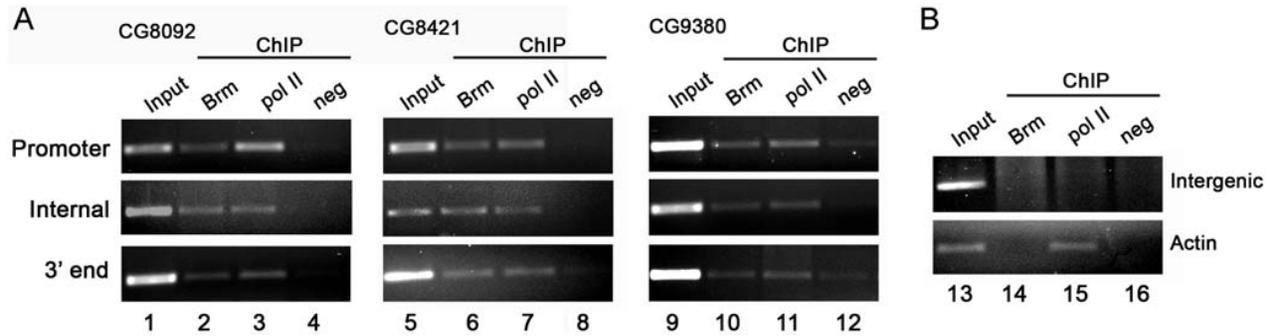
Figure 6. The association of dBrm with its target genes studied by ChIP. (A) Chromatin was extracted from S2 cells after fixation with formaldehyde, and immunoprecipitation reactions were carried out with Ab1 to precipitate chromatin fragments bound to dBrm (lanes 2, 6, 10 and 14). A positive control immunoprecipitation was carried out using an antibody against the C-terminal domain of the largest subunit of RNA pol-II (lanes 3, 7, 11 and 15). A negative control immunoprecipitation was carried out in parallel with an unrelated rabbit antibody against mouse immunoglobulins (lanes 4, 8, 12 and 16). For each gene analyzed, PCR reactions were carried out with primers specific for the proximal promoter, the internal region involved in the alternative processing and the 3'-end of the gene, as indicated in the figure. (B) The actin 5C gene and an intergenic region were analyzed in parallel to assess the specificity of the ChIP (lanes 13–16).
doi:10.1371/journal.pgen.1000470.g006

microarray experiment and asked whether depletion of other SWI/SNF subunits had any effects on the processing of the pre-mRNAs derived from the CG8092, CG8421 and CG9380 genes. Depletion of either Mor or Snr1, two SWI/SNF core subunits, induced changes very similar to those induced by dBrm, whereas depletion of the signature subunits Osa, Bap170 or PB gave milder and in many cases non-significant effects (Figure 7). The effect of Snr1 depletion on the abundances of the CG8421 and CG9380 transcripts was validated by RNAi and RT-qPCR (Figure S7). These results suggest that dBrm does not regulate pre-mRNA processing alone: it is part of the core SWI/SNF complex.

The Ontology of the Genes Regulated Post-Transcriptionally by SWI/SNF

The 15 genes identified above were tested for enrichment of gene ontology (GO) terms for biological processes and molecular functions [31]. The expected number of genes associated with a given term by random was compared with the observed number of SWI/SNF-regulated genes that were associated with that particular term using the Fisher's exact test. Several GO terms for biological processes were very significantly enriched, including positive regulation of developmental process (GO:0051094, p = 0.00003), programmed cell death (GO:0008219, p = 0.0005) and instar larval or pupal morphogenesis (GO:0048707, p = 0.0008).

The test for enrichment of molecular functions also revealed significant associations. Seven of the genes are predicted to code for proteins with catalytic activity. Three of them have phosphatase activity (GO:0016791, p = 0.0009) and three have RNA polymerase II transcription factor activity (GO:0003702, p = 0.002). And nine out of the fifteen gene products were found to have metal ion binding activity (GO:0046872, p = 0.00002). In agreement with this finding, a search for conserved protein domains revealed that five of the genes regulated by SWI/SNF post-transcriptionally, including the known transcription factors broad/CG11491, lola/CG12052, mod(mdg4)/CG32491 and hr39/CG8676, code for proteins that contain zinc finger domains.

In summary, the tests for ontology enrichments indicate that the genes that are regulated post-transcriptionally by SWI/SNF are primarily enzymes and transcription factors that function in the regulation of postembryonic developmental processes. It is worth mentioning that one of them, broad/CG11491, is a key regulator of metamorphosis [32].

Discussion

Brm Binds to the Nascent Pre-mRNP during Transcription

The hBrm and hBrg1 proteins are the catalytic subunits of the SWI/SNF chromatin remodeling complexes and much of what is known about their function comes from studies of transcriptional regulation [33,34]. SWI/SNF participates in regulatory networks that can result in either the activation or the repression of a gene, depending on the genomic context and the activities of additional co-regulators [35]. One of the functions of SWI/SNF is to remodel the structure of nucleosomes at promoter regions in an ATP-dependent manner [36,37]. In some genes, SWI/SNF is associated also with downstream regions of the genes and influences transcription elongation [38]. Recent studies have shown that hBrm and hBrg1 regulate the alternative splicing of several pre-mRNAs in human cells [12,39]. The current proposed model (Figure 8A) suggests that hBrm acts together with mRNA-binding proteins such as Sam68 or p54[nrb] to decrease the elongation rate of RNA pol II and to induce the accumulation of RNA pol II at specific positions in the gene. This facilitates the assembly of the splicing machinery at weak splice sites, which favors the inclusion of proximal exons [12,39]. Inactivation of the ATPase activity of hBrm does not affect its ability to regulate alternative splicing [12], which indicates that the mechanism by which hBrm regulates pre-mRNA processing is independent of its nucleosome remodeling activity.

We have now studied Brm in two insect model systems, *D. melanogaster* and *C. tentans*, and our results show that Brm becomes incorporated into nascent pre-mRNPs during transcription. This conclusion is based on several observations. Firstly, immunofluorescence experiments combined with RNase A digestion show that the association of ctBrm with the polytene chromosomes of *C. tentans* is partially mediated by RNA. Secondly, immuno-EM reveals that a fraction of ctBrm is associated with the BR pre-mRNPs, not with the chromatin. Thirdly, biochemical fractionation experiments show that Brm is present in the chromosomal RNP fraction in *C. tentans*, *D. melanogaster* and *H. sapiens*. The fact that Brm interacts directly with the nascent pre-mRNP suggests that Brm regulates gene expression post-transcriptionally.

An interesting question is whether the post-transcriptional role of Brm is mediated by the Brm protein alone or in complex with other SWI/SNF subunits. Fractionation of nuclear extracts from
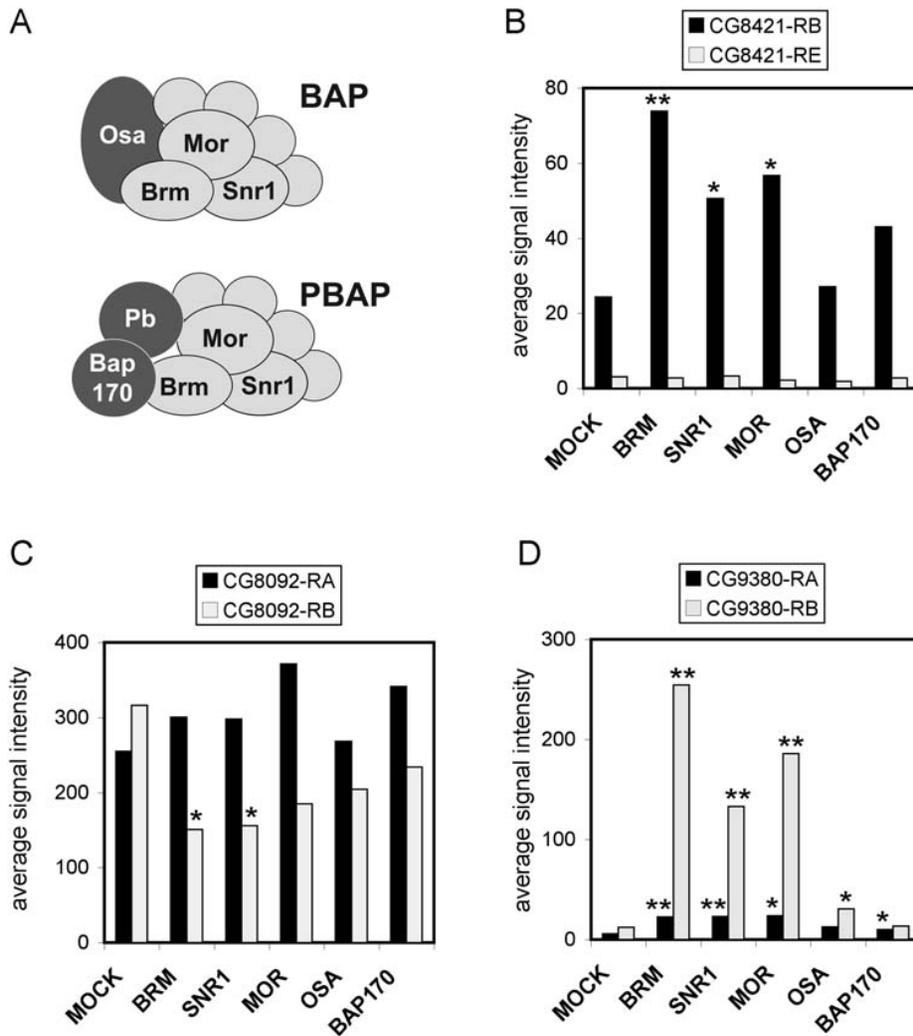
Figure 7. Depletion of other SWI/SNF subunits changes the relative abundances of alternatively processed transcripts in S2 cells.
(A) Schematic representation of the composition of the BAP and PBAP complexes. The light grey subunits constitute the SWI/SNF core whereas the dark grey ones represent signature subunits. (B–D) The histograms show the abundances (expressed as average signal intensities) for alternatively processed transcripts from three selected genes, as indicated, in mock-treated cells and in cells treated with dsRNA for Brm, Snr1, Mor, Osa or Bap170. Data from ArrayExpress E-TABM-169 [28]. * indicates $p < 0.05$. ** indicates $p < 0.005$.
doi:10.1371/journal.pgen.1000470.g007

HeLa cells showed that several SWI/SNF core subunits, Brm, Brg1, Baf155, Baf170 and INI1/SNF5, are associated with nuclear RNPs. This observation suggests that, at least in human cells, the hBrm and hBrg1 proteins that are present in the RNP-associated fraction are part of a SWI/SNF complex. This seems to be the case also in insect cells, as judged by the effects of depletion of individual SWI/SNF subunits on pre-mRNA processing (see below).

SWI/SNF Regulates the Abundances of Alternatively Processed Transcripts

We have analyzed the expression profiles of S2 cells in which either Brm or other subunits of SWI/SNF have been silenced by RNAi, and we have identified 15 genes that show changes in the relative abundance of alternatively processed transcripts. The number of genes that are affected by dBrm depletion is likely to be underestimated because our study is based on the use of gene expression arrays that do not fully cover all the transcriptome. Our

results clearly show that Brm influences the levels of alternatively processed mRNAs in *Drosophila* cells. Interestingly, the analysis of the expression profiles of S2 cells depleted of other SWI/SNF core subunits revealed effects similar to those induced by Brm depletion. This finding suggests that the role of Brm in pre-mRNA processing is mediated by a core SWI/SNF complex.

What is the functional significance of the alternative processing events regulated by SWI/SNF? For some of the identified genes, the alternative transcripts code for different protein isoforms. For instance, the CG8092 gene encodes two different proteins. The longer isoform, CG8092-PA, contains an AT-hook motif and a zinc finger domain whereas the shorter isoform, CG8092-PB, lacks the zinc finger. Considering that CG8092 is an essential gene in *D. melanogaster* (http://flybase.org/reports/FBal0211894.html) and that the CG8092 proteins resemble transcription factors, changes in the relative abundance of the CG8092 isoforms are likely to be biologically significant.

48
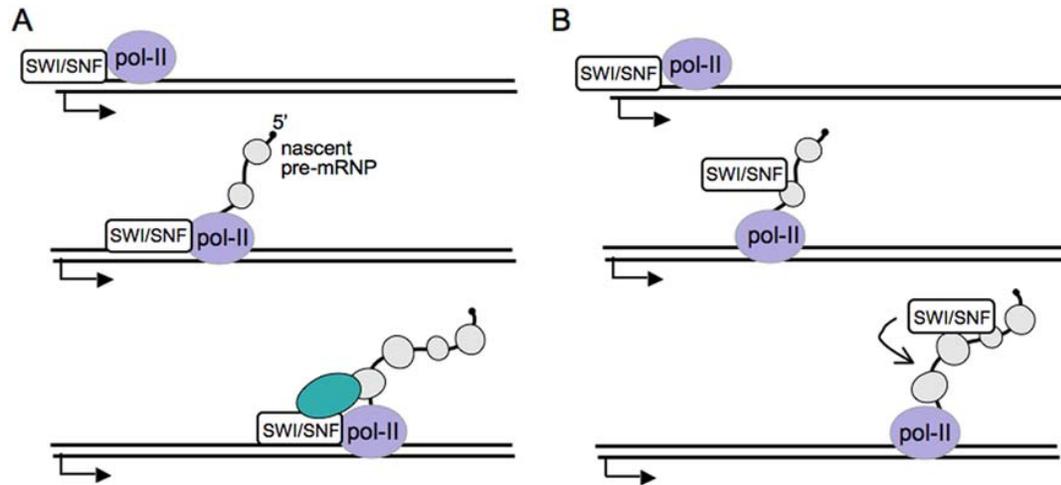
Figure 8. Two models for the post-transcriptional role of SWI/SNF in pre-mRNA processing. (A) SWI/SNF interacts with components of the nascent pre-mRNP complex and with the transcription machinery. These interactions cause a delay in the transcription elongation rate, which in turn affects splice site selection. Based on data from Batsché et al. [12]. (B) SWI/SNF becomes incorporated into the nascent pre-mRNP complex and modulates the interactions between the pre-mRNP and the processing machineries.
doi:10.1371/journal.pgen.1000470.g008

In other cases, the differences among the alternative transcripts regulated by SWI/SNF lie outside the ORF, in the 3′ UTRs of the mRNAs. The proteins encoded are thus identical, but the stability of the transcripts may differ. A search in miRBase (http://microrna.sanger.ac.uk), a database for microRNA (miRNA) data, revealed that several of the genes identified in our study are predicted targets for miRNA regulation and that in most cases the miRNAs are specific for each alternative transcript (data not shown). One example of this is broad/CG11491, a gene with seven alternative mRNAs with five different 3′ UTRs. Interestingly, all five 3′ UTRs contain predicted miRNA targets [40]. These observations lead us to speculate that in some cases the regulation of alternative processing mediated by SWI/SNF acts in concert with the miRNA pathway to fine-tune the abundances of key proteins with catalytic and/or regulatory activities.

Transcription versus Post-Transcriptional Regulation

Depletion of SWI/SNF does not change the levels of all the mRNAs that have originated from a pre-mRNA in the 15 genes identified, but only a few. Indeed, in most cases only one mRNA is affected, which indicates that the step that is affected is not only the synthesis of the pre-mRNA but its processing into alternative mRNAs.

SWI/SNF plays an important role in the transcription of many genes, and it may be that the alterations of pre-mRNA processing that we have observed are a consequence of alterations in the synthesis of specific pre-mRNA processing factors. However, we have shown by ChIP that dBrm is physically associated with both the proximal promoter and downstream sequences of the genes affected. We have also shown by immuno-EM of the BR genes of *C. tentans* that ctBrm is associated with nascent pre-mRNPs. These observations strongly suggest that Brm acts directly at the mRNA level.

Our results indicate that dBrm affects both alternative splicing and alternative polyadenylation sites. Many of the alternative processing events regulated by SWI/SNF involve the use of mutually exclusive splicing and polyadenylation sites. The use of the proximal site is favored by dBrm in some of the genes, as would be expected if dBrm acts by reducing the elongation rate of

the Pol-II at certain positions, as has been proposed for the regulation of the CD44 pre-mRNA in human cells [12]. However, depletion of dBrm has the opposite effect in other cases (such as CG18251, CG9380, CG3665), which is difficult to reconcile with a model of kinetic regulation based on modulation of the Pol-II elongation rate. This observation, and the finding that a significant fraction of Brm is associated with nascent pre-mRNPs, lead us to propose that dBrm regulates pre-mRNA processing in a more direct manner. Several mechanisms can be envisioned by which Brm, being part of the pre-mRNP complex, could influence the usage of alternative splicing or polyadenylation sites (Figure 8B). In one possible scenario, SWI/SNF could work as an RNP remodeling factor to modulate interactions between specific processing factors and their target RNA sequences. However, this possibility is unlikely because experiments in human cells suggest that the role of hBrm in pre-mRNA processing does not require a functional ATPase domain [12]. Alternatively, SWI/SNF could influence the structure of the pre-mRNP. For instance, SWI/SNF could recruit pre-mRNA processing factors to the nascent transcript, or prevent the interactions of processing factors with their target RNA sequences. The interaction of Brm with each pre-mRNP is likely to depend on the sequence of the transcript and/or the specific combination of proteins that forms the pre-mRNP. As a consequence, the action of Brm and the specific outcome of the processing reactions will depend in each case on the specific features of the pre-mRNP. The same type of context-dependent mechanism has been proposed to explain the complex function of SWI/SNF in transcription regulation [35 and references therein]. SWI/SNF acts as a co-activator in the transcription of certain genes, but represses the transcription of certain other genes [for example 41], and these two opposite effects are mediated by interactions with different types of co-regulators. In a similar manner, SWI/SNF can either repress or activate the choice of splice sites and/or polyadenylation sites in a gene-specific manner.

The role of SWI/SNF in pre-mRNA processing affects a specific subset of pre-mRNAs [12 and our present results]. Ontology analysis revealed that many of the transcripts regulated post-transcriptionally by SWI/SNF in *D. melanogaster* code for

proteins that are implicated in postembryonic developmental processes. One of them, broad, is a transcription factor that plays a central role in the cross-talk between ecdysone and juvenile hormone, the two hormones that coordinate insect growth and development [reviewed in 42] Dubrovsky 2005. This is particularly interesting, since hBrm and hBrg1 have also been identified as key regulators of growth control and differentiation in mammals [43,44]. hBrm and hBrg1 are differentially expressed during development, and their expression is altered in cancer cells, which leads to deregulation of genetic programs [reviewed in 45]. In summary, SWI/SNF appears to act both transcriptionally and post-transcriptionally to fine-tune the expression of genes with key regulatory functions in development. In this way, SWI/SNF can regulate gene expression at two levels by determining not only the amount of mRNA synthesized from a given promoter but also the type of alternative transcript produced. Acting at the pre-mRNA processing level, SWI/SNF can rapidly modulate the abundance and activity of the resulting protein products by acting on genes that are already active.

Materials and Methods

Animals and Cell Culture

Chironomus tentans were cultured as described by Meyer et al. [46]. The salivary glands used for study were isolated from 4th instar larvae. *C. tentans* tissue culture cells were grown in ZO medium at 24°C as described by Wyss et al. [47]. *D. melanogaster* S2 cells were cultured at 28°C in Schneider's medium (Invitrogen).

Isolation and Expression of a cDNA That Encodes ctBrm

Degenerate primers for nested PCR were designed based on conserved residues in the C-terminal part of *D. melanogaster* Brm (CG5942) and *Anopheles gambiae* Brm (AGAP010462). The sequences of the primers are provided in the Supplementary Materials and Methods. A PCR product of about 750 bp was amplified from a total cDNA preparation made from *C. tentans* tissue culture cells. The sequence of the PCR product encoded a partial protein corresponding to amino acids 1252–1455 in dBrm. The PCR product was cloned into pET21b (Novagen), expressed in BL21 *E. coli* cells (Novagen) and used to immunize rabbits.

Antibodies

The antibody against the C-terminal part of ctBrm was raised in rabbit following standard procedures (AgriSera, Sweden). The anti-rat Brg1 antibody was raised and characterized by Östlund Farrants et al. [24]. The anti-dBrm antibody was raised and characterized by Zraly et al. [25]. The antibodies against hBrm, Baf47/INI1, Baf180, Baf155, Baf170 and Sap155 were characterized by Sif et al. [48], de la Serna et al. [49], Nicolas et al. [50], Shanahan et al. [51] and Will et al. [52], respectively. The anti-Baf170 was from Santa-Cruz (sc-10757). The mAbs 10:3G1 and 2E4 against Hrp36 and Hrp45, respectively, have previously been characterized [53,54]. The Y12 antibody against the Sm epitope of snRNP core proteins was characterized by Lerner et al. [55]. The rabbit anti-Hrp59 was the Y38 antibody raised by Falk et al. [56]. The anti-Pol-II antibody was purchased from Abcam (Ab5408). The anti-TBP was from Santa-Cruz Biotechnology (sc-204). FITC-conjugated, Texas Red-conjugated and gold-conjugated secondary antibodies were from Jackson ImmunoResearch Laboratories. The secondary antibodies conjugated to alkaline phosphatase and horseradish peroxidase were from DakoCytomation.

Immunofluorescence

Salivary glands were pre-fixed with 2% formaldehyde in TKM buffer (10 mM triethanolamine-HCl, 100 mM KCl and 1 mM MgCl$_2$), permeabilized and disrupted by pipetting in 0.25% NP40 in TKM. Individual chromosomes were isolated and fixed with 4% parafomaldehyde in TKM. The chromosomes were then blocked in 2% bovine serum albumin (BSA) in TKM and incubated with antibodies following standard procedures. The secondary antibodies were conjugated to FITC or Texas Red. The immunostained chromosomes were mounted in Vectashield (Vector Laborarories).

Acquisition and Processing of Confocal Images

Preparations were analyzed and images were taken with a laser scanning microscope (model LSM 510; Carl Zeiss MicroImaging, Inc.) equiped with PlanApochromat objectives 40×/1.0 oil and 63×/1.4 oil, using immersion oil Immersol 518F (Carl Zeiss MicroImaging, Inc.). The optical sections were approximately 1 μm thick. Photoshop software (Adobe) was used for the preparation of composite images and for adjustment of intensity and contrast.

Immuno-Electron microscopy

Salivary glands were prefixed and permeabilized, and the polytene chromosomes were isolated by pipetting in the same way as those intended to be used in immunofluorescence experiments. The isolated chromosomes were fixed with freshly prepared 4% paraformaldehyde in TKM. The chromosome preparations were blocked in 2% BSA in TKM for 30 min, incubated with primary antibody, washed and incubated with an anti-rabbit IgG conjugated to 6-nm gold particles. The control chromosomes were incubated with either secondary antibody only or with a pre-immune serum. The stained chromosomes were fixed with 2% glutaraldehyde and embedded in Agar 100. Thin sections (70 nm) of plastic-embedded chromosomes were stained with uranyl acetate and examined in a FEI 120 kV TECNAI electron microscope. Images were recorded using a Gatan US 1000P CCD camera. For quantitative purposes, the BR genes were photographed at random areas and the numbers of gold markers in the proximal, middle, and distal segments of the BR genes were counted. The specificity of the immuno-EM results was supported by negative control preparations that were processed in parallel and incubated with either the pre-immune serum or with only secondary antibody. The labeling density in the negative controls was calculated and was found to be between 18.5 and 7.5%, respectively.

