Isolation and functional characterization of Hrp65-binding proteins in *Chironomus tentans*

Eva Kiesler

Department of Molecular Biology and Functional Genomics
Stockholm University, Sweden

Stockholm 2004
The cover picture shows an immunolabeling experiment in a salivary gland cell of *Chironomus tentans*, demonstrating that the RNA-binding protein Hrp59 is present at a subset of active transcription sites. Bands on the polytene chromosomes that correspond to sites of active transcription have been visualized by incorporation of BrUTP followed by staining with an antibody against BrUTP (red). The cell has also been stained with an antibody against Hrp59 (green).
To my parents
Abstract

It is well-established that the organization of nuclear components influences gene expression processes, yet little is known about the mechanisms that contribute to the spatial co-ordination of nuclear activities. The salivary gland cells of *Chironomus tentans* provide a suitable model system for studying gene expression *in situ*, as they allow for direct visualization of the synthesis, processing and export of a specific protein-coding transcript, the Balbiani ring (BR) pre-mRNA, in a nuclear environment in which chromatin and non-chromatin structures can easily be distinguished. The RNA-binding protein Hrp65 has been identified in this model system as a protein associated with non-chromatin nucleoplasmic fibers, referred to as connecting fibers (CFs). The CFs associate with BR RNP particles in the nucleoplasm, suggesting that Hrp65 is involved in mRNA biogenesis at the post-transcriptional level. However, the function of Hrp65 is not known, nor is the function or the composition of CFs. In the work described in this thesis, we have identified by yeast two-hybrid screening and characterized different proteins that bind to Hrp65. These proteins include a novel hnRNP protein in *C. tentans* named Hrp59, various isoforms of Hrp65, the splicing and mRNA export factor HEL/UAP56, and a RING-domain protein of unknown function. Immuno-electron microscopy experiments showed that Hrp59 and HEL are present in CFs, and in larger structures in the nucleoplasm of *C. tentans* salivary gland cells.

Hrp59 is a *C. tentans* homologue of human hnRNP M, and it associates cotranscriptionally with a subset of pre-mRNAs, including its own transcript, in a manner that does not depend quantitatively on the amount of synthesized RNA. Hrp59 accompanies the BR pre-mRNA from the gene to the nuclear envelope, and is released from the BR mRNA at the nuclear pore complex. We have identified the preferred RNA targets of Hrp59 in *Drosophila* cells, and we have shown that Hrp59 binds preferentially to exonic splicing enhancer sequences.

Hrp65 self-associates through an evolutionarily conserved domain that can also mediate heterodimerization of Hrp65 homologues. Different isoforms of Hrp65 interact with each other in all possible combinations, and Hrp65 can oligomerize into complexes of at least six molecules. The interaction between different Hrp65 isoforms is crucial for their intracellular localization, and we have discovered a mechanism by which Hrp65-2 is imported into the nucleus through binding to Hrp65-1.

Hrp65 binds to HEL/UAP56 in *C. tentans* cells. We have analyzed the distribution of the two proteins on polytene chromosomes and in the nucleoplasm of salivary gland cells, and our results suggest that Hrp65 and HEL become associated during post-transcriptional gene expression events. HEL binds to the BR pre-mRNP cotranscriptionally, and incorporation of HEL into the pre-mRNP does not depend on the location of introns along the BR pre-mRNA. HEL accompanies the BR mRNP to the nuclear pore and is released from the BR mRNP during translocation into the cytoplasm.
# Table of contents

List of articles included in the thesis .......................................................... 8  
Commonly used abbreviations ........................................................................ 9

**GENERAL INTRODUCTION ........................................................................... 11**

Organization of the eukaryotic cell nucleus ................................................. 11  
The nuclear envelope ..................................................................................... 12  
  *The nuclear pore complex* ......................................................................... 13  
Chromatin ........................................................................................................ 14  
The nucleolus ................................................................................................... 15  
Other subnuclear structures .......................................................................... 16  
  *Splicing factor compartments/speckles* ..................................................... 16  
  *Cajal bodies* .............................................................................................. 17  
  *PML bodies* .............................................................................................. 17  
  *Nuclear stress bodies* ............................................................................... 18  
  *Paraspeckles* ............................................................................................ 18  
  *The nuclear matrix* .................................................................................. 19

Expression of protein-coding genes in eukaryotic cells ............................ 20

Synthesis and processing of pre-mRNA transcripts .................................. 21  
  *RNA polymerase II transcription* ............................................................ 21  
  *Pre-mRNA processing* ............................................................................. 22  
Export of mRNA to the cytoplasm ............................................................... 23  
  *The mRNA export receptor NXF1/Mex67* ............................................. 24  
  *Adapters for NXF1/Mex67-mediated mRNA export* ............................. 24  
  *Other mRNA export factors* ................................................................... 24  
  *Directionality of mRNA transport* .......................................................... 26  
Coupling of mRNA export to pre-mRNA synthesis and processing ............. 27  
  *Splicing and mRNA export* ................................................................. 27  
  *Transcription and mRNA export* .......................................................... 28  
  *Pre-mRNA surveillance and nuclear retention* ....................................... 28  
Intranuclear movement and localization of (pre-) mRNPs ............................ 29  
  *Dynamics of intranuclear (pre-) mRNPs* ............................................... 29  
  *(Pre-) mRNPs and nuclear structures* ..................................................... 31
An introduction to the model system and to the present investigation...... 32

The Balbiani ring genes of Chironomus tentans: a model system for in situ studies of gene expression........................................................................................................... 32

The active BR genes........................................................................................................... 32
Biogenesis of BR pre-mRNAs.............................................................................................. 33
Proteins associated with BR pre-mRNAs ........................................................................... 34

BR particles associate with nucleoplasmic structures in salivary gland cells of C. tentans.......................................................................................................................... 36
Composition of CFs and FGCs............................................................................................ 36
The C. tentans Hrp65 protein .............................................................................................. 37
Proteins related to Hrp65..................................................................................................... 40
The Drosophila NonA/Bj6 protein......................................................................................... 40
The mammalian DBHS proteins......................................................................................... 41

PRESENT INVESTIGATION......................................................................................... 44

Aim and experimental approach......................................................................................... 44

Results and discussion ...................................................................................................... 44

Studies on Hrp59 (Paper I) .............................................................................................. 45
Studies on Hrp65 self-interaction (Paper II) ...................................................................... 48
Studies on HEL/UAP56 (Papers III and IV) ..................................................................... 50
Studies on Y9 .................................................................................................................... 52
Ongoing projects and future perspectives ........................................................................ 56
Investigation of CFs and FGCs.......................................................................................... 56
Functional studies on Hrp65/NonA and Hrp59 ................................................................. 57

Concluding remarks ......................................................................................................... 58

Acknowledgements .......................................................................................................... 59

References .......................................................................................................................... 60
List of articles included in the thesis

This thesis is based on the following articles, which will be referred to with their Roman numerals.