Preparation of Nuclear Protein Extracts

C. tentans tissue culture cells, *Drosophila* S2 cells or human HeLa cells were homogenized in PBS containing 0.2% NP-40. The homogenate was centrifuged at 1500 *g* for 10 min at 4°C. The pellet containing the nuclei was resuspended in PBS, sonicated and centrifuged at 16,300 *g* for 10 min at 4°C. The resulting supernatant was the soluble nuclear extract. The pellet was resupended in PBS, digested with RNase A (100 μg/ml) and centrifuged at 16,300 *g* for 10 min at 4°C. The supernatant was the chromosomal RNP fraction and contained proteins that were retained in the pellet through RNA-dependent interactions. For the experiment shown in Figure 4G, the RNase A digestion was allowed to run for 15 min at either 4°C or 25°C.

Immunoprecipitation

Immunoprecipitation experiments were carried out following standard procedures. Soluble nuclear extracts and chromosomal RNP extracts were prepared as described above, supplemented with 0.1% NP40 and used as input. The bound proteins were eluted, precipitated with acetone and analyzed by SDS-PAGE and Western blotting.

SDS-PAGE and Western Blotting

Protein extracts were separated by SDS-PAGE and transferred to polyvinylidenefluoride membranes (Millipore) following standard procedures. The NBT/BCIP system was used for detecting secondary antibodies conjugated to alkaline phosphatase. The ECL system (GE Healthcare) was used for the chemiluminiscent detection of horseradish peroxidase.

Chromatin Immunoprecipitation

Chromatin was prepared from S2 cells after cross-linking with 2% formaldehyde. The chromatin was sheared by sonication to a DNA size of 250–1000 bp and pre-cleared. Chromatin fragments were precipitated with antibodies against either rBrg1 (Ab1) or Pol-II (Abcam) using protein A/G-Sepharose beads (50% of each). The precipitated DNA fragments were purified and amplified by PCR using primers for the CG8092, CG8421 and CG9380 genes. Actin 5C (CG4027) was used as a control. The PCR conditions were optimized to avoid saturation. See

Microarray Data Analysis

The microarray data was extracted from Array Express, E-TABM 169 (http://www.ebi.ac.uk/microarray-as/aew/). *Drosophila* Genome 2.0 Arrays (Affymetrix) were hybridized with total RNA purified from *Drosophila* S2 cells treated with dsRNA corresponding to dBrm or to other subunits of SWI/SNF [28].

RNA Interference in S2 Cells

dsRNAs against dBrm and GFP were prepared by *in vitro* transcription from PCR products with T7 promoters on both ends of the amplimers, using the Megascript RNAi kit (Ambion). The sequences of the PCR primers are provided in the Supplementary Materials and Methods. The RNAi treatment was performed as described by Clemens et al. [57]. In brief, 20 μg of dsRNA was applied to S2 cells and the cells were harvested after 48 h. Total RNA from S2 cells was extracted, reverse transcribed and used as a template for PCR reactions using primers specific for selected transcripts. Quantitative real-time PCR was carried out in an ABI7000 system using SYBR Green (Applied Biosystems).The RNAi experiments were repeated three times to confirm the reproducibility of the observations.

Supporting Information

A detailed description of the Materials and Methods, including primer sequences, are provided as Supporting Information (Text S1). Seven Supporting Figures are also provided (Figures S1, S2, S3, S4, S5,S6, and S7).

Supporting Information

Figure S1 The amino acid sequence of ctBrm. Multiple sequence alignment of the carboxy terminal portion of the Brm proteins of *C. tentans* (ctBrm, FM211186), *D. melanogaster* (dBrm, CG5942), *Anopheles gambiae* (agBrm, XM_311484) and *Aedes aegypti* (aaBrm, XM_001650039). The multiple sequence alignment was done with using CLUSTAL W at the Biology WorkBench 3.2

(http://workbench.sdsc.edu/). The ctBrm sequence was deduced from a partial cDNA obtained by nested PCR using degenerate primers based on the amino acid sequences of dBrm and agBrm as described in the Supporting Materials and Methods. The amino acid sequence of the carboxy terminal portion of ctBrm shares 60.4% identity with dBrm and 69% with aaBrm and agBrm.
Found at: doi:10.1371/journal.pgen.1000470.s001 (1.14 MB JPG)

Figure S2 Specificity of three anti-Brm/Brg1 antibodies in *C. tentans*. (A) Three independent antibodies were tested by Western blot against nuclear protein extracts prepared from *C. tentans* cultured cells. Ab1 (*lane 2*) was raised against the rat Brg1 protein. Ab2 (*lane 3*) was raised against the C-terminal part of the ct-BRM protein (Figure S1). Ab3 (*lane 4*) was raised against dBrm. A negative control without primary antibody was processed in parallel (*lane 1*). The three antibodies detected a major band of approximate molecular mass 200 kDa (*arrow*). The mobility of molecular mass standards is shown to the left in kDa. (B) A preparation of total proteins from larval salivary glands was probed with Ab1. The antibody recognized a band with the expected mobility of ctBrm (*arrow*). (C) A nuclear protein extract was prepared from *C. tentans* cultured cells and ctBrm was immunoprecipitated using the Ab1 antibody. A negative control immunoprecipitation without primary antibody was processed in parallel to assess the specificity of the experiment. The immunoprecipitated protein was probed by Western blot using Ab2 and Ab3, as indicated in the figure. Ab2 and Ab3 detect the 200 kDa protein immunoprecipitated by Ab1.
Found at: doi:10.1371/journal.pgen.1000470.s002 (0.14 MB JPG)

Figure S3 The association of ctBrm with the polytene chromosomes studied by immunofluorescence. (A–B) The images show confocal sections of isolated polytene chromosomes stained with either Ab2 or Ab3, as indicated. Multiple loci were intensely stained in the chromosomes, including the BR puffs BR1, BR2 and BR3 in chromosome IV. (C–D) Confocal sections of isolated polytene chromosomes stained with either Ab2 or Ab3 and co-stained with a mAb against Hrp45. Hrp45 is an hnRNP protein used as a marker to visualize the BRs. The merged images show that the BR puffs are stained by the anti-Brm antibodies. The scale bars represent approximately 10 μm.
Found at: doi:10.1371/journal.pgen.1000470.s003 (0.33 MB JPG)

Figure S4 Patterns of alternative pre-mRNA processing affected by dBrm depletion. The figure summarizes the different types of alternative pre-mRNA processing reactions changed in S2 cells treated with Brm-dsRNA: (A) intron retention and alternative use of polyadenylation signals, (B) alternative 3′ slice sites and alternative polyadenylation, (C) alternative use of polyadenylation signals, (D) alternative 5′ slice sites and alternative polyadenylation, and (E) exon skipping. Note that the alternative splicing and polyadenlation events are exclusive in most cases and the choice of a given splice site determines the site of a given polyadenylation and vice versa. Examples of representative genes are given for each case to the right. The figure is based on the annotations available at FlyBase (http://flybase.bio.indiana.edu/).
Found at: doi:10.1371/journal.pgen.1000470.s004 (0.18 MB JPG)

Figure S5 The abundance of alternative CG8421 transcripts is affected by Brm depletion. (A) Schematic representation of the alternative transcripts derived from the CG8421 gene. F and R indicate the positions of the forward and reverse PCR primers, respectively. (B) dBrm was knocked-down in S2 cells by RNAi. Control cells were treated in parallel with dsRNA for GFP, as in Figure 5. RT-PCR reactions were carried out using primers F and R to amplify simultaneously the alternatively spliced CG8421

mRNAs. The PCR products were analyzed in agarose gels stained with ethidium bromide. The mobility of molecular mass standards is shown to the left, in nt.

Found at: doi:10.1371/journal.pgen.1000470.s005 (0.25 MB JPG)

Figure S6 Optimization of PCR reactions for the proximal, middle and 3′ end regions of the CG9380, CG8092 and CG8421 genes. DNA purified from ChIP experiments was analyzed by PCR using primer pairs specific for each region of interest. For each primer pair, different amounts of DNA template were tested as indicated in the figure, and the conditions of the PCR reactions were optimized in order to determine the linear range of the PCR amplification and to avoid saturation. For each ChIP experiment, the optimal conditions were established and all the samples, including the negative control immunoprecipitation, were run under the same conditions.

Found at: doi:10.1371/journal.pgen.1000470.s006 (0.45 MB JPG)

Figure S7 The effect of SNR1 depletion on the abundance of alternative transcripts from the CG8421 and CG9380 genes. The expression of SNR1 in S2 cells was silenced by RNAi. Control cells were treated in parallel with GFP-dsRNA. Total RNA was purified from cells treated with either SNR1-dsRNA or GFP-dsRNA, reverse transcribed and analyzed by qPCR. The relative abundance of each mRNA was expressed relative to the actin 5C mRNA levels. (A) The expression of SNR1 was significantly reduced in cells treated by SNR1-dsRNA. (B) Depletion of SNR1 affected the levels of the CG8421 and CG9380 mRNAs. The effects were transcript-specific and were very similar to those observed in BRM-depleted cells (compare with Figure 5C).

Found at: doi:10.1371/journal.pgen.1000470.s007 (0.20 MB JPG)

Text S1 Supporting materials and methods.

Found at: doi:10.1371/journal.pgen.1000470.s008 (0.07 MB DOC)

Author Contributions

Conceived and designed the experiments: AT JR DB AKÖF NV. Performed the experiments: AT JR DB AKÖF NV. Analyzed the data: AT JR DB AKÖF NV. Contributed reagents/materials/analysis tools: AT DB AKÖF NV. Wrote the paper: AT AKÖF NV.

References

1. Sorek R, Shamir R, Ast G (2004) How prevalent is functional alternative splicing in the human genome? Trends Genet 20: 68–71.
2. Stolc V, Gauhar Z, Mason C, Halasz G, van Batenburg MF, et al. (2004) A gene expression map for the euchromatic genome of Drosophila melanogaster. Science 306: 655–660.
3. Allemand E, Batsché E, Muchardt C (2008) Splicing, transcription, and chromatin: a ménage à trios. Curr Opin Genet Dev 18: 145–151.
4. Cramer P, Cáceres JF, Cazalla D, Kadener S, Muro AF, et al. (1999) Coupling of transcription with alternative splicing: RNA pol II promoters modulate SF2/ASF and 9G8 effects on an exonic splicing enhancer. Mol Cell 4: 251–258.
5. Auboeuf D, Batsché E, Dutertre M, Muchard C, O'Malley BW (2007) Coregulators: transducing signal from transcription to alternative splicing. Trends Endocrinol Metab 18: 122–129.
6. Iwasaki T, Chin WW, Ko L (2001) Identification and characterization of RRM-containing coactivator activator (CoAA) as TRBP-interacting protein, and its splice variant as a coactivator modulator (CoAM). J Biol Chem 276: 33375–33383.
7. Auboeuf D, Dowhan DH, Li X, Larkin K, Ko L, et al. (2004) CoAA, a nuclear receptor coactivator protein at the interface of transcriptional coactivation and RNA splicing. Mol Cell Biol 24: 442–453.
8. Dowhan DH, Hong EP, Auboeuf D, Dennis AP, Wilson MM, et al. (2005) Steroid hormone receptor coactivation and alternative RNA splicing by U2AF65-related proteins CAPERalpha and CAPERbeta. Mol Cell 17: 429–439.
9. Meng Q, Rayala SK, Gururaj AE, Talukder AH, O'Malley BW, et al. (2007) Signaling-dependent and coordinated regulation of transcription, splicing, and translation resides in a single coregulator, PCBP1. Proc Natl Acad Sci U S A 104: 5866–5871.
10. de la Mata M, Kornblihtt AR (2006) RNA polymerase II C-terminal domain mediates regulation of alternative splicing by SRp20. Nat Struct Mol Biol 13: 973–980.
11. Kornblihtt AR (2005) Promoter usage and alternative splicing. Curr Opin Cell Biol 17: 262–268.
12. Batsché E, Yaniv M, Muchardt C (2006) The human SWI/SNF subunit Brm is a regulator of alternative splicing. Nat Struct Mol Biol 13: 22–29.
13. Mohrmann L, Verrijzer CP (2005) Composition and functional specificity of SWI2/SNF2 class chromatin remodeling complexes. Biochim Biophys Acta 1681: 59–73.
14. Saha A, Wittmeyer J, Cairns BR (2006) Chromatin remodelling: the industrial revolution of DNA around histones. Nat Rev Mol Cell Biol 7: 437–447.
15. Armstrong JA, Papoulas O, Daubresse G, Sperling AS, Lis JT, et al. (2002) The Drosophila BRM complex facilitates global transcription by RNA polymerase II. EMBO J 21: 5245–5254.
16. Zraly CB, Middleton FA, Dingwall AK (2006) Hormone-response genes are direct in vivo regulatory targets of Brahma (SWI/SNF) complex function. J Biol Chem 281: 35305–35315.
17. Chalkley GE, Moshkin YM, Langenberg K, Bezstarosti K, Blastyak A, et al. (2008) The transcriptional coactivator SAYP is a trithorax group signature subunit of the PBAP chromatin remodeling complex. Mol Cell Biol 28: 2920–2929.
18. Wieslander L (1994) The Balbiani ring multigene family: coding repetitive sequences and evolution of a tissue-specific cell function. Prog Nucleic Acid Res Mol Biol 48: 275–313.
19. Visa N, Izaurralde E, Ferreira J, Daneholt B, Mattaj IW (1996) A nuclear cap-binding complex binds Balbiani ring pre-mRNA cotranscriptionally and accompanies the ribonucleoprotein particle during nuclear export. J Cell Biol 133: 15–14.
20. Baurén G, Wieslander L (1994) Splicing of Balbiani ring 1 gene pre-mRNA occurs simultaneously with transcription. Cell 76: 183–192.
21. Daneholt B (2001) Assembly and transport of a premessenger RNP particle. Proc Natl Acad Sci U S A 98: 7012–7017.
22. Skoglund U, Andersson K, Björkroth B, Lamb MM, Daneholt B (1983) Visualization of the formation and transport of a specific hnRNP particle. Cell 34: 847–855.
23. Botelho SC, Tyagi A, Hessle V, Ostlund Farrants AK, Visa N (2008) The association of Brahma with the Balbiani ring 1 gene of Chironomus tentans studied by immunoelectron microscopy and chromatin immunoprecipitation. Insect Mol Biol. In press.
24. Östlund Farrants AK, Blomquist P, Kwon H, Wrange Ö (1997) Glucocorticoid receptor-glucocorticoid response element binding stimulates nucleosome disruption by the SWI/SNF complex. Mol Cell Biol 17: 895–905.
25. Zraly CB, Marenda DR, Nanchal R, Cavalli G, Muchardt C, et al. (2003) SNR1 is an essential subunit in a subset of Drosophila brm complexes, targeting specific functions during development. Dev Biol 253: 291–308.
26. Wetterberg I, Zhao J, Masich S, Wieslander L, Skoglund U (2001) In situ transcription and splicing in the Balbiani ring 3 gene. EMBO J 20: 2564–2574.
27. Dellaire G, Makarov EM, Cowger JJ, Longman D, Sutherland HG, et al. (2002) Mammalian PRP4 kinase copurifies and interacts with components of both the U5 snRNP and the N-CoR deacetylase complexes. Mol Cell Biol 22: 5141–5156.
28. Moshkin YM, Mohrmann L, van Ijcken WF, Verrijzer CP (2007) Functional differentiation of SWI/SNF remodelers in transcription and cell cycle control. Mol Cell Biol 27: 651–661.
29. Neugebauer KM (2002) On the importance of being co-transcriptional. J Cell Sci 115: 3865–3871.
30. Görnemann J, Kotovic KM, Hujer K, Neugebauer K (2005) Cotranscriptional spliceosome assembly occurs in a stepwise fashion and requires the cap binding complex. Mol Cell 19: 53–63.
31. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, et al. (2000) Gene ontology: tool for the unification of biology. The Gene Ontology Consortium Nat Genet 25: 25–29.
32. von Kalm L, Crossgrove K, Von Seggern D, Guild GM, Beckendorf SK (1994) The Broad-Complex directly controls a tissue-specific response to the steroid hormone ecdysone at the onset of Drosophila metamorphosis. EMBO J 13: 3505–3516.

33. Katsani KR, Mahmoudi T, Verrijzer CP (2003) Selective gene regulation by SWI/SNF-related chromatin remodeling factors. Curr Top Microbiol Immunol 274: 113–141.

34. Bouazoune K, Brehm A (2006) ATP-dependent chromatin remodeling complexes in Drosophila. Chromosome Res 14: 433–449.

35. Zhang B, Chambers KJ, Faller DV, Wang S (2007) Reprogramming of the SWI/SNF complex for co-activation or co-repression in prohibitin-mediated estrogen receptor regulation. Oncogene 26: 7153–7157.

36. Soutoglou E, Talianidis I (2002) Coordination of PIC assembly and chromatin remodeling during differentiation-induced gene activation. Science 295: 1901–1904.

37. Racki LR, Narlikar GJ (2008) ATP-dependent chromatin remodeling enzymes: two heads are not better, just different. Curr Opin Genet Dev 18: 137–144.

38. Corey LL, Weirich CS, Benjamin IJ, Kingston RE (2003) Localized recruitment of a chromatin-remodeling activity by an activator in vivo drives transcriptional elongation. Genes Dev 17: 1392–1401.

39. Ito T, Watanabe H, Yamamichi N, Kondo S, Tando T, et al. (2008) Brm transactivates the telomerase reverse transcriptase (TERT) gene and modulates the splicing patterns of its transcripts in concert with p54(nrb). Biochem J 411: 201–209.

40. Enright AJ, John B, Gaul U, Tuschl T, Sander C, et al. (2003) MicroRNA targets in Drosophila. Genome Biology 5. Available: http://genomebiology.com/2003/5/1/R1.

41. Wang F, Zhang B, Faller DV (2004) BRG1/BRM and prohibitin are required for growth suppression by estrogen antagonists. EMBO J 23: 2293–2303.

42. Dubrovsky EB (2005) Hormonal cross talk in insect development. Trends Endocrinol Metab 16: 6–11.

43. Inayoshi Y, Miyake K, Machida Y, Kaneoka H, Terajima M, Dohda T, Takahashi M, Iijima S (2006) Mammalian chromatin remodeling complex SWI/SNF Is essential for enhanced expression of the albumin gene during liver development. J Biochem 139: 177–188.

44. Yan Z, Wang Z, Sharova L, Sharov AA, Ling C, Piao Y, Aiba K, Matoba R, Wang W, Ko MSH (2008) BAF250B-associated SWI/SNF chromatin-remodeling complex is required to maintain undifferentiated mouse embryonic stem cells. Stem Cells 26: 1155–1165.

45. Reisman D, Glaros S, Thompson EA (2009) The SWI/SNF complex and cancer. Oncogene, in press.

46. Meyer B, Mähr R, Eppenberger HM, Lezzi M (1983) The activity of Balbiani rings 1 and 2 in salivary glands of Chironomus tentans larvae under different modes of development and after pilocarpine treatment. Dev Biol 98: 265–277.

47. Wyss C (1982) Chironomus tentans epithelial cell lines sensitive to ecdysteroids, juvenile hormone, insulin and heat shock. Exp Cell Res 139: 309–319.

48. Sif S, Stukenberg PT, Kirschner MW, Kingston RE (1998) Mitotic inactivation of a human SWI/SNF chromatin remodeling complex. Genes Dev 12: 2842–2851.

49. de La Serna IL, Carlson KA, Hill DA, Guidi CJ, Stephenson RO, et al. (2000) Mammalian SWI-SNF complexes contribute to activation of the hsp70 gene. Mol Cell Biol 20: 2839–2851.

50. Nicolas RH, Goodwin GH (1996) Molecular cloning of polybromo, a nuclear protein containing multiple domains including five bromodomains, a truncated HMG-box, and two repeats of a novel domain. Gene 175: 233–240.

51. Shanahan F, Seghezzi W, Parry D, Mahony D, Lees E (1999) Cyclin E associates with BAF155 and BRG1, components of the mammalian SWI-SNF complex, and alters the ability of BRG1 to induce growth arrest. Mol Cell Biol 19: 1460–1469.

52. Will CL, Schneider C, MacMillan AM, Katopodis NF, Neubauer G, et al. (2001) A novel U2 and U11/U12 snRNP protein that associates with the pre-mRNA branch site. EMBO J 20: 4536–4546.

53. Kiseleva E, Visa N, Wurtz T, Daneholt B (1997) Immunocytochemical evidence for a stepwise assembly of Balbiani ring premessenger ribonucleoprotein particles. Eur J Cell Biol 74: 407–416.

54. Wurtz T, Kiseleva E, Nacheva G, Alzhanova-Ericcson A, Rosén A, et al. (1996) Identification of two RNA-binding proteins in Balbiani ring premessenger ribonucleoprotein granules and presence of these proteins in specific subsets of heterogeneous nuclear ribonucleoprotein particles. Mol Cell Biol 16: 1425–1435.

55. Lerner EA, Lerner MR, Janeway CA Jr, Steitz JA (1981) Monoclonal antibodies to nucleic acid-containing cellular constituents: probes for molecular biology and autoimmune disease. Proc Natl Acad Sci U S A 78: 2737–2741.

56. Falk R, Agaton C, Kiesler E, Jin S, Wieslander L, et al. (2003) An improved dual-expression concept, generating high-quality antibodies for proteomics research. Biotechnol Appl Biochem 38: 231–239.

57. Clemens JC, Worby CA, Simonson-Leff N, Muda M, Maehama T, et al. (2000) Use of double-stranded RNA interference in Drosophila cell lines to dissect signal transduction pathways. Proc Natl Acad Sci U S A 97: 6499–6503.