   Hrp59, an hnRNP M-like protein in *Chironomus* and *Drosophila*,
   binds to exonic splicing enhancers in a subset of pre-mRNAs
   *Submitted*


    The Hrp65 self-interaction is mediated by an evolutionarily con-
    served domain and is required for nuclear import of Hrp65 iso-
    forms that lack a nuclear localization signal


    HEL/UAP56 Binds to Hrp65 in *Chironomus tentans* and is Present
    in Nucleoplasmic Fibers that Interact Post-Transcriptionally with
    mRNP Particles
    *Manuscript*


    HEL/UAP56 binds cotranscriptionally to the Balbiani ring pre-
    mRNA in an intron-independent manner and accompanies the BR
    mRNP to the nuclear pore

The articles were reproduced with permission from the copyright holders.
Commonly used abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>AD</td>
<td>(trans-) activation domain</td>
</tr>
<tr>
<td>BR</td>
<td>Balbiani ring</td>
</tr>
<tr>
<td>Br-UTP</td>
<td>bromo-deoxyuridine triphosphate</td>
</tr>
<tr>
<td>CB</td>
<td>Cajal body</td>
</tr>
<tr>
<td>CF</td>
<td>connecting fiber</td>
</tr>
<tr>
<td>CTD</td>
<td>carboxyterminal domain (of RNA pol II)</td>
</tr>
<tr>
<td>C. tentans</td>
<td>Chironomus tentans</td>
</tr>
<tr>
<td>DBD</td>
<td>DNA-binding domain</td>
</tr>
<tr>
<td>DNA</td>
<td>deoxyribonucleic acid</td>
</tr>
<tr>
<td>D. melanogaster</td>
<td>Drosophila melanogaster</td>
</tr>
<tr>
<td>EJC</td>
<td>exon junction complex</td>
</tr>
<tr>
<td>EM</td>
<td>electron microscopy</td>
</tr>
<tr>
<td>ESE</td>
<td>exonic splicing enhancer</td>
</tr>
<tr>
<td>ET</td>
<td>electron tomography</td>
</tr>
<tr>
<td>FGC</td>
<td>fibrogranular cluster</td>
</tr>
<tr>
<td>GFP</td>
<td>green fluorescent protein</td>
</tr>
<tr>
<td>hnRNP</td>
<td>heterogenous nuclear RNP</td>
</tr>
<tr>
<td>IEM</td>
<td>immuno-electron microscopy</td>
</tr>
<tr>
<td>IGC</td>
<td>interchromatin granule cluster</td>
</tr>
<tr>
<td>mRNA</td>
<td>messenger RNA</td>
</tr>
<tr>
<td>mRNP</td>
<td>messenger RNP</td>
</tr>
<tr>
<td>NLS</td>
<td>nuclear localization signal</td>
</tr>
<tr>
<td>NPC</td>
<td>nuclear pore complex</td>
</tr>
<tr>
<td>Nup</td>
<td>nucleoporin</td>
</tr>
<tr>
<td>ORF</td>
<td>open reading frame</td>
</tr>
<tr>
<td>PBD</td>
<td>protein-binding domain</td>
</tr>
<tr>
<td>PML</td>
<td>promyolytic leukemia</td>
</tr>
<tr>
<td>pol II</td>
<td>RNA polymerase II</td>
</tr>
<tr>
<td>poly(A)</td>
<td>poly-adenosine</td>
</tr>
<tr>
<td>pre-mRNA</td>
<td>precursor messenger RNA</td>
</tr>
<tr>
<td>RNA</td>
<td>ribonucleic acid</td>
</tr>
<tr>
<td>RNAi</td>
<td>RNA interference</td>
</tr>
<tr>
<td>RNase</td>
<td>ribonuclease</td>
</tr>
<tr>
<td>RNP</td>
<td>ribonucleoprotein</td>
</tr>
<tr>
<td>RRM</td>
<td>RNA recognition motif</td>
</tr>
<tr>
<td>SFC</td>
<td>splicing factor compartment</td>
</tr>
<tr>
<td>snoRNA</td>
<td>small nucleolar RNA</td>
</tr>
<tr>
<td>snRNP</td>
<td>small nuclear ribonucleoprotein</td>
</tr>
<tr>
<td>TEM</td>
<td>transmission electron microscopy</td>
</tr>
<tr>
<td>TF</td>
<td>transcription factor</td>
</tr>
</tbody>
</table>
General introduction

Organization of the eukaryotic cell nucleus

The eukaryotic cell nucleus contains the genomic information of the cell and is the site of a number of metabolic events, including DNA replication and repair, pre-mRNA synthesis and processing, and ribosome subunit synthesis and assembly. Over the last decades, the development of methods to visualize proteins and nucleic acids in fixed cells by fluorescence microscopy has both enabled detailed, high-resolution observations of cellular structures and revealed an intricate spatial organization within the cell nucleus. The expression of proteins with fluorescent tags in living cells has established the existence of subcompartments in the nucleus, and has also unveiled the highly dynamic properties of nuclear components (reviewed by Carmo-Fonseca 2002, Chubb and Bickmore 2003, Belmont 2003).

In contrast to cytoplasmic organelles, nuclear subcompartments are not enclosed by membranes and many of their resident proteins are highly dynamic (Misteli 2001). Still, like their cytoplasmic counterparts, nuclear compartments can be identified morphologically by light microscopy and electron microscopy, and they are characterized by a distinct composition of proteins. In recent years, several nuclear compartments have been biochemically purified and subjected to proteomic analysis (Mintz et al. 1999, Rout et al. 2000, Dreger et al. 2001, Andersen et al. 2002, Cronshaw et al. 2002, Lam et al. 2002, Scherl et al. 2002).

The most discernible features of the cell nucleus are the nuclear envelope, the chromatin, and the nucleoli. In addition, eukaryotic cells display a large number of other structures that have been characterized in various cell types under different metabolic conditions. I will here give a brief description of the prominent constituents of the nucleus, as well as of some of the other nuclear compartments and structures that have been described.
The nuclear envelope

The nucleus is surrounded by a double lipid bilayer termed the nuclear envelope, which on the cytoplasmic face continues into the endoplasmic reticulum, with which ribosomes are associated. The inner membrane of the nuclear envelope displays an underlying two-dimensional filamentous structure known as the nuclear lamina, which is composed of intermediate filament proteins called lamins. The lamins and additional proteins of the inner nuclear membrane provide structural support for the inner nuclear membrane, and they also anchor chromosomes to the nuclear periphery and regulate nuclear envelope formation (reviewed by Holaska et al. 2002). In addition, these proteins are found in the nuclear interior and participate in various nuclear events, including DNA replication and transcription, cell cycle regulation, cell development and differentiation, nuclear migration, and apoptosis (reviewed by Hutchinson 2002, Mattout-Drubezki and Gruenbaum 2003).
The nuclear pore complex

The nuclear envelope is interspersed with large proteinaceous channels that perforate the two membranes, termed the nuclear pore complexes (NPCs). The NPCs mediate all exchange of molecules that take place between the nuclear compartment and the cytoplasmic compartment. While small molecules such as ions can diffuse through the nuclear pores, translocation of larger molecules is facilitated by signal-dependent mechanisms that are mediated by soluble transport receptors (pages 23-24).

The structure of the NPC has been extensively studied by electron microscopy (EM) and is highly conserved in eukaryotes. Its main parts include the cytoplasmic fibrils, a central core, and the nuclear basket. The central core of the vertebrate NPC displays an eight-fold symmetry of spokes that surround the central channel through which all active transport occurs. The spokes are sandwiched between two ring structures that anchor the cytoplasmic fibrils and the nuclear fibrils. On the nuclear side, fibrils extending from the nuclear ring are joined at the distal end, forming a basket-like structure. It has been suggested that these filamentous structures, extending approximately 50 nm on the cytoplasmic face and twice as far into the nuclear interior, are docking sites for transport complexes (reviewed by Stoffler et al. 1999, Kiseleva et al. 2000, Suntharalingam and Wente 2003).

The protein components of the NPC are collectively referred to as nucleoporins (Nups). Many of the Nups are characterized by multiple repeats of the amino acids phenylalanine and glycine (FG-repeats), which mediate interactions with transport receptors. The yeast NPC contains approximately 30 different Nups, many of which are present in multiple copies. The Nups are distributed either symmetrically or asymmetrically between the nuclear and the cytoplasmic side of the NPC, with some Nups located more towards the central core and some more towards the periphery of the structure (Rout et al. 2000). Comprehensive proteomics-based approaches have been used to map the Nups in purified NPC-enriched nuclear fractions (Rout et al. 2000, Cronshaw et al. 2002). In yeast, each Nup has also been epitope-tagged and localized by immuno-EM (IEM) (Rout et al. 2000). Moreover, various efforts have been made to describe in detail the architectural framework of the NPC by elucidating the nearest-neighbor interaction for each Nup. For example, a large-
scale mass spectrometry analysis has been made of proteins that interact with a panel of Nups (Allen et al. 2001, 2002). In addition, NPC interactions have been analyzed in vivo by fluorescence resonance energy transfer (FRET), using Nups tagged with green fluorescent protein (GFP) (Damelin and Silver 2002).

Chromatin

In eukaryotic cells, chromatin is composed of nucleosomes, comprising octamers of four different histone proteins (H2A, H2B, H3 and H4) around which about 146 bp of DNA is wrapped. The nucleosomes are interconnected by the linker histone H1, and the chromatin fiber is coiled into a 30 nm fiber (reviewed by Widom 1989). It has been shown that chromatin is further organized into large-scale structures in vivo (reviewed by Belmont et al. 1999).

Chromosomes are confined into distinct regions in the nuclear space referred to as “chromosome territories” (reviewed by Cremer and Cremer 2001, Spector 2003). The fact that chromosomes occupy distinct nuclear regions during interphase was originally demonstrated by UV-irradiation of specific nuclear areas of Chinese hamster cells, which destroyed discrete chromosomal regions (Cremer et al. 1982). The chromatin domain surfaces enclose a chromatin-free space that is referred to as the nucleoplasm, the interchromatin space, or the interchromosomal domain (ICD). Chromosome territories comprise regions of closely compacted chromatin, or heterochromatin, and regions having a more open chromatin structure, referred to as euchromatin. While most genes are located in the transcriptionally active euchromatic regions, the heterochromatic regions are relatively gene-poor and transcriptionally inactive, and their replication in S-phase is delayed in relation to the euchromatic regions. Heterochromatin results from the spreading of chromatin remodeling complexes along the chromatin fiber (reviewed by Grewal and Moazed 2003). These protein complexes can mediate gene silencing by modifying the amino-terminal ends of the histone proteins, predominantly by deacetylation or methylation of lysine residues, thereby causing a more tight association of nucleosomes with the DNA (reviewed by Jenuwein and Allis 2001). Large regions of heterochromatin surround functional chromosome structures such as centromeres and telomeres,
whereas smaller heterochromatic regions are interspersed throughout the chromosome (reviewed by Grewal and Elgin 2002).

When active transcription sites are visualized by pulse-labeling techniques, thousands of discrete transcription sites appear that are scattered throughout the entire nucleus (Jackson et al. 1993, Wansink et al. 1993). Active genes localize both at the periphery of the chromosome territories and within these territories, suggesting that chromosome domains are porous and allow diffusion of transcriptional activators within the territory (Mahy et al. 2002).

The positioning of genes within the nucleus can influence their transcriptional activity. For example, proper silencing of genes involved in B cell and T cell development depends on the ability of the genes to relocalize to centromeric heterochromatin (Brown et al., 1997, 1999). In Drosophila, association of the brown locus with heterochromatin correlates with its transcriptional silencing (Dernburg et al. 1996). In yeast, silencing of the mating-type loci depends on the ability of these loci to associate with the nuclear periphery (Andrulis et al., 1998, 2002). In addition, a recent study in yeast indicates that the transcriptional activity of genes is coupled to their association with the NPC and the nuclear transport machinery (Casolari et al. 2004).

The nucleolus

The most well-characterized subnuclear compartment, the nucleolus, is organized around clusters of repeated ribosomal DNA (rDNA) genes in the so-called nucleolar organizer regions (NORs). In the nucleolus, the rDNA is transcribed by RNA polymerase I, and the pre-rRNA is subsequently modified and processed by a set of components that includes the small nucleolar ribonucleoproteins (snoRNAs) (reviewed by Kiss 2002). Each nucleolus contains three distinct components: the fibrillar center (FCs), which contains the rDNA genes; the dense fibrillar center (DFC), into which the nascent rRNAs extend and undergo processing; and the granular component (GC), where the last steps of rRNA maturation take place (reviewed by Lafontaine and Tollervey 2001, Fatica and Tollervey 2002). In addition to its fundamental role in rRNA biogenesis, the nucleolus is implicated in various other functions, including viral replication, apoptosis, cell cycle regulation, and modification of small
RNAs (reviewed by Carmo-Fonseca et al. 2000, Olson et al. 2002). Recent proteomic analyses have identified over 300 human proteins that co-purify with the nucleolus, reflecting the complexity of nucleolar function (Andersen et al. 2002, Scherl et al. 2002, http://www.lamondlab.com).

Other subnuclear structures

Splicing factor compartments/speckles

Compartmentsthat are referred to as splicing speckles, or splicing factor compartments (SFCs) (Spector 1990, Spector et al. 1991), occupy a large volume of the nuclear space. As indicated by their name, splicing speckles show high concentrations of pre-mRNA splicing components, including small nuclear ribonucleoproteins (snRNPs) and SR proteins (see page 22). Transmission electron microscopy (TEM) analysis of splicing speckles reveals two distinct structures: the interchromatin granule clusters (IGCs) in the central part of the compartment, and the more peripheral perichromatin fibrils which are thought to mainly represent nascent transcripts (Monneron and Bernhard 1969, Fakan 1994).

Most active genes are found at the periphery of splicing speckles, rather than within the compartment (Xing et al. 1993, 1995). Splicing factors accumulate in splicing speckles upon inhibition of transcription (Xing et al. 1993, Melcak 2000), and both expression of intron-containing genes and viral infection can mediate redistribution of splicing factors from the speckles towards transcription sites (Jimenez-Garcia and Spector 1993, Huang and Spector 1996, Misteli et al. 1997). Based on this, it has been proposed that splicing speckles function as storage and/or assembly sites for spliceosomal components (Misteli 2000, Lamond and Spector 2003). The recruitment of splicing factors to sites of nascent transcription may involve a cycle of phosphorylation and dephosphorylation that acts on SR proteins (Misteli and Spector 1996, Misteli et al. 1998). In addition to spliceosomal components, splicing speckles also contain transcription factors and 3’-end processing factors (Mintz et al. 1999). Furthermore, IGCs contain both cellular poly (A)+ RNA and viral RNAs (Visa et al. 1993, Huang et al. 1994, Bridge et al. 1996, Puvion and
Puvion-Dutilleul 1996), although the significance of mRNA constituents in splicing speckles is presently unclear.

Cajal bodies
As early as the beginning of the twentieth century, the “nucleolar accessory bodies” were described by the Spanish cytologist Ramón y Cajal. Later on, the same structures have been referred to as “coiled bodies”, until they were renamed “Cajal bodies” (CBs) (Gall 1999). CBs are small spherical structures that are usually present in 1-10 copies per nucleus. A number of different components are found in CBs, including spliceosomal snRNPs, snoRNAs, transcription factors, factors involved in 3´-end formation, and nucleolar components (reviewed by Gall 2000). Although CBs have been observed inside nucleoli in some cell types, they are more commonly found within the nucleoplasm. Moreover, CBs are highly mobile in the nucleoplasm, and can move to and from the nucleolar periphery as well as within nucleoli (Boudonck et al. 1999, Platani et al. 2000, 2002). CBs are formed in a transcription-dependent manner and they dissemble and reassemble during each cell cycle (Andrade et al. 1993, Ferreira et al. 1994). Several lines of evidence suggest that CBs play a role in the maturation of small RNAs. For instance, snRNPs and snoRNAs localize to CBs before moving to splicing speckles or nucleoli (Sleeman and Lamond 1999, Verheggen 2002). CBs also associate with certain U snRNA genes (Jacobs et al. 1999) and with histone gene clusters (Shopland 2001). In addition, it has been proposed that CBs function in the pre-assembly of the RNA polymerase II (RNA pol II) transcription machinery and pre-mRNA processing factors into a multi-subunit transcriptosome complex (Gall et al. 1999, Morgan et al. 2000).