53

Supporting Information

Supporting Figures

```
aaBrm     MFDQKSTGSERQQFLQSILHQDEMDEEEENEVPDDEMINLMISRNDDELE
agBrm     MFDQKSTGSERQQFLQSILHQDEMDEEEENEVPDDEMINLMISRTDDELE
ctBrm     MFDQKSTGSERQQFLQSILTQDEGDDEEENEVPDDEMINFMIARSDEELE
dBrm      MFDQKSTGSERQQFLQTILHQDDNEEEEENEVPDDEMINMMIARSEEEIE
          ***************:** **: ::***************:**:*.::*:*

aaBrm     LFKKMDAERKAEE-----VKPRLIDESELPDWLVKEEEEVDRWDYEEDNS
agBrm     LFKKMDAERKAEE-----VKPRLLDEAELPDWLVKDDEEVDRWDYEEETS
ctBrm     LFKQMDSDRKKSE-----TKKRLVEECELPDWLTNNDDD-DRWNDDEDDS
dBrm      IFKRMDAERKKEDEEIHPGRERLIDESELPDWLTKDDDEVERFHYQYDED
          :**:**::** .:      : **::*.******.:::::: :*:. : : .

aaBrm     -ILGRGSRQRKEVDYTDSLTEKEWLKAID-DGADFDEELEEEEREKKRKG
agBrm     -ILGRGSRQRKEVDYTDSLTEKEWLKAID-DGADFDEELEEEEREKKRKG
ctBrm     TILGRGSRQRKEVDYTDSLTEKEWLKAIDGDDIDFDEELENEERERKRKS
dBrm      TILGRGSRQRKEVDYTDSLTEKEWLKAID-DGAEFDEEEEEDDSKRKRRK
           ************************** *. :**** *::: ::**:

aaBrm     RKRKSR-REDSDDESVISTVTKRRKGTTAPKVKKQMKKVIDAVIKYTDQD
agBrm     RKRKGRGRDDSDDESIVSTMTKRKKGITDPKIRKQMKKVLKAVIKYTDAD
ctBrm     KKR----RKDSDDESVTSNITKRKKGTTDPKMKKQMKKIMSIVMKY-EIE
dBrm      RKNR---KEESDDDSLI--LKRRRQNLDKRSKKQMHKIMSAVIKHNQ-D
          :*.     ::.:***:*:  ::.:*::      : :***:*::. *:*: : :

aaBrm     GRVLSEPFMRLPSRRDLPDYYEVIKRPIDIKKIVTRIEEAKYADFPDLEK
agBrm     GRVLSKPFLKLPSRRDLPDYYEVIKKPIDIKKILARIEEAKYVDFADLEN
ctBrm     GRILSEPFMKLPSRRELPDYYEVIKKPLDIKKILIRIDEAKYIDFADLEK
dBrm      GRTLSEPFMKLPSRQRLPDYYEIIKRPVDIKKILQRIEDCKYADLNELEK
          ** **:**::****: ******:**:*:***** **::.** *: :**:

aaBrm     DFFQLCQNAQIYNEEASLIYEDSLELQQIFTTSKQK-
agBrm     DFFLLCQNAQTYNEEASLIYEDSLELQSVFSNAKAKI
ctBrm     DFFQLCQNAQIYNEAR----ANSSTLAAVTSDPSSY-
dBrm      DFMQLCQNAQIYNEEASLIYLDSIALQKVFVGARQRI
          **: ****** ***      :*  *  :    .
```

```
*  -  single, fully conserved residue
:  -  conservation of strong groups
.  -  conservation of weak groups
   -  no consensus
```

Figure S1. The amino acid sequence of ctBrm

Multiple sequence alignment of the carboxy terminal portion of the Brm proteins of *C. tentans* (ctBrm, FM211186), *D. melanogaster* (dBrm, CG5942), *Anopheles gambiae* (agBrm, XM_311484) and *Aedes aegypti* (aaBrm, XM_001650039). The multiple sequence alignment was done with using CLUSTAL W at the Biology WorkBench 3.2 (http://workbench.sdsc.edu/). The ctBrm sequence was deduced from a partial cDNA obtained by nested PCR using degenerate primers based on the amino acid sequences of dBrm and agBrm as described in the Supporting Materials and Methods. The amino acid sequence of the carboxy terminal portion of ctBrm shares 60.4 % identity with dBrm and 69% with aaBrm and agBrm.

1

Figure S2. Specificity of three anti-Brm/Brg1 antibodies in *C. tentans*

(A) Three independent antibodies were tested by Western blot against nuclear protein extracts prepared from *C. tentans* cultured cells. Ab1 (*lane 2*) was raised against the rat Brg1 protein. Ab2 (*lane 3*) was raised against the C-terminal part of the ct-BRM protein (Figure S1). Ab3 (*lane 4*) was raised against dBrm. A negative control without primary antibody was processed in parallel (*lane 1*). The three antibodies detected a major band of approximate molecular mass 200 kDa (*arrow*). The mobility of molecular mass standards is shown to the left in kDa. **(B)** A preparation of total proteins from larval salivary glands was probed with Ab1. The antibody recognized a band with the expected mobility of ctBrm (*arrow*). **(C)** A nuclear protein extract was prepared from *C. tentans* cultured cells and ctBrm was immunoprecipitated using the Ab1 antibody. A negative control immunoprecipitation without primary antibody was processed in parallel to assess the specificity of the experiment. The immunoprecipitated protein was probed by Western blot using Ab2 and Ab3, as indicated in the figure. Ab2 and Ab3 detect the 200 kDa protein immunoprecipitated by Ab1.

Figure S3. The association of ctBrm with the polytene chromosomes studied by immunofluorescence

(A-B) The images show confocal sections of isolated polytene chromosomes stained with either Ab2 or Ab3, as indicated. Multiple loci were intensely stained in the chromosomes, including the BR puffs BR1, BR2 and BR3 in chromosome IV. **(C-D)** Confocal sections of isolated polytene chromosomes stained with either Ab2 or Ab3 and co-stained with a mAb against Hrp45. Hrp45 is an hnRNP protein used as a marker to visualize the BRs. The merged images show that the BR puffs are stained by the anti-Brm antibodies. The scale bars represent approximately 10 μm.

2

Figure S4. Patterns of alternative pre-mRNA processing affected by dBrm depletion

The figure summarizes the different types of alternative pre-mRNA processing reactions changed in S2 cells treated with Brm-dsRNA: **(A)** intron retention and alternative use of polyadenylation signals, **(B)** alternative 3' slice sites and alternative polyadenylation, **(C)** alternative use of polyadenylation signals, **(D)** alternative 5' slice sites and alternative polyadenylation, and **(E)** exon skipping. Note that the alternative splicing and polyadenlation events are exclusive in most cases and the choice of a given splice site determines the site of a given polyadenylation and *vice versa*. Examples of representative genes are given for each case to the right. The figure is based on the annotations available at FlyBase (http://flybase.bio.indiana.edu/).

Supporting Materials and Methods

Animals and cell culture

Chironomus tentans were cultured as described by Meyer et al. [1]. The salivary glands used for study were isolated from 4th instar larvae. *C. tentans* tissue culture cells were grown in ZO medium at 24 °C as described by Wyss et al. [2]. *D. melanogaster* S2 cells were cultured at 28 °C in Schneider's *Drosophila* medium (Invitrogen) containing 10% heat-inactivated fetal calf serum, 50 µg/ml penicillin and 50 µg/ml streptomycin.

Isolation and expression of a cDNA that encodes ctBrm

Degenerate primers for nested PCR were designed to match the conserved sequence of the C-terminal part of *D. melanogaster* Brm (CG5942) and *Anopheles gambiae* Brm (AGAP010462). The sequences of the forward primers F1 and F2 were 5´-TAYAARYTNAAYATGGAYGARAA-3´ and 5´-ATHCARGCNGGNATGTTYGAYCARAA-3´, respectively. The reverse primers R1 and R2 were 5´-GCYTCRTTRTANRKYTGNGCRTT-3´and 5´-TTNAYYTCRTARTARTCNGG-3´. The primers were used to amplify Brm-homologous sequences from a total cDNA preparation made from *C. tentans* tissue culture cells using *Taq* Polymerase (Fermentas). A PCR product of about 750 bp was obtained. The PCR product was purified using NucleoSpin extract II columns (Machery-Nagel), ligated into a TOPO vector (Invitrogen) and transformed into Top10 *E. coli* cells. The recombinant plasmid Topo-ctBrm was purified and sequenced at Eurofins MWG Operon (Ebersberg, Germany). The sequence was analyzed using software available at the Biology Workbench (http://workbench.sdsc.edu/). The cloned cDNA encoded a partial protein corresponding to amino acids 1252-1455 in dBrm. The insert of the Topo-ctBrm plasmid was further amplified with forward 5´-CCGAATTCATCAACTGGTAGTGAACGTCAA-3´and reverse 5´-GTCCCCAAGCTTTAATTCCCGACGTGATGGCAA-3´ primers containing restriction sites for HindIII and EcoRI, respectively. The PCR product was cleaved, purified and ligated into pET21b (Novagen). The resulting plasmid, pET21-ctBrm, was transformed into BL21 *E. coli* cells (Novagen) for the expression of recombinant protein. The expression was induced with 1 mM IPTG for 2 h at 37 °C. The expressed protein was separated by SDS-PAGE in 15% polyacrylamide gels and stained with Coomassie Brilliant Blue. The band corresponding to recombinant ct-Brm was excised from preparative gels, destained, and used for antibody production in rabbits.

Antibodies

The anti-rat Brg1 antibody was raised and characterized by Östlund Farrants et al. [3]. The antibody against the C-terminal part of *ct*-BRM was raised in rabbit at AgriSera (Vännäs, Sweden) according to standard procedures. The anti-dBrm antibody was raised and characterized by Zraly et al. [4]. The monoclonal antibodies 10:3G1 and 2E4 against Hrp36 and Hrp45, respectively, have been previously characterized [5, 6]. The Y12 antibody against the Sm epitope of snRNP core proteins has been characterized by Lerner et al. [7]. The rabbit anti-Hrp59 was the Y38 antibody raised and characterized by Falk et al. [8]. The anti-Pol-II antibody was purchased from Abcam (Ab5408). FITC-conjugated, Texas Red-conjugated and gold-conjugated secondary antibodies were from Jackson ImmunoResearch Laboratories. The secondary antibodies conjugated to alkaline phosphatase and horseradish peroxidase were from DakoCytomation.

Immunofluorescence

Salivary glands were isolated from 4th instar larve and pre-fixed with 2% formaldehyde in TKM buffer (10 mM triethanolamine-HCl, 100 mM KCl and 1 mM $MgCl_2$) for 5 min. The glands were subsequently permeabilized with 2% Nonidet P40 (NP40) in TKM and disrupted by pipetting in 0.25% NP40 in TKM. The chromosomes IV were isolated and transferred to 8-well slides. The chromosomes were washed briefly in TKM and fixed with 4%

4

parafomaldehyde in TKM. The chromosomes were then blocked in 2% bovine serum albumin (BSA) in TKM for 30 min and incubated overnight in a solution of primary antibody diluted in 0.5% BSA in TKM. Affinity purified antibodies were used at 2.5 µg/ml. Sera were diluted 1/2000. Control preparations were incubated with either 0.5% BSA in TKM or pre-immune serum diluted 1/2000 in 0.5% BSA in TKM. The chromosomes were washed with TKM containing 0.01% Tween-20 for 25 min with several changes of buffer and incubated with FITC-conjugated or Texas Red-conjugated secondary antibodies for 60 min. The immunostained chromosomes were washed as above and mounted in Vectashield (Vector Laborarories).

Acquisition and processing of confocal images

Preparations were analyzed and images were taken with a laser scanning microscope (model LSM 510; Carl Zeiss MicroImaging, Inc.) equiped with PlanApochromat objectives 40x/1.0 oil and 63x/1.4 oil, using immersion oil Immersol 518F (Carl Zeiss MicroImaging, Inc.). The optical sections were approximately 1 µm thick. Photoshop software (Adobe) was used for the preparation of composite images and for adjustment of intensity and contrast.

Immuno-electron microscopy

Salivary glands were prefixed and permeabilized, and the polytene chromosomes were isolated by pipetting in the same way as those intended to be used in immunofluorescence experiments. The isolated chromosomes were transferred to 8-well slides and fixed with freshly prepared 4% paraformaldehyde in TKM. After fixation, the chromosomes were washed in TKM three times for 5 min each, and blocked in 2% BSA in TKM for 30 min. The blocked chromosomes were incubated with primary antibody for 1 h at room temperature or overnight at 4 °C, washed and incubated with an anti-rabbit IgG conjugated to 6-nm gold particles. The control chromosomes were incubated with either secondary antibody only or with the pre-immune serum. The stained chromosomes were fixed again with 2% glutaraldehyde in TKM for 1 h, dehydrated in ethanol and embedded in Agar 100. Thin sections (70 nm) of plastic-embedded chromosomes were mounted on nickel grids, stained with 2% uranyl acetate in 50% ethanol, washed with 50% ethanol and air-dried. The preparations were examined and photographed in an FEI 120 kV TECNAI electron microscope using a Gatan US 1000P CCD camera. For quantitative purposes, the BR genes were photographed at random areas. These areas contained portions from multiple BR transcription units. The number of gold markers in the proximal, middle, and distal segments of the BR genes was counted, and the percentage in each segment was calculated. Results from at least two independent chromosomes were pooled. The average labeling in the negative controls labeled with a pre-immune serum was less than 20%. Photoshop software (Adobe) was used for the preparation of composite images and for adjustment of intensity and contrast.

Preparation of nuclear protein extracts

C. tentans tissue culture cells, *Drosophila* S2 cells or human HeLa cells were homogenized in PBS (137 mM NaCl, 3 mM KCl, 8 mM Na_2HPO_4, 2 mM NaH_2PO_4 at pH 7.2) containing 0.2% NP-40 and protease inhibitor tablets (Roche Diagnostics) using a glass homogenizer. The homogenate was centrifuged at 1500 *g* for 10 min at 4 °C. The pellet containing the nuclei was resuspended in PBS containing protease inhibitor, sonicated three times for 4-5 sec each time, and centrifuged at 16,300 *g* for 10 min at 4 °C. The resulting supernatant was the soluble nuclear extract. The pellet was resupended in PBS containing protease inhibitors, digested with RNase A (100 µg/ml) at room temperature for 15 min and centrifuged at 16,300 *g* for 10 min at 4 °C. The supernatant was the chromosomal RNP fraction and contained proteins that were retained in the pellet through RNA-dependent interactions. For the experiment shown in Figure 4G, the RNase A digestion was allowed to run for 15 min at either 4 °C or room temperature.

5

Immunoprecipitation

Immunoprecipitation experiments were carried out following standard procedures. Soluble nuclear extracts and chromosomal RNP extracts were prepared as described above, supplemented with 0.1% NP40 and used as input. Primary antibodies, either mAb Y12 or anti-rBrg1, were added to the extracts (final concentration approximately 2 μg/ml) and the samples were incubated for 90 min at 4 °C with gentle agitation. 30 μl of protein G-Sepharose slurry was added and the incubation was continued for additional 90 min at 4 ºC. The beads were washed four times with PBS containing 0.1% NP-40, once with PBS and the proteins were eluted with 1% SDS at room temperature. The eluted proteins were precipitated with acetone and subsequently analyzed by SDS-PAGE and Western blotting.

SDS-PAGE and Western blotting

Protein extracts were separated by SDS-PAGE using the Mini-Protean II system (BioRad) and transferred to polyvinylidenefluoride (PVDF) membranes (Millipore) in Tris-glycine buffer with 0.02% SDS and 4 M urea using a semi-dry electrophoretic transfer cell (BioRad). The membranes were blocked with 10% non-fat dry milk in PBS. The antibodies were diluted in 0.05% Tween-20 and 1% milk in PBS and antibody incubations were carried out according to standard procedures. The NBT/BCIP system was used to detect secondary antibodies conjugated to alkaline phosphatase. The ECL system (GE Healthcare) was used for chemiluminiscent detection of horseradish peroxidase.

Chromatin immunoprecipitation

ChIP analyses were performed as described by Takahashi et al. [9]. Chromatin was prepared from S2 cells after cross-linking with 2% formaldehyde. The chromatin was sheared by sonication to a DNA size of 250-1000 bp and pre-cleared. Chromatin fragments were precipitated with antibodies against either rBrg1 (Ab1) or Pol-II (Abcam) in 10 mM Tris-HCl at pH 8.0, 1 mM EDTA, 0.5 mM EGTA containing 0.2% DOC and 1% Triton X-100. The bound complexes were pulled-down with protein A/G-Sepharose beads (50% of each). The beads were washed in RIPA buffer (50 mM Hepes at pH 7.6, 1 mM EDTA, 0.7% DOC, 1% NP40, 0.5 M NaCl). Two controls were used: beads only without primary antibody, and rabbit IgG (DakoCytomation). The precipitated DNA fragments were purified and amplified by PCR with specific primers for the CG8092, CG8421 and CG9380 genes. Actin 5C (CG4027) was used as a control. The PCR conditions were optimized to avoid saturation.

Microarray data analysis

The microarray data was extracted from Array Express, experiment E-TABM 169 carried out by Moshkin et al. [10] (http://www.ebi.ac.uk/microarray-as/aew/). *Drosophila* Genome 2.0 Arrays (Affymetrix) were hybridized with total RNA purified from *Drosophila* S2 cells treated with dsRNA corresponding to dBrm or to other subunits of SWI/SNF. Mock RNAi experiments were carried out in parallel [10]. We selected genes with multiple probe sets (974 genes). Among them, we selected genes that showed changed expression levels specific for a subset of alternatively processed transcripts with p values below 0.02.

RNA interference in S2 Cells

Double-stranded RNAs (dsRNAs) against dBrm and GFP were prepared by *in vitro* transcription from PCR products with T7 promoters on both ends of the amplimers, using the Megascript RNAi kit (Ambion). The sequences of the PCR primers are provided below. The RNAi treatment was performed as described by Clemens et al. [11]. In brief, 20 μg of dsRNA was applied to S2 cells and the cells were harvested after 48 h. Total RNA from S2 cells was extracted using the RNAqueous kit (Ambion). Reverse transcription was performed with Superscript-III (Invitrogen) and hexamer primers (Roche) on 5 μg total RNA. The resulting cDNA was used as a template for PCR reactions using primers specific for selected transcripts. Each RNAi experiment was repeated three times to confirm the reproducibility of the observations.

6

Primer sequences

<u>Primers used for the synthesis of dsRNA</u>

Brm-dsRNA (two different regions targeted)
BKN F: TGAACTGTATCAGCCGCTTG
BKN R: AAGCCCAATCGCATTACAAC
HFA F: GTTTCGCTGTACAATAACAATC
HFA R: ATGTGGAGCAGGACTTAAAG

GFP-dsRNA:
GFP F: ATGGTGAGCCAAGGGCGAGGAGCTG
GFP R: GCGGTCACGAACTCCAGCAG

<u>Primers used for RT-PCR analysis</u>

Actin F: GAACCACTCCCACCAAGAAA
Actin R: TGATCACTTTCAGCACTGGC
18S F: AGCTAGCAATTGGGTGTAGC
18S R: TGAGTCTCGTTCGTTATCGG
Asph1 F1: CCTGGCTCAGCAATTCGAAAA
Asph1 R1: CAACTCTTGCTCGAAAGGAT
Asph1 F2: GATCGAAGAATCAGTAGAGCCG
Asph1 REE2: TCTTTCGGCTTTATTTGTGATGCCAACC
CG9380 F1: GCGAATCGCTATTTGACCAT
CG9380 REE1: TCCATATTTCGCACATACGT
CG9380 F2: CAACTACCAGTGTCGAATGGCAAT
CG9380 REE2: AACGATTGTGGCACATACGT
G8092 F1: TTATGCGGAGAAGTCCATCC
CG8092 R1: AGTGAGAAACGCAGGCACTT
CG8092 REE: TGGTTGCCACCTGATACAGAC

<u>Primers used for ChIP</u>

Actin F: GAACCACTCCCACCAAGAAA
Actin R: TGATCACTTTCAGCACTGGC
CG8092 Prox F: ATCCACGCAAGGAAAGCTAA
CG8092 Prox R: ACCCAGAACTCCATGTCCAG
CG8092 Middle F: CAACCAATCCGTTCAAGGAC
CG8092 Middle R: CTCACGTGTTGGGTTCGATT
CG8092 End F: CATATTCATCGCACATTGGC
CG8092 End R: AATCCCAACTCACATCCAGC
CG8421 Prox F: CGCTCGAAAAGGAAAAACAA
CG8421 Prox R: TCACACCAAATAACGGTCCA
CG8421 Middle F: GTCGCAACCAAAAGGTGTCT
CG8421 Middle R: GGGTCTCTTGAAATGAGCGT
CG8421 End F: GCTTTAGAACCTGGG´CGTGAG
CG8421 End R: ACCGGTCAAATGGGTGATAA
CG8421 Inter F: TGGGTGTGGGTGAATGTATG
CG8421 Inter R: CAGATGGGCGTGTAAATGTG
CG9380 Prox F: GGTGTGGTAGGATGCGGTAG
CG9380 Prox R: TGAAACTTTGCAATCGGTGA
CG9380 Middle F: GCGAATCGCTATTTGACCAT
CG9380 Middle R: CAAGTCTGGACACTGGGTGA
CG9380 End F: CAACTACCAGTGTCGAATGGCAAT
CG9380 End R: AACGATTGTGGCACATACGT

References

1. Meyer, B., Mähr, R., Eppenberger, H.M., and M. Lezzi. 1983. The activity of Balbiani rings 1 and 2 in salivary glands of Chironomus tentans larvae under different modes of development and after pilocarpine treatment. Dev Biol. 98:265-277.

2. Wyss, C., 1982. Chironomus tentans epithelial cell lines sensitive to ecdysteroids, juvenile hormone, insulin and heat shock. Exp Cell Res. 139 : 309-319.

3. Östlund Farrants, A.K., Blomquist, P., Kwon, H., and O. Wrange. 1997. Glucocorticoid receptor-glucocorticoid response element binding stimulates nucleosome disruption by the SWI/SNF complex. Mol Cell Biol. 17:895-905.

4. Zraly, C.B., Marenda, D.R., Nanchal, R., Cavalli, G., Muchardt, C., and Dingwall, A.K. (2003). SNR1 is an essential subunit in a subset of Drosophila brm complexes, targeting specific functions during development. Dev. Biol. *253*, 291-308.

5. Kiseleva, E., Visa, N., Wurtz, T., and B. Daneholt. 1997. Immunocytochemical evidence for a stepwise assembly of Balbiani ring premessenger ribonucleoprotein particles. Eur J Cell Biol. 74:407-416.

6. Wurtz, T., Kiseleva, E., Nacheva, G., Alzhanova-Ericcson, A., Rosén, A., and B. Daneholt. 1996. Identification of two RNA-binding proteins in Balbiani ring premessenger ribonucleoprotein granules and presence of these proteins in specific subsets of heterogeneous nuclear ribonucleoprotein particles.Mol Cell Biol. 16:1425-1435.

7. Lerner, E.A., Lerner, M.R., Janeway, C.A. Jr, and J.A. Steitz. 1981. Monoclonal antibodies to nucleic acid-containing cellular constituents: probes for molecular biology and autoimmune disease. Proc Natl Acad Sci U S A. 78:2737-2741.

8. Falk, R., C. Agaton, E. Kiesler, S. Jin, L. Wieslander, N. Visa, S. Hober, and S. Stahl. 2003. An improved dual-expression concept, generating high-quality antibodies for proteomics research. Biotechnol. Appl. Biochem. 38:231-239.

9. Takahashi ,Y., Rayman, J.B., Dynlacht, B.D. 2000. Analysis of promoter binding by the E2F and pRB families in vivo: distinct E2F proteins mediate activation and repression. Genes Dev 14: 804-816.