PML bodies
Promyolytic leukemia (PML) bodies are small spherical domains that vary in size and number through the cell cycle (Koken et al. 1995, Terris et al. 1995). These structures are specifically disrupted in human acute promyolytic leukemia cells. The most prominent component of PML bodies is the protein PML, which appears to be essential for the formation and integrity of these structures (Wang et al. 1998). The function of
PML bodies is unclear, but they have been implicated in various processes including transcriptional regulation, nuclear storage, and apoptosis (reviewed by Borden 2002).

**Nuclear stress bodies**

It has been reported that a small number of nuclear bodies form in transformed cells in response to various stress treatments. These bodies have been named nuclear stress bodies, HAP bodies, or stress-induced SAM68 nuclear bodies (reviewed by Biamonti 2004). At the ultrastructural level, nuclear stress bodies consist of clusters of perichromatin granules, morphological structures that may correspond to transient storage sites for incompletely or abnormally processed mRNAs (Puvion and Lange 1980, Chiodi et al. 2000). Nuclear stress bodies contain heat-shock factors, a subset of RNA processing factors including the hnRNP M protein (Chiordi et al. 2000), the hnRNP A1-interacting protein (HAP), and some splicing factors of the SR protein family (Denegri et al. 2001). The assembly of nuclear stress bodies depends on the synthesis of non-coding RNAs from a heterochromatic region and subsequent recruitment of RNA-processing factors to this region (reviewed by Biamonti 2004, Sandqvist and Sistonen 2004).

**Paraspeckles**

In mammalian cells, punctuate structures of unknown function have been observed that are adjacent to, but distinct from, the splicing speckles. These structures have been named paraspeckles, and contain the proteins PSP-1, PSP-2 and p54\textsuperscript{nrh} (Fox et al. 2002). The fact that all three proteins contain RNA recognition motifs suggests a role for paraspeckles in RNA biogenesis. Interestingly, PSP-1 and p54\textsuperscript{nrh} are structurally related and share a conserved element, which is also present in the mammalian protein PSF and in its insect homologues (Shav-Tal and Zipori 2003, Kiesler et al. 2003). The function of PSP-1 is currently not known, while p54\textsuperscript{nrh} has been implicated in several aspects of gene expression, including transcriptional regulation and pre-mRNA splicing (see pages 41-42). PSP-2, which is not structurally related to PSP-1 and p54\textsuperscript{nrh}, has been identified as a transcriptional coactivator activator (Iwasaki et al. 2001). The components of paraspeckles have associate with nucleoli in a tran-
scription-dependent manner, indicating that a functional relationship may exist between the paraspeckle and the nucleolus (Fox et al. 2002, Andersen et al. 2002).

The nuclear matrix

The nuclear matrix has been described as a non-chromatin structure of interconnected fibers and granules that can be observed in cells by EM (Fawcett 1966). This structure consists of two parts: the nuclear lamina (see above) and an interior part that is thought mainly to correspond to a ribonucleoprotein network (Nickerson 2001). The concept of the nuclear matrix has often rested on operational definitions, describing it as an insoluble protein residue that remains after the extraction of nuclei with a high concentration of salt or detergent (Berezney and Coffey 1974). The concern that treatment of cells under non-physiological conditions may cause artificial aggregation of proteins and RNPs has led to doubt whether a nuclear non-chromatin network actually exists in the living cell (Pederson 2000). Methods for extracting cell nuclei have improved significantly in recent years (reviewed by Nickerson 2001, Pederson 1998) and some observations have indicated the existence of filamentous non-chromatin structures in living cells (Oegama 1997). However, depictions of the nuclear matrix as a “nuclear cytoskeleton” that functions in organizing nuclear subcompartments and influences gene regulation (for example, as described by Berezney 1991, Agutter 1994) have more recently been subjected to reservations (Pederson 2000, Cremer et al. 2000). It may be more useful to revise the concept of the nuclear matrix as a rigid and inflexible scaffold than it is to dispute the existence of insoluble filamentous structures in the nucleus. Recent discoveries of the kinetic properties of nuclear components have shown that many, if not all, of the structures in the nucleus are maintained by transient interactions of proteins that are highly dynamic (reviewed by Misteli 2001). A relevant example is provided by the ICGs, which are part of the nuclear matrix (Monneron and Bernhard 1969, Spector et al. 1983) and for which the dynamic properties are nowadays well-established (see page 16). The existence and the functional significance of a nuclear matrix in living cells remains an open question.
Expression of protein-coding genes in eukaryotic cells

In eukaryotic cells, the separation of the nuclear compartment and the cytoplasmic compartment leads to a segregation of different steps in gene expression. Messenger RNA precursors are synthesized and processed into translatable mRNA transcripts in the cell nucleus, after which the mRNAs are exported to the cytoplasm where they guide the synthesis of proteins by the ribosomes (Figure 2). Although the different steps in the gene expression pathway are mediated by distinct multimolecular complexes, and most of these steps can be reproduced individually in vitro, there is substantial evidence that these activities are coordinated and functionally integrated in the living cell (reviewed by Bentley 2002, Maniatis and Reed 2002, Proudfoot et al. 2002).

(Pre-)mRNA is dynamically associated with many RNA-binding proteins throughout its biogenesis. These RNA-binding proteins, together with the transcript itself, form a large ribonucleoprotein (RNP) complex. Proteins known as hnRNP proteins become associated with the pre-mRNP particle already during transcription (reviewed by Dreyfuss et al. 2003). As many of these factors remain on the RNP complex throughout its presence in the nucleus, or even after translocation of the mRNP to the cytoplasm (see Table 1), the hnRNP proteins are able to play a role in linking the various steps of gene expression into a connected pathway. Throughout the course of their biogenesis, mRNP complexes are altered.
significantly with respect to the nucleic acid sequence of the transcript and with respect to the composition of associated proteins. Accordingly, the assembly and biogenesis of pre-mRNAs in eukaryotic cells can be envisioned as a chain of metabolic events occurring consecutively, which to some extent are spatially segregated, but nevertheless, functionally integrated.

Synthesis and processing of pre-mRNA transcripts

RNA polymerase II transcription

Protein-coding genes are transcribed in the cell nucleus exclusively by RNA polymerase II (RNA pol II), a multi-subunit enzyme of approximate molecular weight 4 MDa, into pre-messenger RNA precursors (pre-mRNAs). The largest subunit of RNA pol II contains a carboxyterminal domain (CTD) that is composed of tandem repeats of the heptapeptide sequence YSPTSPS. This sequence is conserved among eukaryotes, although the number of repeats varies between species. Phosphorylation of the CTD is a key event during the transcription cycle. During initiation of transcription, RNA pol II containing a hypophosphorylated form of the CTD (pol IIa) is recruited to the promoter sequence of the gene to be transcribed and forms a pre-initiation complex together with general transcription factors (GTFs). Upon phosphorylation of serine residues in the CTD heptapeptides, the polymerase enters its second form (pol IIo) and transcription elongation begins.

The control of RNA pol II transcription is mediated by complex combinatorial networks of transcription factors (TFs) that function at different levels. Sequence-specific DNA-binding TFs bind to transcriptional regulatory regions, such as proximal promoter regions or distal enhancers and silencers, and either favor or impede the assembly of the pre-initiation complex (reviewed by Kadonaga 2004). Through protein-protein interactions, the DNA-binding TFs can recruit a wide variety of additional TFs, collectively termed coregulators, to the regulatory sequences. Coregulators are either activators or repressors of transcription and their regulation of RNA pol II activity can be either direct, by interacting with the general transcription machinery, or indirect. For instance, chromatin-related TFs, such as histone acetyl transferases
(HATs) and chromatin remodelling complexes, are indirect coregulators that alter the structure of chromatin and regulate access of the transcriptional machinery to the DNA (reviewed by Belotserkovskaya and Berger 1999, Vignali et al. 2000). In addition to the factors that are implicated in the formation of the pre-initiation complex, there is another type of TF that acts on the level of transcription elongation. Some transcription elongation factors can modify elongation efficiency by interacting directly with RNA pol II. Furthermore, recent experiments in yeast have revealed the existence of yet another type of elongation factor that acts by binding to the nascent pre-mRNA and destabilizing RNA:DNA hybrids. The THO/TREX complex, also implicated in mRNA export (see page 28), belongs to this last type (Huertas and Aguilera 2003).

**Pre-mRNA processing**

Before leaving the nucleus, the pre-mRNA undergoes several enzymatically catalyzed modification events. Shortly after transcription initiation, a 7-methylguanosine cap is added to the 5´ end of the nascent transcript. The 3´ end of most pre-mRNAs is generated by cleavage and subsequent addition of a poly(A) tail. In the majority of genes in higher eukaryotes, the coding exon sequences are interrupted by introns that are removed from the pre-mRNAs by splicing. The process of pre-mRNA splicing is a two-step transesterification reaction that is catalyzed by the spliceosome, a large ribonucleoprotein complex that is composed of four small nuclear ribonucleoprotein particles (snRNPs U1, U2, U4/U6 and U5) and a large number of non-snRNP splicing factors (reviewed by Nilsen 2003, Jurica and Moore 2003). The non-snRNP splicing factors include a family of proteins called SR proteins, which can promote spliceosome assembly by binding to exonic sequences in the pre-mRNA. These sequences are called exonic splicing enhancers (ESEs) (reviewed by Blencowe 2000, Gravely 2000). Many pre-mRNAs of higher eukaryotes are alternatively spliced, so that a single pre-mRNA molecule can encode a variety of protein isoforms. Consequently, the process of alternative splicing generates a high degree of diversity in the genomes of higher eukaryotes, and allows for the expression of specific protein isoforms in a cell-type specific or developmentally regulated manner (reviewed by Maniatis and Tasic 2002, Caceres and Kornblihtt 2002, Black 2003).
The procedures of capping, 3'-end formation and splicing all stimulate gene expression by increasing the stability, export, and translational yield of the transcript. Although each of these processing events can be reconstituted in isolation in vitro, it appears that they commonly take place cotranscriptionally in vivo, and that they can influence each other’s efficiency and specificity. The phosphorylated CTD plays a key role in the assembly and regulation of processing factors during transcriptional elongation (reviewed by Hirose and Manley 2000, Bentley 2002, Proudfoot et al. 2002). Direct involvement of the CTD in pre-mRNA processing has been further supported by structure determination of the yeast RNA pol II, which revealed that the newly synthesized RNA leaves the pol II complex in the vicinity of the CTD (Cramer et al. 2001).

A notable example of the coupling of early transcription events to downstream processing events is given by the finding that changes in promoter structure can regulate alternative splicing by affecting splice site selection (Cramer et al. 1997, 1999). It has been proposed that transcriptional coregulators can mediate this coupling, as some steroid hormone receptor coregulators affect splice site selection in a promoter-dependent manner (Auboeuf et al. 2002).

Export of mRNA to the cytoplasm

The exchange of all macromolecules between the nucleus and the cytoplasm occurs through the NPCs (see pages 13-14). Large cargoes, including most proteins, transfer RNAs, ribosomal RNAs and U snRNAs, are transported from the nucleus to the cytoplasm by an energy-dependent, signal-mediated process that is mediated by soluble transport receptors of the karyopherin family (reviewed by Weis 2003). These receptors are shuttling proteins that recognize different types of nuclear transport signals in the cargo and mediate movement of the cargo by association with NPC components. Nucleocytoplasmic transport mediated by karyopherins depends on the small GTPase Ran, which regulates receptor-cargo interactions and determines the directionality of transport in and out of the nucleus (reviewed by Görlich and Kutay 1999, Nakielny and Dreyfuss 1999).
The mRNA export receptor NXF1/Mex67

It appears that export of most mRNAs from the nucleus to the cytoplasm is not mediated by any member of the karyopherin family, but instead by an evolutionarily conserved protein named NXF1 in metazoa (often referred to as TAP in vertebrates) and Mex67p in yeast. NXF1/Mex67 was originally identified as a factor that mediates export of some viral RNAs by binding to the constitutive transport element (CTE), an RNA element used by certain retroviruses for the export of non-spliced viral RNA (Gruter et al. 1998, Kang and Cullen 1999). Subsequent studies performed in yeast, C. elegans, and D. melanogaster have shown that NXF1/Mex67 also mediates export of cellular mRNA (Segref et al. 1997, Tan et al. 2000, Herold et al. 2001). NXF1/Mex67 are shuttling proteins that, together with their co-factor p15/Mtr2, interact with nucleoporins and are thought to promote mRNA export by mediating interactions between the mRNP and the NPC (Bachi et al. 2000, Strässer et al. 2000).

Adapters for NXF1/Mex67-mediated mRNA export

In contrast to the CTE RNAs, cellular mRNAs do not seem to interact directly with the NXF1/Mex67 - p15/Mtr2 complex, but are thought to do so through interaction with adapter proteins on the mRNP particle. It has been proposed that proteins belonging to the REF family, including the murine protein Aly and its yeast homologue Yra1, mediate the binding of NXF1/Mex67 to the mRNP (Stutz et al. 2000, Rodrigues et al. 2001). Aly/Yra1 interacts directly with RNA and with TAP/Mex67 (Strässer and Hurt 2000, Stutz et al. 2000). It has also been proposed that some of the shuttling SR proteins serve as adapter proteins for TAP-mediated mRNA export (Caceres et al. 1998, Huang and Steitz 2001, Huang et al. 2003). It has recently been shown that the yeast protein Npl3, an essential mRNA export factor belonging to the SR protein family (Lee et al. 1996, Russel and Tollervey 1995, Singleton et al. 1995), binds directly to Mex67 in vitro (Gibert and Guthrie 2004).

Other mRNA export factors

Apart from the mRNA export factors described above, a number of additional factors have been identified that are crucial for mRNA export. Most of these factors have initially been identified by screenings in yeast...
and subsequently been shown to have functional homologues in vertebrates. One highly conserved protein that is known to participate in the export of mRNA is a putative DExD box helicase protein called Sub2 in yeast, UAP56 in mammals, and HEL in insects (reviewed by Linder and Stutz 2001). The human UAP56 and its yeast homologue Sub2 were originally identified as splicing factors (Fleckner et al. 1997, Libri et al. 2001, Zhang and Green 2001, Kistler and Guthrie 2001). Soon after, it was discovered that Sub2 and its homologues in higher eukaryotes are essential for mRNA export (Strässer and Hurt 2001, Jensen et al. 2001a, Luo et al. 2001 Gatfield et al. 2001), and that they interact directly with Aly/REF/Yra1 (Luo et al. 2001, Strässer and Hurt 2001) (see page 27). A genome-wide analysis of mRNA export pathways in Drosophila revealed that HEL/UAP56 acts in the same mRNA export pathway as NXF1-p15, and that the majority of mRNAs are exported through this pathway (Herold et al. 2003).