10. Moshkin, Y.M., Mohrmann, L., van Ijcken, W.F., and C.P. Verrijzer. 2007. Functional differentiation of SWI/SNF remodelers in transcription and cell cycle control. Mol Cell Biol. 27:651-661.

11. Clemens, J. C., Worby, C. A., Simonson-Leff, N., Muda, M., Maehama, T., Hemmings, B. A., Dixon, J. E. 2000. Use of double-stranded RNA interference in Drosophila cell lines to dissect signal transduction pathways. Proc. Natl. Acad. Sci. U. S. A. 97: 6499-6503.

8

8. Discussion

SWI/SNF is known for its function in transcription regulation by activation and repression of several genes. The complex regulates transcription by modulating chromatin structure in an ATP dependent manner. Recently it has been found that hBrm acts as a regulator of alternative splicing. This suggests that Brm can play roles other than transcription regulator as a chromatin remodeler. The aim of this study is to investigate the role of ctBrm and dBrm in pre-mRNA processing. The results from immuno-EM, immunofluorescence and biochemical fractionation clearly show that Brm is associated with nascent pre-mRNPs. This suggests that Brm regulates gene expression.

The biochemical fractionation studies in human cells also show that several of the core subunits of the SWI/SNF complex are associated with the RNP fraction. The microarray hybridization studies in S2 cells of *D. melanogaster* also show that SWI/SNF regulates the abundance of alternatively processed transcripts. These results suggest that the core SWI/SNF complex plays an important role in pre-mRNA processing. The microarray hybridization studies also show that dBrm affects both alternative splicing and alternative polyadenylation site usage. The alternative processing events regulated by SWI/SNF involve the use of mutually exclusive splicing and polyadenylation sites. The use of the proximal site is favored by dBrm in some of the genes, as expected if dBrm acts by reducing the elongation rate of the Pol-II at certain positions (proposed for the regulation of the CD44 pre-mRNA in human cells (Batsché *et al.,* 2006). However, depletion of dBrm has the opposite effect in other cases (such as CG18251, CG9380, CG3665), which is difficult to explain by modulation of Pol-II elongation rate. This observation that a significant fraction of Brm is associated with nascent pre-mRNPs, leads to propose that dBrm, as part of the pre-mRNP complex, regulates pre-mRNA processing by influencing directly the structure of the pre-mRNP and/or its interactions with the pre-mRNA processing machineries. The action of Brm and the specific outcome of the processing reactions will depend in each case on the specific features of the pre-mRNP. The same type of context-dependent mechanism has been proposed to explain the complex function of SWI/SNF in transcription regulation (Zang *et al*., 2007 and references therein). SWI/SNF acts as a co-activator in the transcription of certain

genes, but represses the transcription of certain other genes (Wang *et al*., 2004), and these two opposite effects are mediated by interactions with different types of co-regulators. The presence of SWI/SNF on RNA may favor the recruitment of splicing factors to the splice site. Alternatively, it may block the recruitment of splicing factors and/or favor the recruitment of polyadenylation factors.

In summary the results show that SWI/SNF is incorporated in the nascent pre-mRNP and regulates processing of a subset of pre-mRNA in *D. melanogaster.* Since a large number of genes are transcribed into multiple alternative transcripts the post-transcriptional role of SWI/SNF has a vast potential to regulate genes. Depletion of SWI/SNF does not affect the levels of all the mRNA species originated from a pre-mRNA in the 15 genes identified, but only a few. This indicates that it is not the synthesis of the pre-mRNA that is affected but the processing into alternative mRNAs. SWI/SNF appears to act both transcriptionally and post-transcriptionally to fine-tune the expression of genes. In this way, SWI/SNF can regulate gene expression by determining the amount of mRNA synthesized from a given promoter and also by regulating the type of alternative transcripts produced.

9. References

Alkhatib SG and Landry JW (2011). The nucleosome remodeling factor. *FEBS Lett.* 20:3197-3207.

Allo M, Buggiano V, Fededa JP, Petrillo E, Schor I, de la Mata M, Agirre E, Plass M, Eyras E and Abou Elela S (2009). Control of alternative splicing through siRNA-mediated transcriptional gene silencing. *Nat Struct Mol Biol*, 16:717-U743.

Auboeuf D, Batché E, Dutertre M, Muchardt C and O'Malley BW (2007) Coregulators: transducing signal from transcription to alternative splicing. *Trends Endocrinol Metab*. 18: 122-129.

Auboeuf D, Hönig A, Berget SM and O'Malley BW (2002) Coordinate regulation of transcription and splicing by steroid receptor coregulators. *Science*. 298: 416-419.

Balbiani EG (1881). "Sur la structure du noyau des cellules salivaires chez les larves de Chironomus". *Zool. Anz.* 4: 637–641.

Banerjee T and Chakravarti D (2011). A peek into the complex realm of histone phosphorylation. *Mol Cell Biol*. 24: 4858-4873.

Bannister AJ and Kouzarides T (2005) Reversing histone methylation. *Nature*. 436: 1103-1106.

Bannister AJ and Kouzarides T (2011). Regulation of chromatin by histone modifications. *Cell Res*. 3:381-395.

Barabino SM and Keller W (1999) Last but not least: regulated poly(A) tail formation. *Cell*. 99: 9-11.

Barak O, Lazzaro MA, Cooch NS, Picketts DJ and Shiekhattar R (2004) A tissue-specific, naturally occurring human SNF2L variant inactivates chromatin remodeling. *J Biol Chem*. 279: 45130-45138.

Barreau C, Paillard L and Osborne HB (2005). AU-rich elements and associated factors: are there unifying principles? *Nucleic Acids Res.* 33: 7138–7150.

Barski A, Cuddapah S, Cui K, Roh TY, Schones DE, Wang Z, Wei G, Chepelev I, and Zhao K (2007). High-resolution profiling of histone methylations in the human genome. *Cell* 129: 823–837.

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

Baurén G, Belikov S and Wieslander L (1998) Transcriptional termination in the Balbiani ring 1 gene is closely coupled to 3'-end formation and excision of the 3'-terminal intron. *Genes Dev*. 12: 2759-2769.

Belikov S, Gelius B and Wrange O (2001) Hormone-induced nucleosome positioning in the MMTV promoter is reversible. *EMBO J*. 20: 2802-2811.

Bentley D (2002) The mRNA assembly line: transcription and processing machines in

the same factory. *Curr Opin Cell Biol*. 14: 336-342.

Berger SL (2007). The complex language of chromatin regulation during transcription. *Nature*. 7143:407-412.

Borah S, Wong AC, Steitz JA (2009). Drosophila hnRNP A1 homologs Hrp36/Hrp38 enhance U2-type versus U12-type splicing to regulate alternative splicing of the prospero twintron. *Proc Natl Acad Sci U S A*. 106:2577-2582.

Bouazoune K and Brehm A (2006) ATP-dependent chromatin remodeling complexes in Drosophila. *Chromosome Res*. 14: 433-449.

Brehm A, Längst G, Kehle J, Clapier CR, Imhof A, Eberharter A, Müller J and Becker PB (2000) dMi-2 and ISWI chromatin remodelling factors have distinct nucleosome binding and mobilization properties. *EMBO J*. 19: 4332-4341.

Brodsky AS, Meyer CA, Swinburne IA, Giles H, Keenan BJ, Liu XLS, Fox EA and Silver PA (2005) Genomic mapping of RNA polymerase II reveals sites of co-transcriptional regulation in human cells. *Genome Biol* :6.

Carrozza MJ, Utley RT, Workman JL and Côté J (2003) The diverse functions of histone acetyltransferase complexes. *Trends Genet*. 19: 321-329.

Case ST and Daneholt B (1978) The size of the transcription unit in Balbiani ring 2 of Chironomus tentans as derived from analysis of the primary transcript and 75 S RNA. *J Mol Biol*. 124: 223-41.

Chodavarapu RK, Feng S, Bernatavichute YV, Chen PY, Stroud H, Yu Y, Hetzel JA, Kuo F, Kim J, Cokus SJ, Casero D, Bernal M, Huijser P, Clark AT, Krämer U, Merchant SS, Zhang X, Jacobsen SE and Pellegrini M (2010). Relationship between nucleosome positioning and DNA methylation. *Nature*. 7304:388-392.

Cramer P, Srebrow A, Kadener S, Werbajh S, de la Mata M, Melen G, Nogués G and Kornblihtt AR (2001) Coordination between transcription and pre-mRNA processing. *FEBS Lett*. 498: 179-182.

Cuthbert GL, Daujat S, Snowden AW, Erdjument-Bromage H, Hagiwara T, Yamada M, Schneider R, Gregory PD, Tempst P, Bannister AJ and Kouzarides T (2004) Histone deimination antagonizes arginine methylation. *Cell*. 118: 545-553.

Danckwardt S, Hentze MW and Kulozik AE (2008) 3' end mRNA processing: molecular mechanisms and implications for health and disease. *EMBO J*. 27: 482-498.

Daneholt B and Hosick H (1973) Evidence for transport of 75S RNA from a discrete chromosome region via nuclear sap to cytoplasm in Chironomus tentans. *Proc Natl Acad Sci USA*.70: 442-446.

Daneholt B, Case ST, Hyde J, Nelson L and Wieslander L (1976) Production and fate of Balbiani ring products. *Prog Nucleic Acid Res Mol Biol*. 19: 319-334.

Daneholt, B and Edström, JE (1967). The content of deoxyribonucleic acid in individual polytene chromosomes of Chironomus tentans. *Cytogen.* 6, 350-356.

Di Giammartino DC, Nishida K and Manley JL (2011). Mechanisms and consequences of alternative polyadenylation. *Mol Cell*. 43: 853-866.

Eberharter A and Becker PB (2004) ATP-dependent nucleosome remodelling: factors and functions. *J Cell Sci*. 117: 3707-3711.

Eisen JA, Sweder KS and Hanawalt PC (1995) Evolution of the SNF2 family of proteins: subfamilies with distinct sequences and functions. *Nucleic Acids Res.* 23: 2715-2723.

Emre NC, Ingvarsdottir K, Wyce A, Wood A, Krogan NJ, Henry KW, Li K, Marmorstein R, Greenblatt JF, Shilatifard A and Berger SL (2005) Maintenance of low histone ubiquitylation by Ubp10 correlates with telomere-proximal Sir2 association and gene silencing. *Mol Cell*. 17: 585-594.

Ericsson C, Grossbach U, Björkroth B and Daneholt B (1990) Presence of histone H1 on an active Balbiani ring gene. *Cell*. 60: 73-83.

Espel E (2005). The role of the AU-rich elements of mRNAs in controlling translation. *Semin. Cell Dev. Biol*. 16: 59–67.

Feng Q and Zhang Y (2001) The MeCP1 complex represses transcription through preferential binding, remodeling, and deacetylating methylated nucleosomes. *Genes Dev*. 15: 827-832.

Gardner RG, Nelson ZW and Gottschling DE (2005) Ubp10/Dot4p regulates the persistence of ubiquitinated histone H2B: distinct roles in telomeric silencing and general chromatin. *Mol Cell Biol*. 25: 6123-6139.

Gibney ER and Nolan CM (2010). Epigenetics and gene expression. *Heredity (Edinb)*. 1:4-13.

Graveley BR (2000) Sorting out the complexity of SR protein functions. *RNA*. 6: 1197-1211.

Graveley BR (2001) Alternative splicing: increasing diversity in the proteomic world. *Trends Genet*. 17: 100-107.

Gregory PD, Wagner K and Hörz W (2001) Histone acetylation and chromatin remodeling. *Exp Cell Res*. 265: 195-202.

Grüne T, Brzeski J, Eberharter A, Clapier CR, Corona DF, Becker PB, and Müller CW (2003) Crystal structure and functional analysis of a nucleosome recognition module of the remodeling factor ISWI. *Mol Cell*. 12: 449-460.

Halees AS, El-Badrawi R and Khabar KS (2008). ARED Organism: expansion of ARED reveals AU-rich element cluster variations between human and mouse. *Nucleic Acids Res*. 36 (Database issue), D137–D140.

Hassa PO, Haenni SS, Elser M and Hottiger MO (2006) Nuclear ADP-ribosylation reactions in mammalian cells: where are we today and where are we going? *Microbiol Mol Biol Rev*.70: 789-829.

Hertel KJ (2008) Combinatorial control of exon recognition. *J Biol Chem*. 283: 1211-1215.

Hertel KJ and Graveley BR (2005) RS domains contact the pre-mRNA throughout spliceosome assembly. *Trends Biochem Sci*. 30: 115-118.

Holste D and Ohler U (2008) Strategies for identifying RNA splicing regulatory motifs and predicting alternative splicing events. *PLoS Comput Biol*. 2008 Jan;4(1):e21.

Horn PJ and Peterson CL (2001) The bromodomain: a regulator of ATP-dependent chromatin remodeling. *Front Biosci*. 6: 1019-1023.

House AE and Lynch KW (2008) Regulation of alternative splicing: more than just the ABCs. *J Biol Chem*. 283: 1217-1221.

Iizuka M and Smith MM (2003) Functional consequences of histone modifications. *Curr Opin Genet Dev*. 13: 154-160.

Jaenisch R, Bird A. (2003). Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. *Nat Genet*. 33: 245-254.

Jenuwein T and Allis CD (2001) Translating the histone code. *Science*. 293: 1074-1080.

Johnson CN, Adkins NL and Georgel P (2005) Chromatin remodeling complexes: ATP-dependent machines in action. *Biochem Cell Biol*. 83: 405-417.

Kelley DE, Stokes DG and Perry RP (1999). CHD1 interacts with SSRP1 and depends on both its chromodomain and its ATPase/helicase-like domain for proper association with chromatin. *Chromosoma*.1: 10-25.

Kiesler E, Miralles F and Visa N (2002) HEL/UAP56 binds cotranscriptionally to the Balbiani ring pre-mRNA in an intron-independent manner and accompanies the BR mRNP to the nuclear pore. *Curr Biol*. 12: 859-862.

Kiseleva E, Visa N, Wurtz T and Daneholt B (1997) Immunocytochemical evidence for a stepwise assembly of Balbiani ring premessenger ribonucleoprotein particles. *Eur J Cell Biol*. 74: 407-416.

Kiseleva E, Wurtz T, Visa N and Daneholt B (1994). Assembly and disassembly of spliceosomes along a specific pre-messenger RNP fiber. *EMBO J.* 13: 6052-6061.

Kolasinska-Zwierz P, Down T, Latorre I, Liu T, Liu XS and Ahringer J (2009). Differential chromatin marking of introns and expressed exons by H3K36me3. *Nat Genet*, 41:376-381.

Kornblihtt AR (2005) Promoter usage and alternative splicing. *Curr Opin Cell Biol*. 17: 262-268.

Kornblihtt AR (2006) Chromatin, transcript elongation and alternative splicing. *Nat Struct Mol Biol*. 13: 5-7.

Kornblihtt AR (2007) Coupling transcription and alternative splicing. *Adv Exp Med Biol*. 623: 175-189.

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

Krämer A (1996) The structure and function of proteins involved in mammalian pre-mRNA splicing. *Annu Rev Biochem*. 65: 367-409.

Krogan NJ, Keogh MC, Datta N, Sawa C, Ryan OW, Ding H, Haw RA, Pootoolal J, Tong A, Canadien V, Richards DP, Wu X, Emili A, Hughes TR, Buratowski S and Greenblatt JF (2003) A Snf2 family ATPase complex required for recruitment of the histone H2A variant Htz1. *Mol Cell*. 12: 1565-1576.

Kwon CS and Wagner D (2007) Unwinding chromatin for development and growth: a few genes at a time. *Trends Genet*. 23: 403-412.

Lachner M and Jenuwein T (2002) The many faces of histone lysine methylation. *Curr Opin Cell Biol*. 14: 286-298.

Lamb MM and Daneholt B (1979). Characterization of active transcription units in Balbiani rings of Chironomus tentans. *Cell.* 17: 835-848.

Längst G and Becker PB (2004) Nucleosome remodeling: one mechanism, many phenomena. *Biochim Biophys Acta*. 1677: 58-63.

Längst G, Becker PB (2001). ISWI induces nucleosome sliding on nicked DNA. *Mol Cell*. 5:1085-1092.

Lee MS, Edwards RA, Thede GL and Glover JN (2005). Structure of the BRCT repeat domain of MDC1 and its specificity for the free COOH-terminal end of the gamma-H2AX histone tail. *J Biol Chem*. 37:32053-32056.

Legendre M, Ritchie W, Lopez F, and Gautheret D (2006). Differential Repression of Alternative Transcripts: A Screen for miRNA Targets. *PLoS Comput Biol*. 2(5): e43.

Li G and Reinberg D (2011). Chromatin higher-order structures and gene regulation. *Curr Opin Genet Dev*.2:1 75-186.

Lou H and Gagel RF (2001) Alternative ribonucleic acid processing in endocrine systems. *Endocr Rev*. 22: 205-225.

Luco RF and Misteli T (2011). More than a splicing code: integrating the role of RNA, chromatin and non-coding RNA in alternative splicing regulation. *Curr Opin Genet Dev*. 4:366-372.

Luco RF, Allo M, Schor IE, Kornblihtt AR and Misteli T (2011). Epigenetics in alternative pre-mRNA splicing. *Cell*, 144:16-26.

Luco RF, Pan Q, Tominaga K, Blencowe BJ, Pereira-Smith OM and Misteli T (2010). Regulation of alternative splicing by histone modifications. *Science*, 327:996-1000.

Lusser A and Kadonaga JT (2003) Chromatin remodeling by ATP-dependent molecular machines. *Bioessays*. 25: 1192-1200.

Maeshima K, Hihara S, Eltsov M (2010). Chromatin structure: does the 30-nm fibre exist in vivo? *Curr Opin Cell Biol.* 3: 291-297.

Mandel CR, Kaneko S, Zhang H, Gebauer D, Vethantham V, Manley JL and Tong L (2006) Polyadenylation factor CPSF-73 is the pre-mRNA 3'-end-processing endonuclease. *Nature*. 444: 953-956.

Maniatis T and Reed R (2002) An extensive network of coupling among gene expression machines. *Nature*. 416: 499-506.

Martens JA and Winston F (2003) Recent advances in understanding chromatin remodeling by Swi/Snf complexes. *Curr Opin Genet Dev*. 13: 136-142.

Mavrich TN, Ioshikhes IP, Venters BJ, Jiang C, Tomsho LP, Qi J, Schuster SC, Albert I and Pugh BF (2008). A barrier nucleosome model for statistical positioning of nucleosomes throughout the yeast genome. *Genome Res*. 18: 1073–1083.

Mehlin H, Daneholt B and Skoglund U (1992) Translocation of a specific premessenger ribonucleoprotein particle through the nuclear pore studied with electron microscope tomography. *Cell*. 69: 605-613.

Mellor J (2006) Dynamic nucleosomes and gene transcription. *Trends Genet*. 22: 320-329.

Mellor J and Morillon A (2004) ISWI complexes in Saccharomyces cerevisiae. *Biochim Biophys Acta*. 1677: 100-112.

Mellor J, Dudek P and Clynes D (2008). A glimpse into the epigenetic landscape of gene regulation. *Curr Opin Genet Dev*. 2:116-122.

Messner S and Hottiger MO (2011). Histone ADP-ribosylation in DNA repair, replication and transcription. *Trends Cell Biol*. 9: 534-542.

Metzger E, Wissmann M, Yin N, Müller JM, Schneider R, Peters AH, Günther T, Buettner R and Schüle R (2005) LSD1 demethylates repressive histone marks to promote androgen-receptor-dependent transcription. *Nature*. 437: 436-439.

Mohrmann L, Langenberg K, Krijgsveld J, Kal AJ, Heck AJR and Verrijzer CP (2004). Differential targeting of two distinct SWI/SNF-related *Drosophila.* chromatin-remodeling complexes. *Mol. Cell. Biol.* 24: 3077-3088.

Moshkin YM, Mohrmann L, van Ijcken WF, Verrijzer CP (2007). Functional differentiation of SWI/SNF remodelers in transcription and cell cycle control. *Mol Cell Biol.* 27: 651-61.

Muchardt C (2010). A few thoughts about catalyzing transcription. *Epigenomics*. 1:1-3

Nathan D, Ingvarsdottir K, Sterner DE, Bylebyl GR, Dokmanovic M, Dorsey JA, Whelan KA, Krsmanovic M, Lane WS, Meluh PB, Johnson ES and Berger SL (2006) Histone sumoylation is a negative regulator in Saccharomyces cerevisiae and shows dynamic interplay with positive-acting histone modifications. *Genes Dev*. 20: 966-976.

Nelson CJ, Santos-Rosa H and Kouzarides T (2006) Proline isomerization of histone H3 regulates lysine methylation and gene expression. *Cell*. 126: 905-916.

Nogues G, Kadener S, Cramer P, Bentley D and Kornblihtt AR (2002) Transcriptional activators differ in their abilities to control alternative splicing. *J Biol Chem*. 277: 43110-43114.

Nowak SJ and Corces VG (2004) Phosphorylation of histone H3: a balancing act

between chromosome condensation and transcriptional activation. *Trends Genet.* 20: 214-220.

Ostlund Farrants AK, Blomquist P, Kwon H, Wrange O (1997). Glucocorticoid receptor-glucocorticoid response element binding stimulates nucleosome disruption by the SWI/SNF complex. *Mol Cell Biol*. 17: 895-905.

Paulsson G, Lendahl U, Galli J, Ericsson C and Wieslander L (1990) The Balbiani ring 3 gene in Chironomus tentans has a diverged repetitive structure split by many introns. *J Mol Biol*. 211: 331-349.

Perales R, Bentley D (2009). "Cotranscriptionality": the transcription elongation complex as a nexus for nuclear transactions. *Mol Cell*. 362: 178-191.

Percipalle P, Zhao J, Pope B, Weeds A, Lindberg U, Daneholt B (2001). Actin bound to the heterogeneous nuclear ribonucleoprotein hrp36 is associated with Balbiani ring mRNA from the gene to polysomes. *J Cell Biol*. 153: 229-236.

Proudfoot N (2004) New perspectives on connecting messenger RNA 3' end formation to transcription. *Curr Opin Cell Biol*. 16: 272-278.

Proudfoot NJ, Furger A and Dye MJ (2002) Integrating mRNA processing with transcription. *Cell.* 108: 501-512.

Reinke H and Hörz W (2003) Histones are first hyperacetylated and then lose contact with the activated PHO5 promoter. *Mol Cell*. 11: 1599-1607.

Richard P and Manley JL (2009). Transcription termination by nuclear RNA polymerases. *Genes Dev*. 11:1247-69.

Ryan K, Calvo O and Manley JL (2004) Evidence that polyadenylation factor CPSF-73 is the mRNA 3' processing endonuclease. *RNA*. 10: 565-573.

Saint-Andre V, Batsché E, Rachez C and Muchardt C (2011). Histone H3 lysine 9 trimethylation and HP1gamma favor inclusion of alternative exons. *Nat Struct Mol Biol*, 18:337-344.