Dbp5 is another RNA helicase that is essential for mRNA export in yeast and in vertebrates (Snay-Hodge et al. 1998, Tseng et al. 1998, Schmitt et al. 1999). Due to its predominantly cytoplasmic distribution and to its location at the cytoplasmic fibrils of the NPC, Dbp5 is thought to function during late steps of mRNA export (Schmitt et al. 1999, Strahm et al. 1999). However, a fraction of Dbp5 in C. tentans is located at the nucleoplasmic side of the NPC, and the protein is incorporated into Balbiani ring pre-mRNAs early during transcription and remains associated with the mRNP particle during transport to the cytoplasm (Zhao et al. 2002). In addition, recent studies in yeast have shown that Dbp5 interacts with the transcriptional machinery (Estruch and Cole 2003). Taken together, these observations suggest an early role for Dbp5.

Gle1 and Gle2/Rae1 are two other factors that are thought to participate in late steps of the mRNA export pathway. Gle1 is essential for mRNA export in both yeast and vertebrates (Murphy and Wente 1996, Watkins et al. 1998), and has been located in yeast to the cytoplasmic fibrils of the NPC (Strahm et al. 1999). In human cells, the function of Gle1 in promoting mRNA export depends on its ability to shuttle between the nucleus and the cytoplasm (Kendirgi et al. 2003). Gle2/Rae1 is another mRNA export factor with shuttling activity (Brown et al. 1995, Bharathi et al. 1997, Pritchard et al. 1999). Studies from C. tentans suggest that Rae1 initially becomes associated with mRNPs at the NPC (Sabri and Visa 2000), and, based on the interactions of Gle2/Rae1 with TAP/NXF1 and
Nup98, it has been proposed that it mediates mRNA export by aiding the delivery of TAP/NXF1 to the NPC (Blevins et al. 2003).

It has also been suggested that some of the abundant and shuttling hnRNP proteins, e.g. hnRNP A1, mediate mRNA export (Michael et al. 1995, Visa et al. 1996a, Izaurralde et al. 1997). However, hnRNP A1 does not seem to be involved in mRNA export since its shuttling motif is not accessible on the fully assembled mRNP (reviewed by Siomi and Dreyfuss 1997).

**Directionality of mRNA transport**

Unlike other nucleocytoplasmic transport receptors, mRNA export receptors of the NXF1 family do not interact with Ran, and the mechanism that governs the directionality of the mRNA export pathway is not known. Several mechanisms have been suggested that could bring about an irreversible release of the mRNP cargo from its receptor in the cytoplasm (Erkmann and Kutay 2004, Izaurralde 2004). One possibility is that mRNPs undergo compartment-specific remodelling by RNA helicases, and it has been proposed that Dbp5 removes non-shuttling proteins from the mRNP before the release of the particle into the cytoplasm (Snay-Hodge et al. 1998, Tseng et al. 1998, Strahm et al. 1999). In addition, ribosomes seem to be able to remove certain proteins from the mRNP during the first round of translation (reviewed by Maquat 2004). Another mechanism that has been proposed is that shuttling mRNA-binding proteins are subjected to compartment-specific post-translational modifications. In yeast, the shuttling mRNA export factor Npl3 is phosphorylated by the cytoplasmic kinase Sky1 (Gilbert et al. 2001). It has recently been shown that cytoplasmic phosphorylation and nuclear dephosphorylation of Npl3 are required for efficient mRNA export, and that dephosphorylation of Npl3 may affect the recruitment of Mex67 to the mRNP (Gilbert and Guthrie 2004). Based on these findings, a model has been proposed where a cycle of localized dephosphorylation and phosphorylation of Npl3 mediates the binding of the Mex67-Mtr2 receptor complex to the mRNP in the nucleus and its subsequent release from the mRNP in the cytoplasm (Gilbert and Guthrie 2004, Izaurralde 2004).
Splicing and mRNA export

Although it has been known for a long time that introns have a stimulatory effect on gene expression, it is not clear at what step in the gene expression pathway this effect acts. One explanation of how the presence of introns in a pre-mRNA can enhance gene expression is provided by the fact that components of the pre-mRNA splicing machinery are physically coupled to components of the mRNA export machinery (reviewed by Reed and Hurt 2002). In higher eukaryotes, Aly/REF forms a complex called the exon junction complex (EJC) with SRm160, RNPS1, Y14, Mago, DEK and HEL/UAP56. The EJC is deposited onto the mRNA as a direct consequence of splicing (Kataoka et al. 2000, Le Hir et al. 2000a, Le Hir et al. 2000b, McGarvey et al. 2000, Zhou et al. 2000, Luo et al. 2001). It has been suggested that acquisition of an EJC is necessary for efficient export of intron-containing pre-mRNAs to the cytoplasm (Luo and Reed 1999, Zhou et al. 2000). However, Aly/REF and other components of the EJCs were recently found to be dispensable for mRNA export in D. melanogaster and in C. elegans, and it has thus been proposed that they facilitate rather than directly mediate mRNA export (Gatfield and Izaurralde 2002, Longman et al. 2003).

The finding that UAP56/HEL/Sub2p interacts with REF/Yra1, together with the role of HEL/UAP56 in splicing, has led to the proposal that HEL binds to the pre-mRNA during splicing and recruits the EJC to the spliced mRNA (Strässer and Hurt 2001, Luo et al. 2001). However, HEL and Sub2p are also required for the export of intron-less pre-mRNAs (Strässer and Hurt 2001, Jensen et al. 2001a, Gatfield et al. 2001), indicating that the role of HEL in pre-mRNA export is not necessarily linked to splicing. This is supported by investigations of HEL in C. tentans. The results from quantitative IEM studies presented in Paper IV show that HEL binds progressively to the pre-mRNA as transcription proceeds, regardless of the position of introns along the pre-mRNA. These findings are presented in Paper IV and further discussed on pages 51-52.

Some splicing factors interact with downstream cleavage and polyadenylation factors (Vagner et al. 2000, McCracken et al. 2002, and references therein). As discussed below, the release of transcripts after transcription depends on 3′ end formation of the transcript, and thus the
presence of introns may enhance the efficiency of expression of a transcript by stimulating the cleavage and polyadenylation machinery.

Furthermore, recent studies have implied that the removal of introns can increase the expression level of a mRNA by stimulating its translation, or by enhancing mRNA abundance (Nott et al. 2003, Lu et al. 2003). In effect, components of the EJC have been shown to function in processes that influence the cytoplasmic fate of mRNAs, such as nonsense-mediated mRNA decay (NMD - a pathway that degrades mRNAs carrying premature stop codons) (Lykke-Andersen et al. 2001) and cytoplasmic mRNA localization (Mohr et al. 2001, Hachet et al. 2001).

Transcription and mRNA export

It has recently been shown that UAP56/HEL/Sub2 and Aly/NXF1 interact with the components of a four-component complex involved in transcription elongation, the THO complex (Chavez et al. 2000), forming a complex known as the transcription-export (TREX) complex (Strässer et al. 2002). Additional studies in yeast indicate that TREX components become associated with the pre-mRNA during transcription elongation and that the THO complex recruits Yra1 and Sub2 to the nascent pre-mRNA (Zenklusen et al. 2002). In addition, both the THO components and Sub2 are needed for efficient transcription elongation (Rondon et al. 2003), demonstrating a functional connection between transcription and mRNA export.

In conclusion, transcription elongation and pre-mRNA splicing are two distinct mechanisms by which factors can be loaded onto the pre-mRNA. The export fate of a transcript appears to be determined primarily during transcription. The recruitment of factors to a transcript in the process of splicing may have additional effects on the export of the mRNA cargo and on subsequent gene expression events.

Pre-mRNA surveillance and nuclear retention

Various mechanisms are starting to be elucidated by which eukaryotic cells can monitor the quality and accuracy of mRNP formation and prevent expression of aberrant transcripts. In mammalian cells, inefficiently processed pre-mRNAs, or pre-mRNAs containing premature termination
codons, are retained at or near the site of transcription (Custodio et al. 1999, Muhlemann et al. 2001). A similar nuclear retention phenomenon has been observed in yeast cells, where release of transcripts from their transcription sites appears to depend on proper 3’-end formation (Jensen et al. 2001b, Hilleren et al. 2001). The retention and degradation of aberrant transcripts close to the transcription site is thought to be mediated by the exosome, a multisubunit complex of 3’ to 5’ exonucleases (reviewed by Jensen et al. 2003). The mechanisms that mediate the sequestration of transcripts close to the gene are not known, but one way to achieve such retention is transient binding of the unprocessed pre-mRNP to non-diffusible structures inside the nucleus.

**Intranuclear movement and localization of (pre-) mRNPs**

Although the pathways by which mRNP particles are exported from the nucleus to the cytoplasm are being discovered, several aspects of this process remain unresolved. One critical aspect to clarify the mechanisms behind intranuclear mRNA trafficking is to understand how the newly synthesized mRNP particle travels from the transcription site to the nuclear envelope. Early studies suggested that RNPs are transported in a directed manner through the nucleoplasm (Blobel 1985, Agutter 1994), but evidence has accumulated showing that the translocation of mRNAs from the gene to the nuclear pores is mainly accounted for by passive diffusion (e.g., Politz and Pederson 2000, Shav-Tal et al. 2004). Nevertheless, the question remains whether transported mRNPs diffuse throughout their passage through the nucleoplasm, or whether they also interact with nuclear structures.

**Dynamics of intranuclear (pre-) mRNPs**

The kinetic properties of intranuclear mRNP particles have been studied in living cells using different detection systems. In two pioneering studies, Politz et al. (1998, 1999) labeled poly(A)+ RNAs *in vivo* using a fluorescently labeled oligo-dT probe and analyzed the movement of the labeled molecules using two detection methods. These studies showed that the labeled mRNPs formed two populations: a major population that showed diffusion rates comparable to that of an average-sized mRNP in
solution, and a smaller population that moved at a very low speed, suggesting that this fraction of mRNPs is associated with very large macromolecular structures. In a recent study, poly(A)+ RNAs were labeled \textit{in vivo} using a fluorescent 2'O-methyl RNA probe instead of an oligo-d(T) (Molenaar et al. 2004). In this case, analysis of the movement of labeled molecules by photobleaching techniques revealed that poly(A)+ RNAs are not only mobile within the nucleoplasm, but also dynamically associated with splicing speckles in a transcription-independent manner. As an alternative strategy for labeling of mRNPs \textit{in vivo}, mRNA-binding proteins have been fused to GFP and their movement has been analyzed by photobleaching techniques (Calapez et al. 2002). These experiments revealed that energy-depletion reduces the mobility of the RNA-bound GFP fusion proteins, suggesting that the kinetic behaviour of intranuclear mRNPs reflects a combination of passive diffusion and ATP-dependent processes (Calapez et al. 2002). A similar energy-dependence has been observed for the movement of RNAs labeled by 2'O-methyl RNA probes into and out of splicing speckles (Molenaar et al. 2004). Recently, a study of the intranuclear movement of single mRNPs expressed from a transgene was reported, which corroborated a free diffusion model for mRNPs (Shav-Tal et al. 2004). In this study, a fraction of the mRNPs showed corralled movement within the nucleoplasmic space, but no accumulation of mRNPs at non-chromatin structures was observed. Moreover, based on the finding that energy depletion causes rapid changes in chromatin structure, the authors proposed that the changes in mRNP mobility that have been observed after energy depletion result from the formation of chromatin barriers (Shav-Tal et al. 2004). Intranuclear mRNP movement has also been studied by pulse-labeling of newly synthesized RNA in salivary gland cells of \textit{C. tentans} by bromo-deoxyuridine triphosphate (BrUTP)-incorporation (Singh et al. 1999). Immuno-EM analysis of Balbiani ring pre-mRNA transcripts following the fixation of cells at sequential time points revealed that the majority of these transcripts moved away from the gene in a random manner, at rates compatible with free diffusion (Singh et al. 1999).
(Pre-) mRNPs and nuclear structures

There are a number of other studies in addition to the kinetic investigations that implicate the association of mRNPs to nuclear structures. For example, fluorescently labeled pre-mRNA injected into the nuclei of living cells localizes to splicing factor-rich compartments in an intron-dependent manner (Wang et al. 1991). IGCs contain cellular mRNAs in addition to splicing factors (Visa et al. 1993, Huang et al. 1994), and viral RNAs accumulate in IGCs upon infection (Bridge et al. 1996, Puvion and Puvion-Dutilleul 1996). It was recently shown in yeast that nuclear retention of intron-containing mRNAs is mediated by Mlp1, a component of a structural platform at the nuclear periphery (Galy et al. 2004). Furthermore, it has been shown by ultra-structural studies performed in C. tentans that (pre-) mRNP particles interact transiently with non-chromatin nucleoplasmic structures (Miralles et al. 2000). These findings are discussed in more detail on pages 36-37.

To conclude, there is considerable evidence that at least a large fraction of mRNPs move randomly throughout the nucleus, and free diffusion is generally believed to be the mechanism by which mRNPs move from the gene to the nuclear pore. It is still controversial whether the mobility of mRNPs is energy-dependent. There are also experimental results supporting retention of mRNPs by association with non-chromatin structures. In light of the extensive functional couplings between transcription, pre-mRNA processing, surveillance, and mRNA export, it is conceivable that association of mRNPs to nuclear structures allows the cell to control the release of impaired transcripts, or incompletely processed transcripts, thereby preventing their diffusion to the nuclear pores and subsequent translocation to the cytoplasm. Still, very little is known about the mechanisms behind nuclear retention of mRNA, or about the possible role of subnuclear structures in mRNA biogenesis.
An introduction to the model system and to the present investigation

The Balbiani ring genes of *Chironomus tentans*: a model system for *in situ* studies of gene expression

Many of the steps in the gene expression pathway can be reproduced individually *in vitro*, and the molecules or multi-molecular complexes responsible for mediating the enzymatic activities involved in gene expression have been identified and characterized in considerable detail at the biochemical level. However, nucleocytoplasmic transport can only be envisaged in relation to the cellular architecture and for this reason our understanding of the transport mechanisms relies mostly on *in situ* studies. The larval salivary glands of the dipteran *Chironomus tentans* constitute a valuable model system for *in situ* studies of gene expression, as they offer the possibility of directly visualizing the synthesis, processing, and export of a specific transcript: the Balbiani ring (BR) pre-mRNA.

*The active BR genes*

The larval salivary gland cells of *C. tentans* possess a polytene nucleus with four giant chromosomes (reviewed by Hägele 1975, Case and Daneholt 1977). At a gross level, four major structures can be easily identified in these nuclei: the nuclear envelope, the polytene chromosomes, the nucleoli and the nucleoplasm (Figure 3). Each polytene chromosome consists of thousands of chromatids arranged in register and the variable degree of chromatin condensation along each chromosome results in a characteristic banding pattern that allows the identification of specific loci. The chromatin is expanded or puffed at transcriptionally active sites.