Sandberg R, Neilson JR, Sarma A, Sharp PA, and Burge CB (2008). Proliferating cells express mRNAs with shortened 3′untranslated regions and fewer microRNA target sites. *Science* 320: 1643–1647.

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

Schor IE, Rascovan N, Pelisch F, Allo M and Kornblihtt AR (2009). Neuronal cell depolarization induces intragenic chromatin modifications affecting NCAM alternative splicing. *Proc Natl Acad Sci U S A*, 106:4325-4330.

Schwartz S, Meshorer E and Ast G (2009). Chromatin organization marks exon-intron structure. *Nat Struct Mol Biol*. 9:990-995.

Shatkin AJ and Manley JL (2000) The ends of the affair: capping and polyadenylation. *Nat Struct Biol*. 7: 838-842.

Shi Y, Lan F, Matson C, Mulligan P, Whetstine JR, Cole PA, Casero RA and Shi Y

(2004) Histone demethylation mediated by the nuclear amine oxidase homolog LSD1. *Cell*. 119: 941-953.

Shilatifard A (2006) Chromatin modifications by methylation and ubiquitination: implications in the regulation of gene expression. *Annu Rev Biochem*. 75: 243-69.

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

Sims RJ, Belotserkovskaya R and Reinberg D (2004) Elongation by RNA polymerase II: the short and long of it. *Genes Dev*. 18: 2437-2468.

Sims RJ, Millhouse S, Chen CF, Lewis BA, Erdjument-Bromage H, Tempst P, Manley JL and Reinberg D (2007). Recognition of trimethylated histone h3 lysine 4 facilitates the recruitment of transcription postinitiation factors and pre-mRNA splicing. *Mol Cell*, 28:665-676.

Sims RJ, Nishioka K and Reinberg D (2003) Histone lysine methylation: a signature for chromatin function. *Trends Genet*. 19: 629-639.

Sjölinder M, Björk P, Söderberg E, Sabri N, Farrants AK and Visa N (2005) The growing pre-mRNA recruits actin and chromatin-modifying factors to transcriptionally active genes. *Genes Dev*. 19: 1871-1884.

Skoglund U, Andersson K, Strandberg B and Daneholt B (1986) Three-dimensional structure of a specific pre-messenger RNP particle established by electron microscope tomography. *Nature*. 319: 560-564.

Spencer VA and Davie JR (1999) Role of covalent modifications of histones in regulating gene expression. *Gene*. 240: 1-12.

Spies N, Nielsen CB, Padgett RA and Burge CB (2009). Biased chromatin signatures around polyadenylation sites and exons. *Mol Cell*, 36:245-254.

Sterner DE and Berger SL (2000) Acetylation of histones and transcription-related factors. *Microbiol Mol Biol Rev*. 64: 435-459.

Subtil-Rodriguez A and Reyes JC (2010). BRG1 helps RNA polymerase II to overcome a nucleosomal barrier during elongation, in vivo. *EMBO Rep,* 11:751-757.

Sudarsanam P and Winston F (2000) The Swi/Snf family nucleosome-remodeling complexes and transcriptional control. *Trends Genet*. 16: 345-351.

Takagaki Y, Seipelt RL, Peterson ML and Manley JL (1996) The polyadenylation factor CstF-64 regulates alternative processing of IgM heavy chain pre-mRNA during B cell differentiation. *Cell*. 87: 941-952.

Thiagalingam S, Cheng KH, Lee HJ, Mineva N, Thiagalingam A and Ponte JF (2003) Histone deacetylases: unique players in shaping the epigenetic histone code. *Ann N Y Acad Sci*. 983: 84-100.

Tian B, Hu J, Zhang H and Lutz CS (2005) A large-scale analysis of mRNA polyadenylation of human and mouse genes. *Nucleic Acids Res*. 33: 201-212.

Truss M, Candau R, Chávez S and Beato M (1995) Transcriptional control by steroid hormones: the role of chromatin. *Ciba Found Symp*.191: 7-17.

Wang F, Zhang B, Faller DV (2004) BRG1/BRM and prohibitin are required for growth suppression by estrogen antagonists. *EMBO J* 23: 2293-2303.

Wang W (2003) The SWI/SNF family of ATP-dependent chromatin remodelers: similar mechanisms for diverse functions. *Curr Top Microbiol Immunol.* 274: 143-169.

Woodcock CL and Ghosh RP (2010). Chromatin higher-order structure and dynamics. *Cold Spring Harb Perspect Biol*. 5: a000596.

Wu JY and Maniatis T (1993) Specific interactions between proteins implicated in splice site selection and regulated alternative splicing. *Cell*. 75: 1061-1070.

Yan J and Marr TG (2005) Computational analysis of 3'-ends of ESTs shows four classes of alternative polyadenylation in human, mouse, and rat. *Genome Res.* 15: 369-375.

Zhang B, Chambers KJ, Faller DV, Wang S (2007) Reprogramming of the SWI/SNF complex for co-activation or co-repression in prohibitin-mediated estrogen receptor regulation. *Oncogene* 26: 7153-7157.

Zhang H, Lee JY and Tian B (2005). Biased alternative polyadenylation in human tissues. *Genome Biol*. 6: R100.

Zhang Y and Reinberg D (2001) Transcription regulation by histone methylation: interplay between different covalent modifications of the core histone tails. *Genes Dev*. 15: 2343-2360.

Zhao J, Jin SB, Björkroth B, Wieslander L and Daneholt B (2002). The mRNA export factor Dbp5 is associated with Balbiani ring mRNP from gene to cytoplasm. *EMBO J.* 21: 1177-1187.

Zraly CB, Marenda DR, Nanchal R, Cavalli G and Muchardt C (2003). SNR1 is an essential subunit in a subset of Drosophila brm complexes, targeting specific functions during development. *Dev Biol* 253: 291-308.

Project II

Role of LSm proteins in translation and replication of positive-strand RNA Viruses

10. Summary

The Sm and LSm proteins are involved in various aspects of RNA metabolism. These families of proteins have the characterstic Sm fold, which is important for protein-protein and protein-RNA interaction. The LSm complexes play important role in both pre-mRNA splicing and mRNA degradation. Additionally, the LSm protein complex has also been found to play an important role in translation and replication of positive strand viruses. The aim of this study was to investigate if LSm1-7 complex is involved in genomic translation and replication of *Flock House virus.* More precisely if LSm1-7 complex is involved in maintaining the genomic/subgenomic ratio of FHV. The results show that LSm1-7 binds to all the regions of RNA1 targeted in this study. Unlike, the results from the study in BMV genome where the binding of LSm1-7 complex is specific to the 3´UTR and internal A-rich sequences, the results from the electro-mobility shift assays in FHV subgenome show that LSm1-7 most likely covers the whole RNA1-sequence.

11. Introduction

Sm and like Sm (LSm) proteins constitute a family of factors, whose members are implicated in various aspects of RNA metabolism. Sm proteins were first identified as the targets of autoantibodies derived from a patient Stephanie Smith, suffering from systematic lupus erythematosis. The hallmark of Sm and LSm protein family is the presence of a characteristic bipartite Sm motif that is separated by a linker of variable length. This sequence folds into a characteristic domain termed the ''Sm fold'', which is capable of mediating contacts to other Sm/Lsm proteins (Khusial *et al.*, 2005).

Figure 4: Sequence comparison of the Sm-fold in Sm/Lsm and Hfq monomers and the evolution of the eukaryotic Sm and LSm family. (A)Gray boxes highlight the common backbone regions defining a minimal Sm-fold. Overall conserved residues appear in orange, those specific to the Hfq family in green and those characteristic to Archaea and eukaryotes in magenta. Blue boxes indicate conserved patches of amino acid residues. (B) Evolution of the eukaryotic Sm and LSm family. Shown in this figure is the evolutionary relationship of the 86 amino acids of the Sm-fold in the 18 characterized eukaryotic Sm and LSm proteins. Proteins that are marked represent the major auto-antigens in systemic lupus erythematosus. (Modified from (A) Sauter *et al.*, 2003 and (B) Khusial *et al.,* 2005).

A common hallmark of these proteins is that they form hexa- or heptameric rings that are capable of interacting with various RNA targets. This has first been established for spliceosomal snRNPs, the major constituents of the pre-mRNA processing spliceosomes. In this case, seven proteins of the Sm class, termed B/B', D1, D2, D3, E, F and G bind to a single-stranded region of the U1, U2, U4/U6 and U5 small nuclear RNAs (snRNAs), and form a the toroidal "Sm core". The LSm proteins likewise form protein rings of different composition that are structurally very similar to the Sm core domain. The LSm rings bind to various coding and non-coding RNAs and influence their function in different ways (reviewed by Wilusz and Wilusz, 2005). (see Figure 4).

11.1 Functions of proteins of the LSm/Sm family

11.1.1 The Bacterial Hfq

Bacteria contain a single LSm protein, called Hfq. The bacterial Hfq protein contains the characteristic Sm motif and is the only protein of the LSm/Sm family know thus far that forms hexameric rather than heptameric ring structures (reviewed in Wilusz and Wilusz, 2005) (see Figure 5). Hfq is considered as an RNA-binding protein, found in diverse bacterial lineages and plays a key role in the regulation of gene expression (Purijn *et al*., 2005). Hfq facilitates short base-pairing interactions of regulatory small RNAs (sRNAs) with target mRNAs. There are several general mechanisms of Hfq mediated regulation at the levels of translation or RNA stability (reviewed by Vogel and Luisi, 2011). First, Hfq can prevent protein synthesis by making 5′ region of the mRNA inaccessible for translation initiation. On the other hand, Hfq can also promote translation by assisting an sRNA to bind to the 5′ region of its target mRNA in order to disrupt a secondary structure inhibiting ribosomal binding. Second, prior to target recognition, Hfq can protect sRNAs from ribonuclease cleavage or present some RNAs in such a way that cleavage of mRNA is promoted. In most cases, the mechanism of action of Hfq seems to be dependent on the structural information present in the RNA with which the protein interacts (reviewed by Vogel and Luisi, 2011).
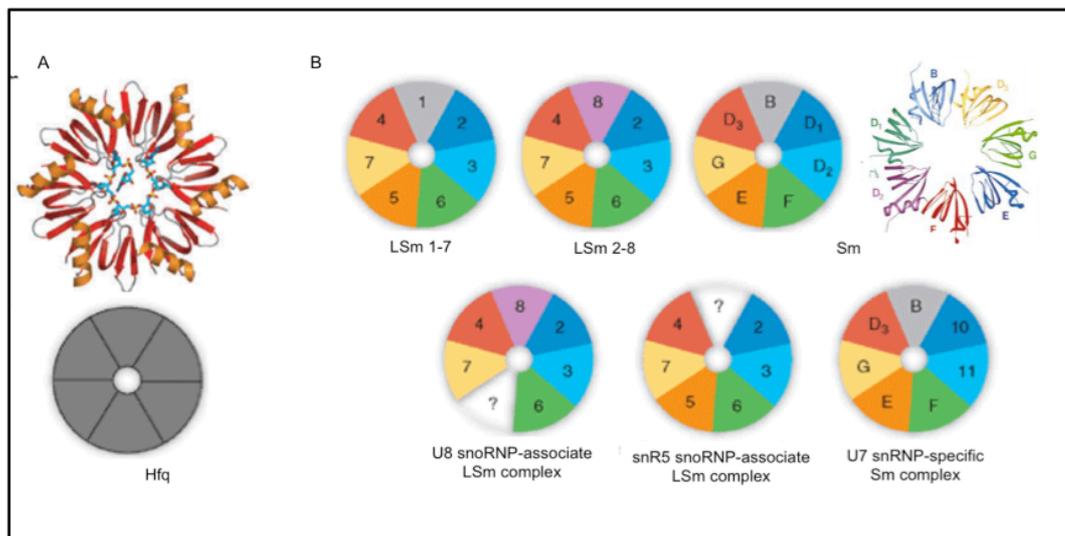
Figure 5: Structure and organization of the Hfq, Sm and LSm protein complexes. (A) Top view shows structure of the Hfq homohexamer from *S. aureus* and below is the schematic representation of Hfq homohexamer (B) Arrangement of LSm and Sm heptamers. The order of the LSm subunits was proposed on the basis of protein interaction and that of the known Sm core. The U8 snoRNA–associated complex and snR5-associated complex are based on the interactions known to occur in other LSm complexes. In U7 snRNP-specific core, LSm10 and LSm11 proteins replace the SmD1 and SmD2. (Modified from Wilusz and Wilusz, 2005)

11.1.2 Sm and LSm Proteins are major building blocks of splicing U snRNPs

The Sm and LSm proteins play important roles in processing of the pre-mRNA during splicing. The newly synthesized pre-mRNA generated by transcription contains intronic sequences. These introns are removed by a multi-subunit complex called the spliceosome to generate the mature mRNAs (reviewed by Kambach *et. al.*, 1999). The spliceosomal snRNPs U1, U2, U4/6 and U5 constitute the major building blocks of the spliceosomes. They bind to the pre-mRNA in a specific order to align the splice sites for cleavage with the help of the Sm proteins. The U1 snRNP binds to the 5' end of a intron and the U2 snRNP binds close to the 3' end of the intron (at the branch point), followed by the binding of the U4/U6 snRNPs and finally the U5 snRNP that helps hold the two exons together. After the intron is spliced out it is rapidly degraded, and the two exons are joined together (for details see section 4.6.2).

Splicosomal snRNPs are composed of a small name-giving snRNA, specific and common protein. The common proteins are the Sm proteins D3-B-D1-D2-F-E-G, which form the structural framework of snRNPs upon binding to the single stranded Sm site. The U4/6 particle is an exception among the snRNPs as this particle contains apart from classical Sm proteins that are bound to U4 snRNA also LSm proteins bound to the 3'-end of the U6 snRNA. While the primary role for the Sm core domain appears to be in facilitating the faithful and efficient biogenesis of snRNPs, the LSm2-8 complex is known to play an important role in pre-mRNA splicing. Although the precise function of LSm2-8 complex in pre-mRNA processing is still unclear experimental evidence suggests that these proteins influence the stability of U6 snRNA. In addition it has been shown that the LSm2-8 ring facilitates conformational changes in U6 snRNAs that promote the formation of higher order splicing complexes such as the U4/6 di-snRNP and the U4/6.U5 tri-snRNP (Achsel *et al*., 1999). This suggests that LSm complexes have a chaperone like activity (reviewed by Wilusz and Wilusz, 2005).

11.1.3 mRNA degradation

In eukaryotes, the process of mRNA decay is often initiated by the shortening and eventually the removal of the poly(A) tail. Two major pathways of decay after deadenylation have been described. Deadenylation is followed by removal of the 5' methyl G cap. Then, the mRNA is decayed either by $5' \rightarrow 3'$ exonuclease or by $3' \rightarrow 5'$ decay mediated by the exosome.

In yeast, it has been shown that a heptameric complex of LSm1-7 proteins can associate with the 3´ end of the deadenylated mRNA. The LSm proteins have been shown to inhibit the trimming of the 3´ end while simultaneously promoting decapping by the Dcp2 enzyme at the 5´ end of the transcript. Once decapped, the body of the mRNA is then digested by a 5´-to-3´ exonuclease Xrn1. In an alternative pathway predominant in mammalian cells, the body of the mRNA is degraded 3´ to 5´ by a large complex known as the exosome and the cap structure is metabolized by the scavenger decapping enzyme.

The LSm1-7 complex plays an important role in mRNA degradation by interacting with both the mRNA and mRNA degradation machinery. It has been shown by Tharun *et al.* that LSm1-7 complex interacts with Dcp1p the decapping

enzyme, Xrn1p the exonuclease and Dcp2p and Pat1p the other two proteins important for mRNA decapping (Tharun *et al.*, 2000). The LSm1-7 has also been shown to co-immunoprecipitate the target mRNA (Tharun *et al.*, 2000). Based on this observation two possible models of function of LSm1-7 complex have been suggested. In the first model it is suggested that the LSm1-7 complex first binds to the target mRNA and then recruits the decapping enzyme to the transcript. In the second model it is proposed that LSm1-7 complex promotes decapping by facilitating the rearrangement of the mRNP structure such that the cap is exposed to the decapping enzyme (reviewed by He and Parker 2001).

11.1.4 Viral RNA stability in host cell

Viruses make use of host cellular translation machinery to replicate in the host cell. Most viruses generate transcripts that mimic cellular mRNAs in having a cap and a poly(A) tail. This includes members of DNA virus family with an exclusively cytoplasmic life cycle. In addition, several viruses including Vaccinia virus have been reported to induce degradation of host-cell mRNAs as a means of reducing competition for the translation machinery (Rice & Roberts, 1983 and Swaminathan, 2005). The question arises how viruses selectively prevent degradation of their own mRNAs while simultaneously ensuring that the host translation machinery can recognize these transcripts efficiently. One example comes from Orthopoxviridae. A unique poly(A) tract is found at the 5´ end of many of its transcripts. The 5´ adenosine tracts are generated due to polymerase 'stuttering' during transcription and it is suggested that 5´ adenosine tracts interfere with the mRNA decay machinery to modulate the stability of viral mRNA (Bergman *et al.*, 2007). They have further shown that LSm1-7 complex interacts with the 5´ poly(A) tract of the viral RNA and prevents exosomal decay. These results indicate importance of the LSm complex in modulating mRNA decay for replication of viruses inside host.

11.2 Viral replication

The bacterial counterpart of LSm proteins, Hfq is known to act as a host factor required for replication of bacteriophage Qß in *E. coli* (Fernandez *et al.*, 2004). It has been shown that specific binding occurs between Hfq and Qß RNA, which then facilitate the association of a replicase at the 3´ end by melting the secondary

structure (Schuppli *et al*., 1997). Intriguingly, LSm proteins not only act in bacteria as cellular cofactors of infectious agents. The first link between LSm proteins and viral propagation came from studies of Bromo mosaic virus (BMV). This virus, which normally infects grasses, can also replicate in the *S. cerevisiae*, allowing the identification of host factors by genetic means. These studies provided a link of viral replication and translation to the LSm1-7 complex.

11.2.1 Positive-strand Viruses

Viruses are divided into seven classes based on their genome and genome replication strategies (reviewed by Ahlquist, 2006). One such class of viruses is positive-strand RNA viruses. The positive-strand viruses are the viruses that contain the genetic information in the form of a translatable messenger-sense RNA. This group includes a large group of plant, animal and human viruses that constitute more than one-third of the virus genera known so far. The positive-strand RNA viruses can be further subdivided into several super-families (reviewed by Ahlquist, 2003). Examples for positive-strand viruses that infect mammals are *Coronaviridae* (eg. SARS coronavirus), *Flaviviridae* (eg. Hepatitis C virus), *Picornaviridae* (eg. Poliovirus) (reviewed by Ahlquist, 2003).

The genome of positive-strand viruses is the template for both translation and replication. This necessitates the interaction with host translation factors for successful translation and subsequent replication. Almost all positive-strand viruses are known to have genes, which code for RNA dependent RNA polymerase to replicate their own genome. This polymerase is not encapsidated in positive-strand viruses. Therefore, the virus cannot replicate until the RNA genome is translated to produce the polymerase and other replication factors (reviewed by Ahlquist *et al*., 2003). To translate the viral genome host factors play an important role. One of the model systems for studying gene expression and RNA replication of the positive strand RNA virus is *Bromo mosaic virus* (BMV). The ability of BMV to replicate in yeast also makes it an ideal model system for positive strand RNA virus biology studies (Alves-Rodrigues *et al*., 2006 and Galão *et al*., 2007). It has been shown that BMV shares fundamental similarities in nucleic acid replication with double stranded-RNA and retro viruses (Schwartz *et al*., 2002). The genome of BMV consists of three RNAs with 5′-terminal m7G-caps and a tRNA-like structure (TLS) located at the end

of the 3′ UTRs (Ahlquist, 1992). The 5′ and 3′ UTRs of all BMV genomic RNAs contain partly overlapping sequences that control translation and initiation of negative-strand (in the 3′ UTR) or positive-strand (in the 5′ UTR) RNA synthesis. A piece of sequence, at the 5′ end of RNA1 and RNA2 called recruitment element (RE) and the intergenic region (IR) of RNA3 are necessary for proper selection and recruitment of the viral RNAs. RNA1 and RNA2 encodes for helicase 1a, and the polymerase 2a, respectively. The only BMV protein required for the recruitment of the BMV genome is protein 1a (Janda and Ahlquist, 1998). RNA3 encodes the movement protein, protein 3a and the coat protein, through a subgenomic RNA generated during replication. Both of these proteins are important for causing systemic infection in plants, but not for replication of virus (Noueiry *et al*., 2003).

Using BMV/yeast system Mas *et al*. previously reported that LSm1-7 complex was required for both the translation and the recruitment of the BMV genome (Diez *et al*., 2000; Noueiry *et al*., 2003 and Mas *et al*., 2006). With the availability of recombinant LSm1-7 complexes, using an *in vitro* binding approach Galão *et al.* have shown that LSm1-7 directly interacts with BMV RNAs via two specific RNA-targeting elements, a tRNA-like structure located in the 3′ UTR and two internal A-rich single-stranded regions. They have shown by *in vivo* analysis that LSm1-7 RNA-target sequences act as regulators of translation and recruitment of the BMV genome (Galão *et al.,* 2010).

11.2.2 *Flock House Virus*

Flock house virus (FHV) is a positive strand virus from Nodavirus family and infects insects (reviewed by Ahlquist *et al*., 2003). Several characteristic features like simplicity of its genome organization with simple genetics, high levels of viral replication and its ability to replicate in a variety of eukaryotic cells makes it a useful model system for the study of positive strand RNA viruses.

The genome of FHV consists of RNA1 (3.1 kb) and RNA2 (1.4 kb), which are packaged into a single virus (Scotti *et al*., 1983 and Venter *et al.,* 2005). The 5'-ends of these RNAs have a 7-methylguanylate (m^7G) cap like cellular mRNAs but their 3′ ends are not polyadenylated (Ball *et at*., 1998). RNA1 codes for a protein called protein A (112 kDa), which is the RNA-dependent RNA polymerase (RdRp) that amplifies both genomic strands as well as a 387 nucleotide subgenomic RNA

(sgRNA), RNA3 (Friesen *et al*., 1982). RNA3 corresponds to the 3' end of RNA1 and codes for proteins B1 and B2. The function of B1 is unknown, where as protein B2 is required for suppression of RNA silencing in infected hosts. RNA2 codes for precursor coat protein, protein alpha (α) which is 43 kDa in size) (Friesen *et al*., 1982).

Translation of protein A is necessary for the synthesis of RNA3 (Freisen *et al*., 1982). The synthesis of RNA3 also requires long-distance base pairing between two *cis* acting elements on RNA1: a short distal subgenome control element (DSCE) located 1.5 kb upstream of the RNA3 start site and a longer proximal subgenome control element (PSCE) located directly upstream of the start site (Lindenbach *et al*., 2002). Computer modeling by Lindenbach *et al*. has predicted that base pairing between DSCE and PSCE results in the formation of two helices. It has been shown that mutations disrupting the helices abolish RNA3 synthesis, while when the helix is restored the synthesis of sgRNA3 is reestablished (Lindenbach *et al*., 2002). RNA3 has an important role in the regulation of FHV gene expression by acting as a transactivator of RNA2 replication (Lindenbach *et al*., 2002 and Eckerle *et al*., 2002). This type of regulation provides a reasonable explanation for the observed organization of the FHV lifecycle into two phases: an early phase in which sufficient levels of protein A molecules are translated to establish functional replication complexes for RNA replication, and a later phase in which coat protein translation from RNA2 is up-regulated to high levels to promote virion assembly (reviewed by Schneemann *et al*., 1998). RNA3 not only transactivates RNA2 replication but is also repressed at the onset of RNA2 replication (Gallagher *et al*., 1983 and Zhong *et al*., 1993). This suggests that RNA3 coordinates the synthesis of RNA1 and RNA2 by a feedback inhibition mechanism. The feedback mechanism is important for viruses with segmented genomes such as FHV, as it ensures the synthesis of genomic RNA segments at optimal ratios for processes like RNA replication, gene expression and virus assembly (Zhong *et al*., 1993).
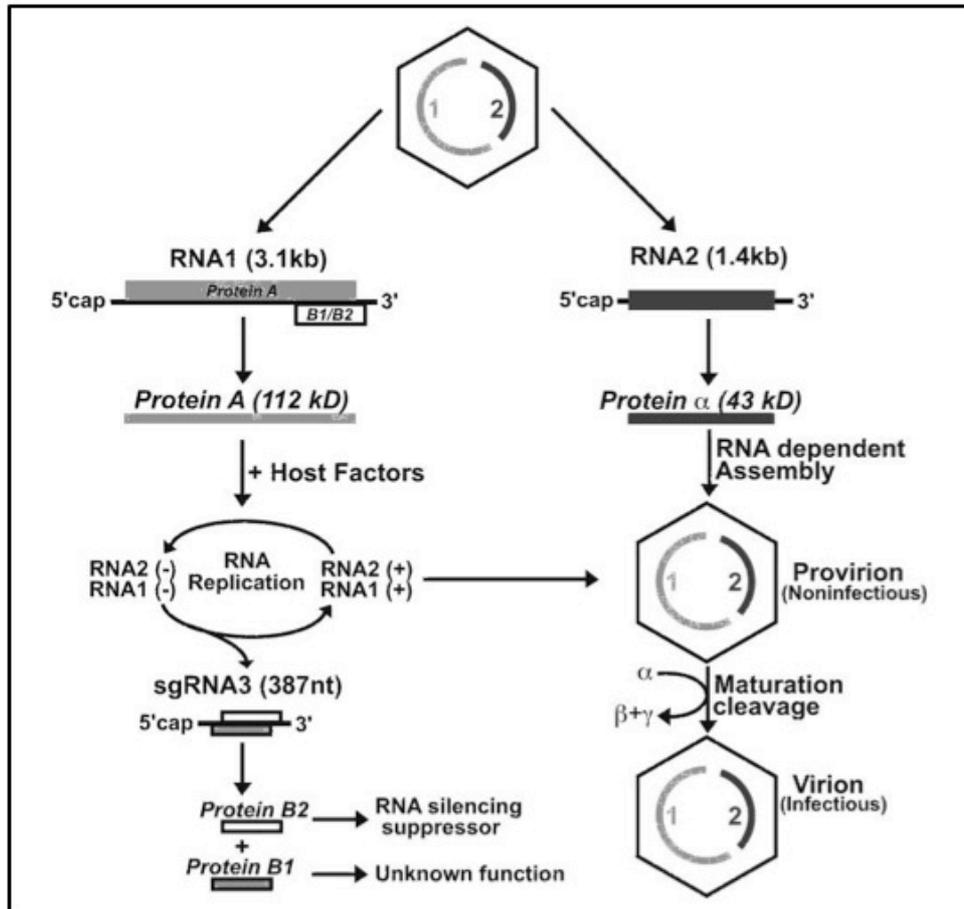
Figure 6: Genome organization of *Flock house Virus*. The genome of FHV contains RNA1 (3.1 kb) and RNA2 (1.4 kb), packed in a single virion. The 5'-ends of these RNAs are capped but their 3' ends are not polyadenylated. RNA1 codes for the synthesis of protein A (112 kDa), which is the RNA-dependent RNA polymerase (RdRp). RdRp is required for the amplification of both genomic strands as well as a 387 nucleotide subgenomic RNA (RNA3). RNA3 corresponds to the 3' end of RNA1 and codes for two proteins B1 and B2. The function of B1 is still unknown and B2 is required for the suppression of RNA silencing in infected hosts. RNA2 encodes the precursor of the coat protein (protein alpha; 43 kDa). (Reproduced from Venter and Schneemann A, 2008).

12. Aim of the Study

The role of LSm proteins as RNA binding factors has been well established both in prokaryotes and eukaryotes. The search for target RNAs and the roles of LSm comlexes have already established that LSm complexes play important roles in pre-mRNA splicing and mRNA decay. It has also been found that LSm complexes are important for positive strand RNA virus replication and translation. Through studies of BMV in yeast, it has been shown by Galão *et al*. that the LSm1-7 complex directly interacts with the BMV genome and acts as a regulator of translation and recruitment of BMV genome (Galão *et al*., 2010). The confirmed role of LSm1-7 complex in BMV/yeast studies compels to ask the question if LSm1-7 complex also play important roles in genome translation and replication of other positive strand viruses. Hence, the aim of this study is to investigate the role of LSm1-7 complex in the replication cycle of the positive strand *Flock house virus* (FHV). Due to its simple genome organization and its ability to replicate in yeast, FHV represents an ideal model system. With the availability of recombinant LSm1-7 complex and established *in vitro* binding assays we studied if LSm1-7 affects the FHV genomic/subgenomic ratio by acting on long distance pairing signals.

13. Materials and Methods

13.1 Materials

13.1.1 Chemicals:

All laboratory chemicals were obtained from Roth, Sigma Aldrich, Merck, Promega, Amersham Bioscience, Qiagen and Serva in p.a. quality.

13.1.2 Chromatographic Matrices

Ni-nitrilotriacetic acid beads (Ni-NTA)	QIAGEN, Hülsen
Protein G-Sepharose beads	Amersham, Freiburg

13.1.3 Markers

GeneRuler™ DNA Ladder Mix	Fermentas, St. Leon-Rot
GeneRuler™ 100 bp Plus DNA Ladder	Fermentas, St. Leon-Rot
Precision Plus Protein Dual Color Standards	Bio-Rad, München
Precision Plus Protein Unstained Standard	Bio-Rad, München

13.1.4 Kits and reagents

Amersham Rapid-hyb™ Buffer	GE Healthcare, München
Bio-Rad Protein Assay	Bio-Rad, München
NucleoSpin® Plasmid	Macherey-Nagel, Düren
NucleoSpin® Plasmid QuickPure	Macherey-Nagel, Düren
Ribonucleotidmix	Roche, Mannheim
Rotiphorese Gel 30	Roth, Karlsruhe
Rotiphorese Gel 40	Roth, Karlsruhe
Rotiphorese Gel A	Roth, Karlsruhe
Rotiphorese Gel B	Roth, Karlsruhe
Roti®-Phenol	Roth, Karlsruhe

13.1.5 Dialysis membranes and Filters

Dialysis membrane, MWCO 30kD	Roth, Karlsruhe

MF-Millipore Membranfilter	Sigma, München
Millipore Nitrocellulosemembran 0,22 µm GSWP	Millipore, Schwalbach
Minisart® Single use filter unit, 0,2 µm, steril	Sartorius Stedim, Göttingen
Slide-A-Lyzer® MINI Dialysis Units, 7.000 MWCO	Thermo Scientific, Rockford (USA)
VIVASPIN 6, 30.000 MWCO PES	Sartorius Stedim, Göttingen
Whatman 3 MM-Filterpaper	Hartenstein, Würzburg

13.1.6 Plastic ware

DC-Plasticfolie PEI Cellulose F	Merck, Darmstadt
Disposable cuvettesPS, 0.5µl, 1,6 mL	Hartenstein, Würzburg
Pipettes 2, 5, 10 and 20 mL	Hartenstein, Würzburg
PARAFILM®, M	Pechiney, Chicago (USA)
Pipette tips	Peske, Aindling-Pichl
CELLSTAR® Tubes, PP, 15 and 50 mL, steril	Greiner, Frickenhausen
Tubes 1, 1,5 and 2,0 mL	Eppendorf, Hamburg
UVette® 220-1600 nm, disposable single sealed	Eppendorf, Hamburg

13.1.7 X-ray films and screens

Amersham Hyperfilm™ MP	GE Healthcare, München
CEA RP NEW Medical X-ray Screen	Agfa Healthcare,

13.1.8 Buffers and Solutions

Agarose gel (0,8 %)	0,8 g Agarose
	100 mL 1x TBE
	10 µL Ethidiumbromid (1%)
Coomassie –Stain	35% Methanol
	5% Acetic acid
	0.2% Serva Blue
Coomassie-Destain	35 % Methanol

	7 %Acetic acid
Ion Exchange Buffer 100	20 mM Hepes pH 7,5
	100 mM NaCl
	5 mM DTT
	filtered and degassed
Ion Exchange Buffer 1000	20 mM Hepes pH 7.5
	1000 mM NaCl
	5 mM DTT
	filtered and degassed
Gel Filtration Buffer	20 mM Hepes pH 7,5
	200 mM NaCl
	5 mM DTT
	filtered and degassed
Ni- Wash Buffer	20 mM Hepes pH 7,5
	50 mM NaCl
	10 mM Imidazole
	5 mM β-ME
Ni-Elution Buffer	20 mM Hepes pH 7,5
	500 mM NaCl
	500 mM Imidazole
	5 mM β-ME

13.1.9 RNA-Polyacrylamid-Gels

Denaturing, 6 %	15 % (v/v) RotiphoreseGel40
	8 M Urea
	0,5x TBE
	450 µL 10% APS
	45 µL TEMED

Native, 6 %

20 % (v/v) Rotiphorese Gel A
3,25 % (v/v)RotiphoreseGelB
0,3x TBE
10 % (v/v) Glycerin 86 %
250 µL 10% APS/50mL
40 µL TEMED/50mL

Native, 7,5 %

25 % (v/v) Rotiphorese GelA
4,69 % (v/v)RotiphoreseGelв
0,5x TBE
10 % (v/v) Glycerin 86 %
250 µL/50 mL APS
30 µL/50 mL TEMED

13.1.10 Gel electrophoresis buffers

10x Laemmli Buffer

440 mM Tris-base
2 M Glycerin
1,5 % (w/v)   SDS
pH 8,3

4x SDS-loading gel buffer

0,5 M Tris-HCl
0,4 % (w/v) SDS
pH 6,8

4x SDS-Separating gel buffer

1,5 M Tris-HCl
0,4 % (w/v)   SDS
pH 8,8
10x TBE
890 mM Tris
890 mM H_3BO_4
20 mM Na_2EDTA

Denaturing RNA loading buffer	80% (v/v) Formamide
	0.1% Xylene Blue
	1.1% (w/v) Bromophenolblue
DNA loading buffer 6X	10mM Tris/HCl pH 8.0
	30% (v/v) Glycerine
	0.1% Bromophenol blue
	0.1% Xylene blue
Native RNA loading buffer 2X	20% Glycerine
	0.1% Xylene blue
	0.1% Bromophenol blue
	1X TBE
Protein Sample loading buffer	0.5 M tris/HCl, pH 6.8
	30% (v/v) 86% Glycerine
	1% (w/v) SDS
	1% (w/v) DTT
	0.1% Bromophenol Blue

13.1.11 Bacterial cells

E. coli BLR pREP4 SG13009
E. coli BLR pREP4
E. coli XL1 BLUE
E. coli DH5α

13.1.12 Medium for Bacterial culture

Ampicillin	Boehringer, Mannheim
Bacto-Agar	Clontech, Heidelberg
Bacto-Trypton	Clontech, Heidelberg
Bacto-Yeast-Extract	Clontech, Heidelberg
Chloramphenicol	Boehringer, Mannheim
1-Isopropyl-β-D-1-thiogalactopyranosid (IPTG)	Roche, Mannheim

Kanamycin	Boehringer, Mannheim
A-D(+)-Glucose Monohydrat	Roth, Karlsruhe
LB-Medium	0,5 %(w/v) Yeast extract
	1 % (w/v) Bacto-Tryptone
	1% (w/v) NaCl
Superbroth-Medium (SB)	3,5 % (w/v) Bacto-Tryptone
	2 % (w/v) Yeast extract
	0,5 % (w/v) NaCl

13.1.13 Instruments

Accument$^®$ pH Meter Model 15	Denver Instruments, NY (USA)
Avanti$^®$ J-HC Coolingcentrifuge	Beckman Coulter, Krefeld
Biofuge pico	Heraeus, Hanau
Biometra$^®$ Standard Power Pack P25	Analytik Jena, Jena
BioPhotometer	Eppendorf, Hamburg
Bio-Rad Mini PROTEAN$^®$ 3 Cell Gelsystem	Bio-Rad, München
Certomat$^®$ BS-1 Bacterialincubator/shuttler	Braun Biotech Melsungen
Chyo JL-180 Fine Balance	Hartenstein, Würzburg
Electrophoresis Power Supply EP S 601	Amersham, Freiburg
Heatingblock	Störktronic, Stuttgart
Horizontal gel electrophoresis system	Instituite of Biochemistry, Würzburg
IKA$^®$ Combimag RC0 Magnetic stirer	IKA, Staufen
IKA$^®$ Vibrax VXR basic	IKA, Staufen
Innova$^®$ 4300 Incubator Shaker Series	New Brunswick Scien., Nürtingen
Innova$^®$ 44 Incubator Shaker Series	New Brunswick Scien., Nürtingen
Jouan VX490S -80 °C Freezer	Schilling, Leipzig
Kühlzentrifuge 5415R	Eppendorf, Hamburg

Kühlzentrifuge 5804R	Eppendorf, Hamburg
Magnetic stirer	Gesellschaft für Labor-bedarf, Würzburg
Microwave	Philips, Böblingen
Pharmacia LKB ECPS 3000/150 Power Supply	Pharmacia, Berlin
Savant Speed Vac SC110	Thermo, Dreieich
Sonicato	Heinemann, Schwäbisch Gmünd
Specord 50 Photometer	Analytik Jena, Jena
Speed Vac Concentrator	Bachofer, Reutlingen
Thermomixer comfort	Eppendorf, Hamburg
Thermomixer compact	Eppendorf, Hamburg
TKA-LAB-HP Millopore System	TKA, Niederelbert
Vertical Gelelectrophoresisapparatus	Institute of Biochemistry, Würzburg
Vortex Genie 2	Scientific Industrie, NY (USA)
Vortexer	Hartenstein, Würzburg
Balance BL1500S	Sartorius Stedim, Göttingen
Balance PG8001-S	Mettler, Giessen
Centrifuge 5415D	Eppendorf, Hamburg
Centrifuge 5424	Eppendorf, Hamburg

13.1.14 Rotors

SW60Ti	Beckman Coulter, Krefeld
Typ45Ti	Beckman Coulter, Krefeld
Typ60Ti	Beckman Coulter, Krefeld
JS 4.2	Beckman Coulter, Krefeld

13.2 Methods

13.2.1 Plasmid Isolation from *E. coli* Cells:

Plasmid isolation from bacterial cultures was performed according to manufacture's protocols (Promega, Qiagen and Macherey-Nagel). This technique is based on the alkaline and SDS/Sodium dodecyl sulfate treatment (Birnboim and Doly, 1979). Plasmid-transformed *E. coli* cells were grown in 5 ml LB medium containing 50µg/µl appropriate antibiotics overnight at 37°C. The cells were pelleted by centrifugation at 14.000 rpm for 30 seconds. The addition of SDS and Sodium hydroxide led to lysis of bacterial cell wall and denaturation of chromosomal DNA, plasmid DNA and protein. After neutralization with potassium acetate, protein, chromosomal DNA as well as bacterial cell debris formed a insoluble complex in the presence of less soluble-potassium dodecyl sulfate. This complex was then precipitated under high salt concentration, while plasmid DNA remains in solution. The plasmid DNA was purified from other cellular components and chromosomal DNA on an anion exchange chromatography column. Following washing with ethanol-containing wash buffer, the plasmid DNA was eluted and precipitated using isopropanol and ethanol. The pellet was dissolved in sterile ddH2O and plasmid concentration was measured using spectrophotometer, the isolated plasmid was examined and analyzed on 1% (w/v) agarose gel.

13.2.2 Expression, and Purification of LSm1-7 Proteins and Subcomplexes

Expression constructs contains an MRGS H_6 tag at the N terminus of the first cistron, followed by a tobacco etch virus (TEV) cleavage site. The *E. coli* SG13009 (pREP4) for LSm2/3 and, BLR (pREP4) for LSm1, LSm4, LSm 5/6/7 were used for the expression. The *E. coli* cells were transformed with plasmid DNA and plated out on selective media. LB starter cultures were grown at 30 °C overnight, and 12 liters of SB media were inoculated the next day. Cultures were grown to an OD600 of 0.8 at 37 °C and induced with 1 mM isopropyl 1-thio-γ-D-galactopyranoside. Induction temperature was between 25 °C and 37 °C. Cells were harvested after 4 – 48 h of induction, depending on construct. Cell pellets were resuspended in lysis buffer containing 20 mM HEPES-Na, pH 7.5, 0.5–1.0 M NaCl, 10 mM imidazole-Cl, pH

7.50, 5 mM β-mercaptoethanol, and sonicated. Insoluble material was removed by ultracentrifugation, and supernatants were purified by immobilized metal ion affinity chromatography (IMAC) on nickel-charged Hi-Trap chelating Sepharose columns (Amersham Biosciences). LSm proteins and sub-complexes were eluted with imidazole. When insufficiently pure, samples were subsequently dialyzed into 100 mM NaCl buffer without imidazole and subjected to ion exchange chromatography (100 mM to 1 M NaCl). Samples were frozen in liquid nitrogen in ion exchange buffer

13.2.3 Reconstitution and Purification of LSm1–7 Complex

Individual LSm protein or sub-complex preparations were incubated in 4 M urea, 1 M NaCl buffer for 2 h at 37 °C and then mixed in equimolar amounts for the assembly of the complex. The mix was incubated again for 2–5 h, and the sample was dialyzed against buffer with progressively less salt (1 M and 0.5 M NaCl) overnight at 4 °C. Reconstituted LSm1–7 were purified by consecutive gel filtration and anion exchange chromatography.

13.2.4 Detection of Proteins on Gel

Visualization of electrophoretically separated proteins requires use of dyes or stains. Organic stain such as Coomassie blue has been adapted for protein detection on gels. In the Coomassie blue staining, an acidic medium is required for generation of an electrostatic attraction between dye molecules and amino groups of protein. This electrostatic attraction, together with van der Waals forces, binds the dye-protein complex together. Proteins are visualized as discrete blue bands. Gel was soaked immediately after electrophoresis in a Coomassie blue solution for 15 minutes and then washed with a large excess of a destaining solution. The destaining solution was changed several times, until the background has been considerably removed.

13.2.5 Quantification of Protein Concentration by Bradford method

Protein concentration was estimated using Bradford assay. This method takes advantage of the fact that the absorbance maximum of the dye in acidic solution will shift from 465 to 595 nm after adding protein. Protein sample was diluted in 800 µl ddH$_2$O, mixed with 200 µl Bradford reagents, incubated at RT for 5 min and absorbance of this mixture was measured at 595 nm. As a reference: instead of

protein sample to be measured, an appropriate buffer used in the protein sample was added into the ddH$_2$O-Bradford reagent mixture.

13.2.6 *In vitro* Transcription

The plasmid was linearized before use. The transcription reaction was assembeled as follows:

1µl   DNA template (~ 1µg of linearized plasmid DNA)

6µl   5x RNA pol buffer

1µl   NTPs (5mM ATP, CTP, GTP, 1mM UTP)

3µl ^{32}P-UTP (30µCi)

1µl   RNase Inhibitor (10U/µl)

1µl   T7 RNA Polymerase

After 2h at 37° C 1µL T7 RNA-Polymerse was added again and incubated for a further 2h at 37° C.

After separating the RNA on a 5% Polyacrylamide/8 M urea gel, the corresponding band was cut after exposure to X-ray film and then incubated in 400µl acetate-EDTA-SDS buffer (AES) at 4° C on the head over tail rotor. After centrifugation, the gel pieces were removed and the RNA eluted in AES buffer was precipitated by adding 1ml 100% EtOH and incubating at -20° C for 2h. After incubation the precipated RNA was collected as pellet after centrifugation at 4° C at 13 000 rpm for 30 min. The supernatant was discarded and the pellet was counter checked. The pellet was washed with 80% EtOH. After drying the pellet in speedvac for ~ 20 min. the pellet was resuspended in DEPC treated water and diluted to 100cps/µl (counts per second/µl). The diluted RNA samples were stored at -20° C.

13.2.7 Electro-mobility Shift assay

In a typical binding experiment, 300cpm (≈50fmol) of labeled transcript is incubated with 20pmol of LSm1-7 complexes in a buffer containing 20mM HEPES-NaOH pH 7.5, 200mM NaCl, 2mM MgCl$_2$, 0.1U/µl RNasin and 0.1µg/µl yeast tRNA in a 5µl assay at 30°C for 1h. A 2x loading buffer containing 20% glycerol was added to each samples and then total volume was loaded on previously pre-run 4% glycerol gels, and run at 4°C for 2h, 30mA. Gels were autoradiographed at -80°C on

maximum sensitivity Kodak Biomalx Films (Sigma-Aldrich, Munich, Germany). Depending on the RNA- LSm1-7 complex titration experiment two different of amounts complex are used.

14. Results

14.1 Purification of LSm proteins and LSm sub-complexes

Expression and purification of individual LSm proteins and subcomplexes was performed as described earlier by Zaric *et al*. LSm2-3 and 5-7 were coexpressed and purified as subcomplexes whereas LSm1 and 4 were expressed and purified as individual proteins (Zaric *et al*., 2005). The protein sequences followed by a His$_6$ tag and a TEV cleavage site were cloned into pQE30 and expressed in BLR (pREP4) *E. coli* cells. After sonication of cells expressing individual LSm proteins and LSm subcomplexes, a clear lysate was prepared by ultracentifugation. The cleared lysate was then subjected to immobilized metal ion affinity chromatography (IMAC) on nickel-charged Hi-Trap chelating Sepharose column (Amersham Biosciences). When insufficiently pure, samples were subsequently dialyzed in a buffer and subjected to ion exchange chromatography. LSm1 protein was purified using cation exchange chromatography where as LSm2-3, LSm4 and LSm5-7 were purified by anion exchange chromatography. The fractions containing most pure LSm proteins were pooled and analysed by SDS-PAGE (see Figure 7). Samples were concentrated upto 10 mg/ml and frozen in liquid nitrogen in ion exchange buffer. Bacterial culture of 12 litres yielded about 15-30 mg of purified proteins.
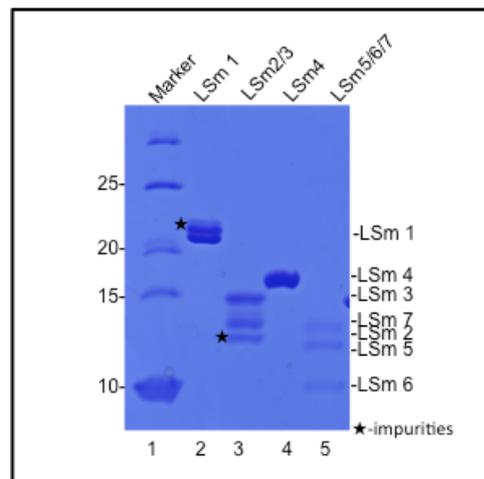
Figure 7: The coomassie stained SDS-polyacrylamide gel of purified LSm proteins and LSm subcomplexes after purification by ion exchange chromatography. Lane 2 shows purified LSm 1 protein; lane 3 LSm2-3 subcomplex; lane 4 and 5 LSm4 protein and LSm5-7 subcomplex respectively. The marker is shown in lane 1. The impurities are marked by asteriks.

14.2 Reconstitution of LSm1-7 Complex

The purified proteins were used to reconstitute LSm1-7 complex. 230 nmoles of each protein was used for reconstitution of the complex. After thawing the individual proteins and subcomplexes at 4°C the dissociation buffer containing 4M Urea was added and incubated at 37°C for 2 h. After mixing equimolar amounts of each subunit/subcomplexes the mixture was further incubated for 2 h at 37°C. Subsequently, urea was removed by dialysing against the buffer containing 20 mM Hepes-NaOH (pH-7.5), 1 M NaCl and 5 mM β-mercaptoethanol for 6h at 4°C. In a subsequent overnight dialysis, the salt concentration was lowered by dialysing against buffer conatining 20 mM Hepes-NaOH (pH-7.5), 500 mM NaCl and 5 mM β-mercaptoethanol. The dialysed reconstitution mixture was purified by Superdex 200 10/300 Gl gel filtration column (see Figure 8). The fractions corresponding to the main peak obtained during gel filtration chromatography were analyzed by SDS-Page (see Figure 9). The fraction containing all the seven subunits of LSm1-7 complex were pooled and subjected to anion exchange chromatography (see Figure 10). The main peak obtained was analyzed by SDS-Page (see Figure 11) and the fraction containing all the subunits of LSm1-7 complex were pooled and stored at 4°C. The final yield of the complex after all purification steps was approximately 6 mg.
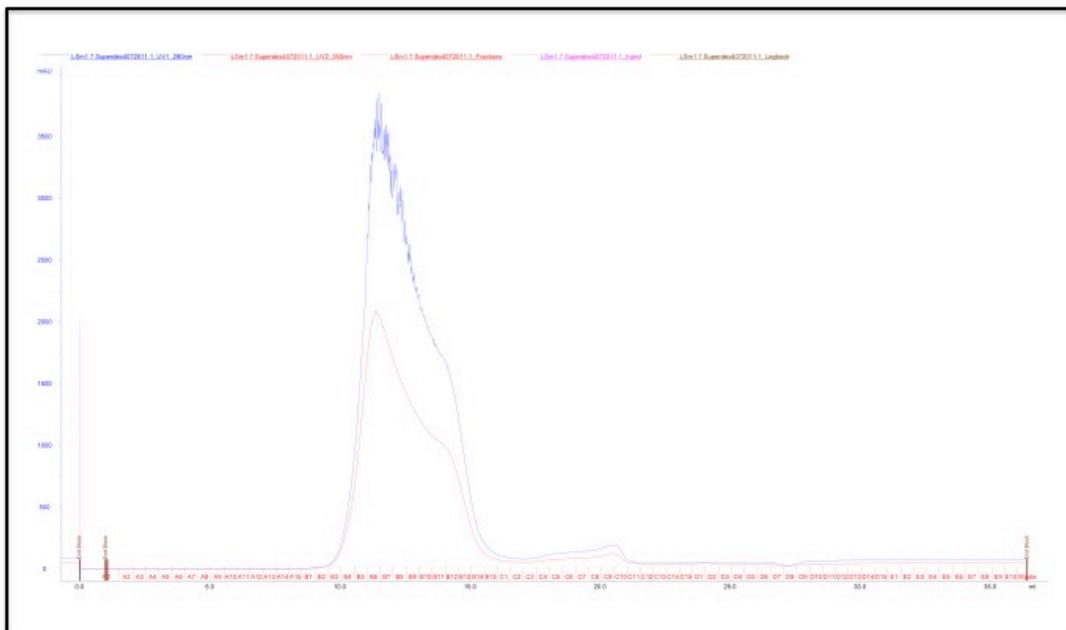
Figure 8: Chromatogram of gel filtration purification of LSm1-7 complex. The peak corresponds to the

fractions containing the reconstituted complex. The eluted fractions in the peak from fraction A15 to C3 were analyzed on SDS-Page (see below Figure 10).
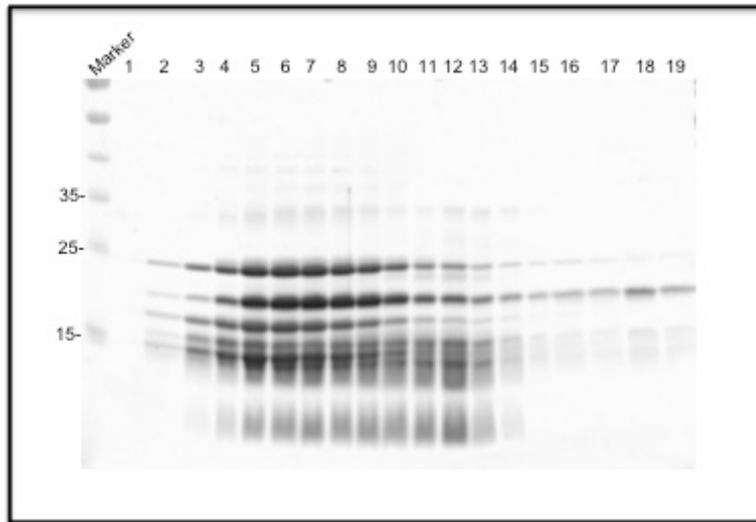
Figure 9: Coomassie stained SDS-Page gel of eluted fractions after gel filtration of the reconstituted LSm1-7 complex. Reconstitution mixture concentrated to a volume of 500 μl was injected into Superdex 200 10/300GL. The eluted fractions containing the peak were analyzed by SDS-Page. Lanes 5-11 shows all the seven subunits of LSm1-7 complex. The corresponding fractions were pooled and further subjected to anion exchange chromatographic purification (see below Figure 10).
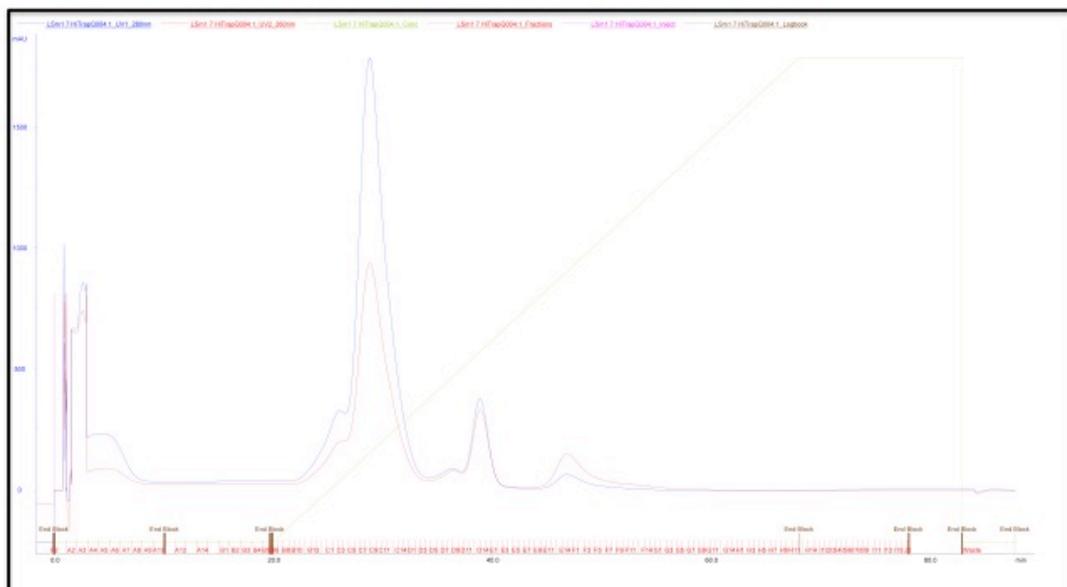
Figure 10: Chromatogram of anion exchange chromatography of LSm1-7. The purified fraction obtained after gel filtration chromatography were pooled and subjected to anion exchange chromatography. The highest peak corresponds to the reconstituted full LSm1-7 complex where as the smaller peaks corresponds to the un-incorporated or dissociated subunits/subcomplexes. Fractions C3-D3, D13, D14, E1, E14, F1, F2 corresponding to the peaks were analyzed by SDS-Page. (see below Figure 11)
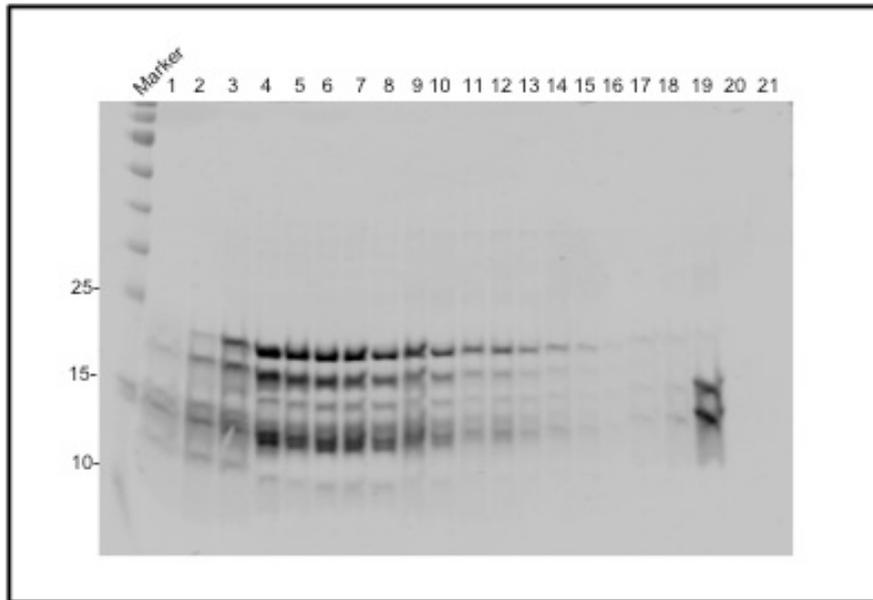
Figure 11: Coomassie stained SDS-Page gel of the fractions obtained after anion exchange purification. The fractions shown in lanes 4-9 show all the seven subunits of LSm1-7 complex. The fractions from lane 4-9 were pooled and stored at 4°C for further experiments. The marker is indicated to the left.

14.3 Preparation of labeled transcripts for electro-mobility shift assays

To map the binding sites of LSm1-7 complex in the FHV subgenome a set of 12 constructs was generated. The transcription of these constructs resulted in FHV RNA derivatives that cover different regions of RNA1, the FHV subgenome (see Figure 12). Sequences of 400-500 nucleotides were cloned into plasmids. Figure 12 shows a schematic represntation of the constructs. Constructs A-J span the different regions of RNA1. Constructs K and M were specially designed to mimic a long-distance interaction required between the regions DSCE and PSCE (through helix 1 and helix 2). Computer modeling by Lindenbach *et al*. has predicted that base pairing between DSCE and PSCE results in the formation of two helices, each consisting of approximately 6 bp important for RNA3 synthesis (Lindenbach *et al*., 2002). Constructs K and M were generated in order to bring these two regions together and mimic this interaction (construct K) or to disrupt it (construct M).

The cloned sequences were first linearized by the respective restriction

enzyme. The linearized templates were then transcribed using T7 polymerase and labeled by P^{32} uridine triphosphate (UTP). Transcription of all the constructs resulted in the transcripts of expected size (see Figure 14). After transcription the radiolabelled transcripts were stored at -20°C and diluted to the concentration of 50 fmol before using in electro-mobility shift assays.
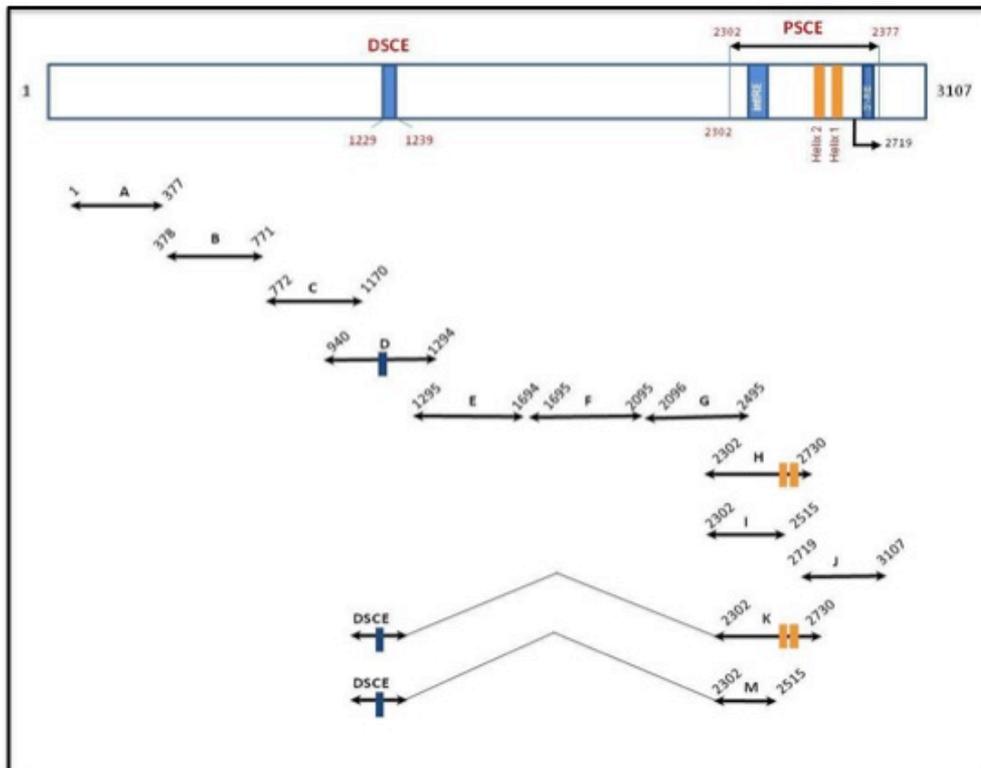
Figure 12: The figure represents the scheme of constructs prepared to perform electro-mobility shift assays. 12 constructs spanning RNA1 were designed to map the binding sites of LSm1-7 complex.
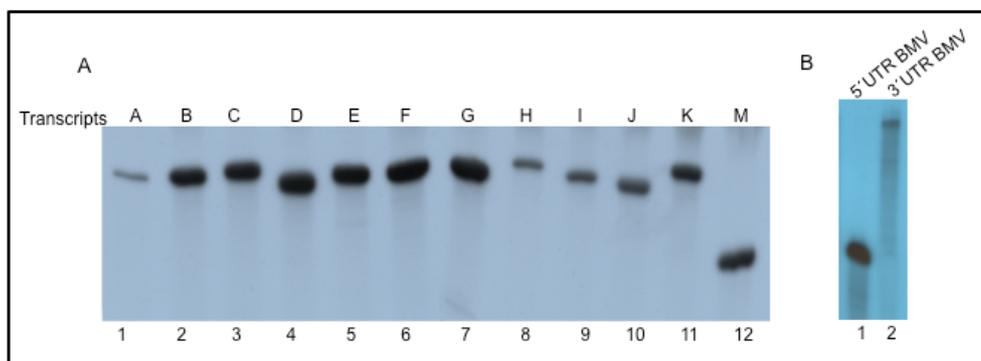
Figure 13: Synthesis of the radiolabelled transcripts. The constructs prepared were linearized with respective enzymes. Linearized templates were used for *in vitro* transcription by T7 polymerase. The transcripts were labeled using P^{32} uridine triphosphate (P^{32} UTP) during transcription. (A) Shows the transcripts from RNA1 of FHV genome and (B) shows the control transcripts from BMV.

14.4 Electro-mobility Shift assay:

To map the binding sites of LSm1-7 complex in the FHV subgenome, electro-mobility shift assays were performed. In a typical binding experiment, 50 fmol of labeled RNA were incubated with 60 and 120 pmol of reconstituted LSm complexes in a buffer containing 20 mM HEPES-NaOH pH 7.5, 200 mM NaCl, 2 mM MgCl$_2$, 0.1 U/µL RNasin, and 0.1 µg/µL yeast tRNA in a 5 µL assay at 30°C for 1 h. Samples were loaded on a prerun 5% native polyacrylamide gels, and run at 4°C for 2 h. Gels were autoradiographed at −80°C on maximum sensitivity Kodak Biomax Films. To assess if the recombinant LSm1-7 complex is functional a titration assay with the control transcripts, the radiolabeled transcripts from the 5′ UTR and 3′ UTR of BMV genome was performed. It has been already shown by Galão *et al* that LSm1-7 complex binds to the 3′ UTR of the BMV genome and not with the 5′ UTR. As expected the complex binds to the 3′ UTR and not to the 5′ UTR, under the same conditions (see Figure 15) (Galão *et al.*, 2010). The results show that the RNA-Protein complex is formed already when the concentration of LSm1-7 complex is 30 pmoles and most of the free RNA is bound when the concentration of the complex is 60 pmoles. Therefore 60 pmoles and 120 pmoles of the complex were choosen to perform the electro-mobility shift assays with FHV transcripts.
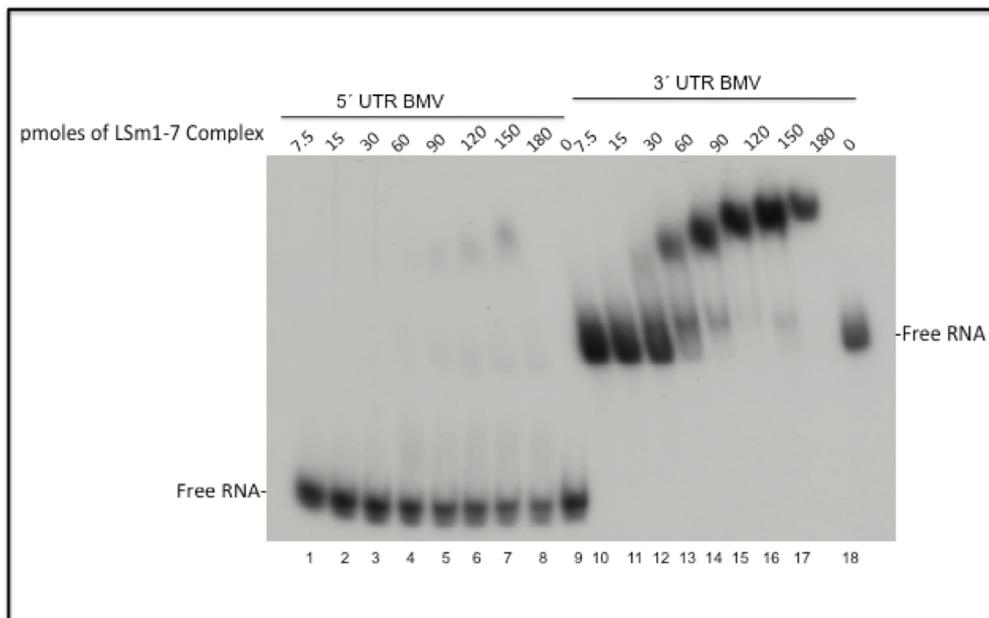
Figure 14: Titration of LSm1-7 complex with the negative (5′ UTR BMV) and positive (3′ UTR BMV) control. The lanes 9 and 18 show the respective free RNA. In lanes 1-8 and 10-17 the complex is added in increasing concentration as described.

The binding of the LSm1-7 complex to radiolabeled FHV RNA derivatives was performed using identical conditions as determined for the LSm1-7 complex-BMV transcripts (Figure 15). Interestingly, the shift assay shows a robust binding to all regions of RNA1. The binding of recombinant LSm1-7 complex to two distinct RNA-target sequences in the BMV genome, a tRNA-like structure at the 3′ UTR and two internal A-rich single-stranded regions has been shown by Galão *et al. In vivo* analysis shows that these sequences regulate the translation and replication of the BMV genome (Galão *et al.*, 2010). However, studies from Wynsberghe and Ahlquist show that 5′ proximal nt 1 to 376 of RNA1 of FHV contains *cis* activity for directing proteinA mediated RNA1 recruitment to mitochondria required for RNA1 replication. Considering the importance of 5′ end of the FHV RNA1, it is not surprising that LSm1-7 complex binds to the 5′ end of the RNA1 of FHV genome.
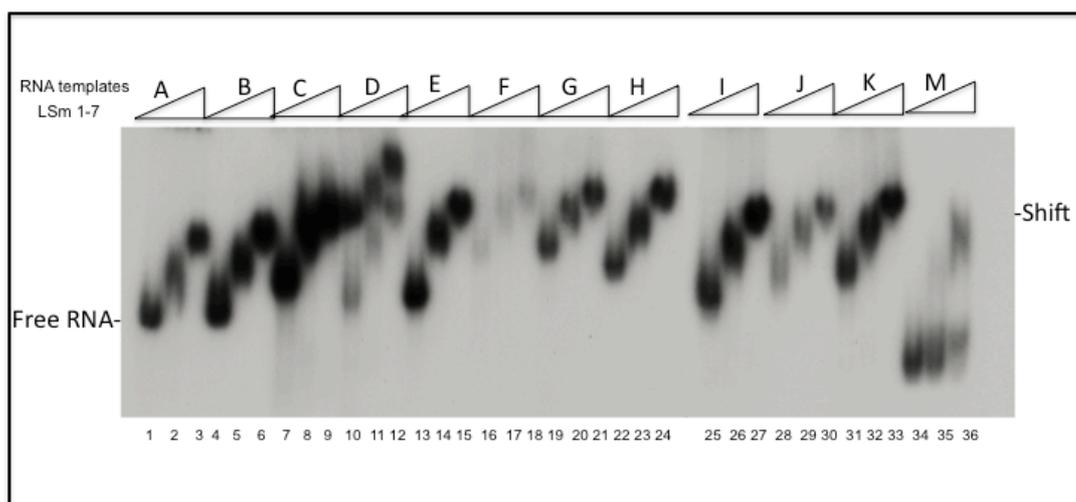
Figure 15: Band shift assay with the transcripts from RNA1 of FHV genome. Each template RNA was incubated with 60 pmoles and 120 pmoles of the complex. Lanes 1, 4, 7, 10, 13, 16, 19, 22, 25, 28, 31 and 34 show the free RNA. All the other lanes show the RNA-protein complex.

To confirm that the electro-mobility shift observed is due to the RNA-Protein complex formation and not due to the secondary structures of RNA, the electro-mobility shift assay was performed in the presence of Proteinase K (PK). To the reaction mix after forming RNA-protein complex PK was added to the final concentration of 100 U/ml and further incubated for 30 min (see Figure 16). The results show that the shift is indeed due to specific binding of the RNA sequences to the LSm1-7 complex. Upon addition of PK, the RNA-protein complex is disrupted and

the RNA migrates like the free RNA. This confirms that LSm1-7 binds to all the regions of RNA1 of FHV genome.
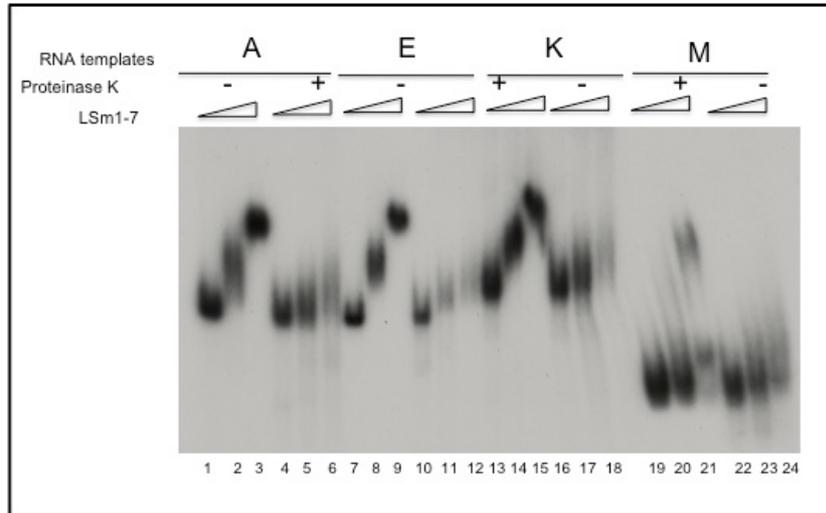

Figure 16: Proteinase K digestion upon formation of RNA-protein complexes. Lanes 1, 2, 3, 7, 8, 9, 13, 14, 15, 19, 20 and 21 show the band shift assay with transcripts as controls and in lanes 4, 5, 6, 10, 11, 12, 16, 17, 18, 22, 23 and 24 show the RNA-protein complex subjected to PK digestion. As a control the free RNAs in lanes 4, 10,16 and 22 also subjected to PK digestion.

15. Discussion

The Sm and LSm proteins belong to a conserved family of proteins that form a complex of hexameric or heptameric ring like structures, which bind to cellular RNAs. In eukaryotes, the LSm proteins form five different complexes containing various LSm proteins. LSm1-7 present in the cytoplasm is involved in RNA metabolism. LSm1-7 complex has been the focus of studies lately since it plays an important role in mRNA degradation by interacting with both the mRNA and mRNA degradation machinery. Apart from their reported role in the degradation of cellular mRNA, the LSm1-7 complex also functions as a host factor for the translation and replication of positive strand RNA viruses.

Viruses require the cellular machinery to multiply and express their genomes and regulate the viral life cycle, Galão *et al*. have shown that LSm1-7 directly interacts with BMV RNAs via two specific RNA-targeting elements, a tRNA-like structure located in the 3′ UTR and two internal A-rich single-stranded regions. They have shown by *in vivo* analysis that the LSm1-7 RNA-target sequences act as regulators of translation and recruitment of the BMV genome (Galão *et al*., 2010). Based on the observation that the LSm1-7 complex plays an important role in the translation and replication of BMV genome this study aimed to investigate if there are similarities in LSm1-7 function in other viral model systems. The simple genome organization and the property to replicate in yeast and *Drosophila* cells has made FHV a valuable model system. The aim of this study was to map the binding sites of LSm1-7 complex in RNA1 of FHV genome by electro-mobility shift assays. An unpublished study from Prof. Diez′s laboratory in Barcelona (Personal communication) shows that deletion of LSm1 gene dramatically increases FHV RNA3 accumulation resulting in altered ratio of genomic/subgenomic RNAs. They have tested that this effect is not related to the differences in stability of the RNA or to the expression of viral proteins. Their data supports that LSm1-7 plays an important role in the regulation of FHV genomic/subgenomic ratio. Based on these observations, one possibility is that LSm1-7 affects the FHV genomic/subgenomic ratio by acting on long distance pairing signals.

The results show that LSm1-7 binds to all the regions of RNA1 targeted in this study. Unlike the results from the study in BMV genome where the binding of LSm1-7

complex is specific to the 3′ UTR and internal A-rich sequences, the results from the electro-mobility shift assays in FHV subgenome show that LSm1-7 most likely covers the whole RNA1-sequence. Electro-mobility shift assays in the presence PK confirm that binding of LSm1-7 complex to the target regions of RNA1 of FHV is not an artifact but due to formation of RNA-protein complexes. Binding of the LSm1-7 complex to the 5′ end of RNA1 is in agreement to the study by Wynsberghe and Ahlquist, where they have shown that the 5′ proximal region of RNA1 of FHV contains *cis* activity required for RNA1 recruitment for replication in mitochondria. They studied RNA1 subgenome by disecting its sequence into five regions (as shown in Figure 17). Authors investigated if there is any effect on translation, recruitment and replication of RNA1 in yeast and *Drosophila* cells upon deleting each of the five regions. They found that deleting either of the 1, 4 or 5 region severely inhibited negative strand RNA1 accumulation in both yeast and *Drosophila* cells. Deletion of region 3 had no effects in either cell types where as deletion of region 2 showed cell type specific effects. Yeast cells showed no effect on the RNA1 replication where as in *Drosophila* cells deletion of region 2 strikingly inhibited both the positive and negative strand RNA1 accumulation. The authors state that *Drosophila*-specific effect of region 2 was anticipated as this region affects the production of FHV subgenomic RNA3 which inturn encodes RNA silencing inhibitor protein B2. B2 protein itself is required for normal FHV RNA accumulation in *Drosophila*. Considering the importance of regions 1, 2, 4 and 5 (Wynsberghe and Ahlquist, 2009), its is not surprising that LSm1-7 binds to the transcripts arising from constructs A-E, G-K and M (Figure 12) since the sequences of these constructs lie in the regions 1, 2, 4 and 5. Further experiments are required to elucidate the importance of transcript from construct F. One of the possible explaination why recombinant LSm1-7 complex also binds to transcript from construct F, which lies in the region 3 for which Wynsberghe and Ahlquist show no effect on RNA1 translation and replication, can be due to the altered secondary structure adopted by RNA *in vitro.* Alteration in the secondary structure of RNA might be responsible for the LSm1-7 binding*.*
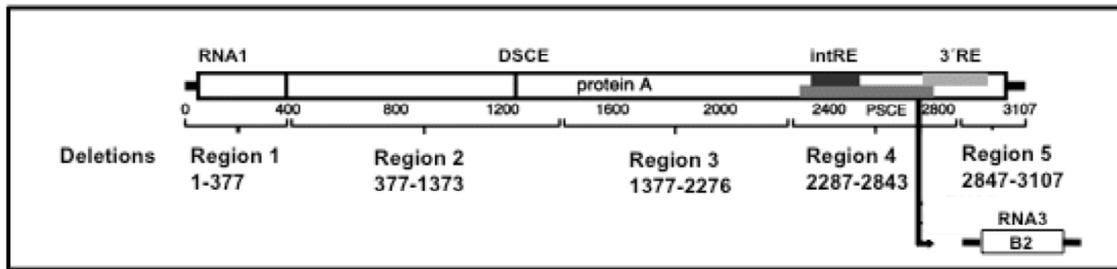
Figure 17: Scheme of known RNA1 and RNA 3´ cis elements present within RNA1. The distal subgenomic control element (DSCE) and the proximal subgenomic control element (PSCE) important for RNA3 synthesis, the internal element (intRE and the 3´ replication element (3´ RE) required for RNA1 replication are marked. Also marked are five regions (region 1-5) used by Wynsberghe and Ahlquist for deletion studies. (Modified from Wynsberghe and Ahlquist, 2009).

16. References

Achsel T, Brahms H, Kastner B, Bachi A, Wilm M and Lührmann R (1999). A doughnut-shaped heteromer of human Sm-like proteins150 binds to the 3'-end of U6 snRNA, thereby facilitating U4/U6 duplex formation in vitro. *EMBO J* 18: 5789-5802.

Ahlquist P (1992). Bromovirus RNA replication and transcription. *Curr. Opin. Genet. Dev* 2: 71-76.

Ahlquist P (2006). Parallels among positive-strand RNA viruses, reverse-transcribing viruses and double-stranded RNA viruses. *Nat. Rev. Microbiol* 4: 371-382.

Ahlquist P, Noueiry AO, Lee W, Kushner DB and Dye BT (2003). Host factors153 in positive-strand RNA virus genome replication. *J. Virol* 77: 8181-8186.

Alves-Rodrigues I, Galão RP, Meyerhans A and Díez J (2006). Saccharomyces cerevisiae: a useful model host to study fundamental biology of viral replication. *Virus Res* 120: 49-56.

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

Ball LA and Johnson KL (1998). Nodaviruses of insects. *The Insect Viruses*, pp. 225–267.

Beggs, J.D (2005). Lsm proteins and RNA processing. *Biochem. Soc. Trans* 33: 433-438.

Bergman N, Moraes KC, Anderson JR, Zaric B, Kambach C, Schneider RJ, Wilusz CJ and Wilusz J (2007). Lsm proteins bind and stabilize RNAs containing 5' poly(A) tracts. *Nat. Struct. Mol. Biol* 14: 824-831.

Birnboim HC, Doly J (1979). A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic Acids Res*. 24:1513-1523.

Coller JM, Tucker M, Sheth U, Valencia-Sanchez MA and Parker R (2001). The DEAD box helicase, Dhh1p, functions in mRNA decapping and interacts with both the decapping and deadenylase complexes. *RNA* 7: 1717-1727.

Conti E & Izaurralde E (2005). Nonsense-mediated mRNA decay: molecular insights and mechanistic variations across species. *Curr. Opin. Cell Biol* 17: 316-325.

Cordin O, Banroques J, Tanner NK and Linder P (2006). The DEAD-box protein family of RNA helicases. *Gene* 367: 17-37.

Díez J, Ishikawa M, Kaido M and Ahlquist P (2000). Identification and characterization of a host protein required for efficient template selection in viral RNA replication. *Proc. Natl. Acad. Sci. U.S.A* 97: 3913-3918.

Eckerle LD and Ball LA (2002). Replication of the RNA segments of a bipartite viral genome is coordinated by a transactivating subgenomic RNA. *Virology* 296: 165–176.

Fernandez CF, Pannone BK, Chen X, Fuchs G and Wolin SL (2004). An Lsm2-Lsm7 complex in Saccharomyces cerevisiae associates with the small nucleolar RNA snR5. *Mol Biol Cell*. 15: 2842-2852.

French, R. & Ahlquist, P (1987). Intercistronic as well as terminal sequences are required for efficient amplification of brome mosaic virus RNA3. *J. Virol* 61: 1457-1465.

Friesen PD and Rueckert RR (1982) Black beetle virus: Messenger for protein B is a subgenomic viral RNA. *J. Virol*. 42, 986–995.

Frischmeyer PA, van Hoof A, O'Donnell K, Guerrerio AL, Parker R and Dietz HC (2002). An mRNA surveillance mechanism that eliminates transcripts lacking termination codons. *Science*. 295: 2258-2261.

Galão RP, Chari A, Alves-Rodrigues I, Lobão D, Mas A, Kambach C, Fischer U, Díez J (2010). LSm1-7 complexes bind to specific sites in viral RNA genomes and regulate their translation and replication. *RNA*. 16: 817-827.

Galão RP, Scheller N, Alves-Rodrigues I, Breinig T, Meyerhans A, Díez J (2007). Saccharomyces cerevisiae: a versatile eukaryotic system in virology. *Microb Cell Fact*. 6:32.

Gallagher TM, Friesen PD and Rueckert RR (1983). Autonomous replication and expression of RNA 1 from black beetle virus. *J. Virol*. 46, 481–489.

Garneau NL, Wilusz J and Wilusz CJ (2007). The highways and byways of mRNA decay. *Nat. Rev. Mol. Cell Biol* 8: 113-126.

He W and Parker R (2001). The yeast cytoplasmic LsmI/Pat1p complex protects mRNA 3' termini from partial degradation. *Genetics* 158: 1445-1455.

He W and Parker R (2001). The yeast cytoplasmic LsmI/Pat1p complex protects mRNA 3′ termini from partial degradation. *Genetics* 158: 1445–1455.

Inada T and Aiba H (2005). Translation of aberrant mRNAs lacking a termination codon or with a shortened 3'-UTR is repressed after initiation in yeast. *EMBO J* 24: 1584-1595.

Janda M & Ahlquist P (1993). RNA-dependent replication, transcription, and persistence of brome mosaic virus RNA replicons in S. cerevisiae. *Cell* 72: 961-970.

Janda M and Ahlquist P (1998). Brome mosaic virus RNA replication protein 1a dramatically increases in vivo stability but not translation of viral genomic RNA3. *Proc. Natl. Acad. Sci. U.S.A* 95, 2227-2232.

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

Khusial P, Plaag R and Zieve GW (2005). LSm proteins form heptameric rings that bind to RNA via repeating motifs.*Trends Biochem Sci*. 30:522-528.

Kufel J, Bousquet-Antonelli C, Beggs JD and Tollervey D (2004). Nuclear premRNA decapping and 5' degradation in yeast require the Lsm2-8p complex. *Mol.*

Cell. Biol. 24: 9646-9657.

Lee T and Feig AL (2008). The RNA binding protein Hfq interacts specifically with tRNAs. *RNA* 14: 514-523.

Lindenbach BD, Sgro JY and Ahlquist P (2002). Long distance base pairing in flock house virus RNA1 regulates subgenomic RNA3 synthesis and RNA2 replication. *J. Virol*. 76: 3905–3919.

Mas A, Alves-Rodrigues I, Noueiry A, Ahlquist P and Díez J (2006). Host deadenylation-dependent mRNA decapping factors are required for a key step in brome mosaic virus RNA replication. *J. Virol* 80: 246-251.

Mayes AE, Verdone L, Legrain P and Beggs JD (1999). Characterization of Smlike proteins in yeast and their association with U6 snRNA. *EMBO J* 18: 4321-4331.

Moll I, Leitsch D, Steinhauser T and Bläsi U (2003). RNA chaperone activity of the Sm-like Hfq protein. *EMBO Rep* 4, 284-289.

Noueiry AO and Ahlquist P (2003). Brome mosaic virus RNA replication: revealing the role of the host in RNA virus replication. *Annu Rev Phytopathol* 41: 77-98.

Noueiry AO, Diez J, Falk SP, Chen J and Ahlquist P (2003). Yeast Lsm1p-7p/Pat1p deadenylation-dependent mRNA-decapping factors are required for brome mosaic virus genomic RNA translation. *Mol. Cell. Biol* 23: 4094-4106.

Pillai RS, Will CL, Lührmann R, Schümperli D and Müller B (2001). Purified U7 snRNPs lack the Sm proteins D1 and D2 but contain Lsm10, a new 14 kDa Sm D1-like protein. *EMBO J*. 20: 5470-5479.

Pruijn GJ (2005). Doughnuts dealing with RNA. *Nat Struct Mol Biol*. 12: 575-581.

Raker VA, Plessel G and Lührmann R (1996). The snRNP core assembly pathway: identification of stable core protein heteromeric complexes and an snRNP subcore particle in vitro. *EMBO J*. 15: 2256-2269.

Rice AP and Roberts BE (1983). Vaccinia virus induces cellular mRNA degradation. *J. Virol.* 47: 529–539.

Sauter C, Basquin J and Suck D (2003). Sm-like proteins in Eubacteria: the crystal structure of the Hfq protein from Escherichia coli. *Nucleic Acids Res.* 14:4091-4098.

Schneemann A, Reddy V and Johnson JE (1998). The structure and function of nodavirus particles: a paradigm for understanding chemical biology. *Adv Virus Res*. 50: 381-446.

Schuppli D, Miranda G, Tsui HC, Winkler ME, Sogo JM and Weber H (1997). Altered 3'-terminal RNA structure in phage Qbeta adapted to host factor-less Escherichia coli. *Proc. Natl. Acad. Sci. U.S.A* 94: 10239-10242.

Schwartz M, Chen J, Janda M, Sullivan M, den Boon J and Ahlquist P (2002). A positive-strand RNA virus replication complex parallels form and function of retrovirus capsids. *Mol. Cell* 9: 505-514.

Schwartz M, Chen J, Lee W, Janda M and Ahlquist P (2004). Alternate, virus induced membrane rearrangements support positive-strand RNA virus genome replication. *Proc. Natl. Acad. Sci. U.S.A* 101: 11263-11268.

Scotti PD, Dearing S and Mossop DW (1983). Flock House virus: a nodavirus isolated from Costelytra zealandica (White) (Coleoptera: Scarabaeidae). *Arch Virol*. 75:181-189.

Sullivan ML and Ahlquist P (1999). A brome mosaic virus intergenic RNA3 replication signal functions with viral replication protein 1a to dramatically stabilize RNA in vivo. *J. Virol* 73: 2622-2632.

Swaminathan S (2005). Post-transcriptional gene regulation by gamma herpesviruses. *J Cell Biochem*. 4:698-711.

Tharun S and Parker R (2001). Targeting an mRNA for decapping: displacement of translation factors and association of the Lsm1p-7p complex on deadenylated yeast mRNAs. *Mol. Cell* 8: 1075-1083.

Tharun S, He W, Mayes AE, Lennertz P, Beggs JD and Parker R (2000). Yeast Sm-like proteins function in mRNA decapping and decay. *Nature* 404: 515-518.

Tharun S, Muhlrad D, Chowdhury A and Parker R (2005). Mutations in the Saccharomyces cerevisiae LSM1 gene that affect mRNA decapping and 3' end protection. *Genetics* 170: 33-46.

Tomasevic N and Peculis BA (2002). Xenopus LSm proteins bind U8 snoRNA via an internal evolutionarily conserved octamer sequence. *Mol Cell Biol*. 22: 4101-4112.

Tomasevic N and Peculis BA (2002). Xenopus LSm proteins bind U8 snoRNA via an internal evolutionarily conserved octamer sequence. *Mol. Cell. Biol*. 22: 4101–4112.

Van Wynsberghe PM and Ahlquist P (2009). 5' cis elements direct nodavirus RNA1 recruitment to mitochondrial sites of replication complex formation. *J Virol*. 83: 2976-2988.

Venter PA and Schneemann A (2008). Recent insights into the biology and biomedical applications of Flock House virus. *Cell Mol Life Sci*. 65:2675-87.

Venter PA, Krishna NK and Schneemann A (2005). Capsid protein synthesis from replicating RNA directs specific packaging of the genome of a multipartite, positive-strand RNA virus. *J Virol*. 79: 6239-6248.

Vogel J and Luisi BF (2011). Hfq and its constellation of RNA. *Nat Rev Microbiol*. 9:578-589.

Wilusz CJ and Wilusz J. Eukaryotic Lsm proteins: lessons from bacteria (2005). *Nat Struct Mol Biol*. 12:1031-1036.

Zaric B, Chami M, Rémigy H, Engel A, Ballmer-Hofer K, Winkler FK and Kambach C (2005). Reconstitution of two recombinant LSm protein complexes reveals aspects of their architecture, assembly, and function. *J. Biol. Chem* 280:

16066-16075.

Zhong W and Rueckert RR (1993). Flock house virus: Down-regulation of subgenomic RNA3 synthesis does not involve coat protein and is targeted to synthesis of its positive strand. *J. Virol*. 67, 2716–2722.

Zúñiga S, Sola I, Cruz JLG and Enjuanes L (2009). Role of RNA chaperones in virus replication. *Virus Res* 139: 253-266.

17. List of Abbreviations

BMV– *Bromo Mosiac Virus*

BR– Balbiani Ring

Brg1– Brahma related gene1

Brm– Brahma

CBC– Cap binding complex

CBP– Cap binding protein

CPSF– Cleavage and polyadenylation specificity factor

CstF– Cleavage stimulating factor

CTD– Carboxy-terminal domain of large subunit of RNAPII

DNA– Deoxyribonucleic acid

DSCE– Distal Subgenome Control Element

EM– Electron microscopy

ESE– Exonic splicing enhancers

ESS– Exonic splicing silencers

EST – Expressed Sequence Tags

FHV– *Flock House Virus*

GT– Guanylyltransferase

hnRNPs– heterogenous nuclear ribonucleoproteins

IEM– Immuno-electron microscopy

IF– Immuno-fluorescence

ISS– Intronic splicing silencers

ISWI– Imitation switch

LSD1– Lysine specific demethylase 1

Lsm– Like-Sm proteins

mRNA– Messenger RNA

MT– Methyltransferase

pre-mRNA precursor-messenger RNA

PSCE– Proximal Subgenome Control Element

RdRp– RNA-dependent RNA polymerase

RNAPII– RNA polymerase II

RNP– Ribonucleoproteins

rRNA – ribosomal RNA

rRNA– ribosomal RNA

snoRNA– small nucleolar RNA

snRNA– small nuclear RNAs

snRNPs– small nuclear ribonucleoprotein particles

SR proteins– Serine/Arginine rich proteins

SWI/SNF– SWItch/Sucrose NonFermentable

tRNA– transfer RNA

UTR– Untranslated region

18. Acknowledgements

19. Affidavit

I hereby confirm that my thesis entitled **Role of SWI/SNF in regulation of pre-mRNA processing in *Drosophila melanogaster*; Funktion von SWI/SNF in der Regulation der prä-mRNA-Prozessierung in *Drosophila melanogaster*** is the result of my own work. I did not receive any help or support from commercial consultants. All sources and/or materials applied are listed and specified in the thesis.

Furthermore, I confirm that this thesis has not been submitted as part of another examination process neither in identical nor in similar forms.

Würzburg

Date   Signature

Eidesstattliche Erklärung

Hiermit erkläre ich an Eides statt, die Dissertation **Role of SWI/SNF in regulation of pre-mRNA processing in *Drosophila melanogaster*; Funktion von SWI/SNF in der Regulation der prä-mRNA-Prozessierung in *Drosophila melanogaster*** eigenständig, d.h. ins besondere selbständig und ohne Hilfe eines kommerziellen Promotionsberaters, angefertigt und keine anderen als die von mir angegebenen Quellen und Hilfsmittel verwendet zu haben.

Ich erkläre außerdem, dass die Dissertation weder in gleicher noch in ähnlicher Form bereits in einem anderen Prüfungsverfahren vorgelegen hat.

Würzburg

Datum   Unterschrift