Chromosome IV contains two puffs of exceptional size, referred to as Balbiani rings (BR) 1 and 2, which arise as a consequence of the expression of the BR genes (reviewed by Wieslander 1994). The pre-mRNA synthesized in BR1 and BR2 is unusually long (35-40 kb). Upon transcription, the BR pre-mRNA is associated with proteins and assembled into a large ribonucleoprotein (RNP) complex called the BR RNP particle. Due to
their structure and extraordinary dimensions, the BR RNP particles can be unambiguously identified by TEM, and their assembly, transport and disassembly can be directly studied in situ (reviewed by Mehlin and Daneholt 1993, Daneholt 2001). Nascent BR pre-mRNA molecules are rapidly packaged into growing RNP fibers that can be observed along the BR genes (Figure 4). As transcription progresses, the growing RNP fibers gradually increase in length, which confers a distinct polarity to the active BR transcription unit (Skoglund et al. 1983). When transcription is complete, the BR particles are released from the chromosome and can be seen in the nucleoplasm as RNP granules with a diameter of about 50 nm. The BR RNP granules move towards the nuclear envelope, where they unfold and become elongated during translocation through the nuclear pores (Mehlin et al. 1992).

![Figure 3](image_url)

*Figure 3.* Electron micrograph showing the nucleus of a salivary gland cell from *C. tentans.* CH chromosome, NE nuclear envelope, NP nucleoplasm, NU nucleolus. The bar represents 5 µm. Reprinted with permission from Kiesler and Visa 2004. © Springer-Verlag.

**Biogenesis of BR pre-mRNAs**

In spite of their unusual length, the BR pre-mRNAs have all the features of typical pre-mRNAs. The 5' end is capped, the pre-mRNA is cleaved and polyadenylated at the 3' end, and the four short introns are spliced prior to export of the mRNA to the cytoplasm (Visa et al. 1996b, Baurén et al. 1998, Baurén and Wieslander 1994). Kinetic experiments *in vivo*
have shown that the BR mRNA is efficiently exported to the cytoplasm and is highly stable (Edström et al. 1978 and references therein).

Figure 4. Schematic representation of BR (pre)mRNP particles showing successive stages of synthesis and assembly of the pre-mRNP at the BR gene, transport through the nucleoplasm, docking at the NPC, translocation to the cytoplasm and translation in polysomes. The polarity of the gene (5' to 3') is indicated. Reprinted with permission from Kiesler and Visa 2004. © Springer-Verlag.

Proteins associated with BR pre-mRNAs

One of the major contributions of the BR system to the study of gene expression in situ is the possibility of analyzing the association of defined RNA-binding proteins with the BR RNP particles during successive stages of synthesis, maturation, and export using IEM (reviewed by Daneholt 2001). Thus, it is possible in the BR system to determine at what step of the gene expression pathway a specific protein interacts with the (pre-)mRNA. As shown in Table 1, 15 proteins have been identified that associate with the BR RNP particles in the nucleus and it is possible to identify a mammalian orthologue for most of them. Some of the proteins found in the BR RNP particles, such as hrp23, hrp36 and hrp45, are abundant proteins present in multiple copies along the BR pre-mRNP and it is likely that these proteins play a role in the packaging of the BR pre-mRNA. However, some other proteins, such as CBP20 and PABPN1, are bound to specific structural elements of the BR (pre-)mRNA and probably have specialized functions in mRNA biogenesis.
It is interesting to note that, although the proteins found in the BR RNP particle are of different types and probably have specific functions in the gene expression pathway, most of them become associated with the BR pre-mRNA as early as transcription. Based on this observation, it has been proposed that the fate of the mRNA is determined concomitantly with transcription, during the early assembly of the RNP at the gene (Daneholt 2001).

Table 1. BR RNA-associated nuclear proteins

<table>
<thead>
<tr>
<th>Protein (1)</th>
<th>Human Orthologs(s)</th>
<th>Putative Function (2)</th>
<th>Association with BR RNP (3)</th>
<th>References (4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hrp23/RSF1</td>
<td>?</td>
<td>packaging/splicing</td>
<td>gene to NPC</td>
<td>Sun et al. 1998</td>
</tr>
<tr>
<td>Hrp36</td>
<td>hnRNP A/B</td>
<td>packaging</td>
<td>gene to polysomes</td>
<td>Visa et al. 1996a</td>
</tr>
<tr>
<td>Hrp45</td>
<td>?</td>
<td>packaging/splicing</td>
<td>gene to NPC</td>
<td>Alzhanova-Ericsson et al. 1996</td>
</tr>
<tr>
<td>Hrp59 (5)</td>
<td>hnRNP M</td>
<td>splicing/transcription</td>
<td>gene to NPC</td>
<td></td>
</tr>
<tr>
<td>Hrp84 (6)</td>
<td>DBX</td>
<td>translation initiation</td>
<td>gene to polysomes</td>
<td></td>
</tr>
<tr>
<td>P40/50</td>
<td>P50/YB-1</td>
<td>translation</td>
<td>gene to polysomes</td>
<td>Soop et al. 2003</td>
</tr>
<tr>
<td>RAE1</td>
<td>mrnp40</td>
<td>nuclear export</td>
<td>NPC proximity</td>
<td>Sabri et al. 2000</td>
</tr>
<tr>
<td>Dbp5</td>
<td>Dbp5</td>
<td>nuclear export</td>
<td>gene to NPC</td>
<td>Zhao et al. 2002</td>
</tr>
<tr>
<td>HEL</td>
<td>UAP56</td>
<td>splicing/export</td>
<td>gene to NPC</td>
<td>Kiesler et al. 2002</td>
</tr>
<tr>
<td>REF</td>
<td>Aly/REF</td>
<td>nuclear export</td>
<td>gene to NPC</td>
<td>Kiesler et al. 2002</td>
</tr>
<tr>
<td>PABPN1</td>
<td>PABP1</td>
<td>poly(A)-binding</td>
<td>gene to NPC</td>
<td>Bear et al. 2003</td>
</tr>
<tr>
<td>CBP20</td>
<td>CBP20</td>
<td>cap-binding</td>
<td>gene to NPC</td>
<td>Visa et al. 1996b</td>
</tr>
<tr>
<td>Hrp65</td>
<td>PSP1, PSF, p54&lt;sub&gt;arb&lt;/sub&gt;</td>
<td>splicing/retention/transcription</td>
<td>gene to NPC</td>
<td>Miralles et al. 2000</td>
</tr>
<tr>
<td>p2D10</td>
<td>TFIICα</td>
<td>transcription/possible post-transcriptional</td>
<td>gene to NPC</td>
<td>Sabri et al. 2002</td>
</tr>
<tr>
<td>actin</td>
<td>actin</td>
<td>transcription/possible post-transcriptional</td>
<td>gene to polysomes</td>
<td>Percipalle et al. 2001</td>
</tr>
</tbody>
</table>

1) Name of the protein in _C. tentans_
2) Based on known function of orthologs in other species and on localization of the protein in _C. tentans_
3) As shown by IEM
4) Reference to the _C.tentans_ protein
5) E. Kiesler, M. Hase and N. Visa (Paper I)
BR particles associate with nucleoplasmic structures in salivary gland cells of *C. tentans*

Early TEM investigations of nucleoplasmic BR particles suggested that many particles that are in transit from the gene to the NPC are associated with thin fibers that extend into the surrounding nucleoplasm (Figure 5a). The nucleoplasmic BR RNP particles have been analyzed by electron tomography (ET) of cryosections of fixed salivary gland cells in order to obtain reliable 3D information about possible interactions between the BR RNP particles and other nuclear structures. Miralles and co-workers (2000) showed using this method that approximately one third of the nucleoplasmic BR particles do not show any sign of binding interactions, and the morphology of these BR particles is thus compatible with a free diffusion model. However, an equally large fraction of nucleoplasmic BR RNP particles was in direct contact with thin fibers, referred to as connecting fibers, or CFs (Figure 5b). The CFs have a diameter of approximately 7 nm and a length in the 15-50 nm range. In some cases the CFs merged into larger structures of unknown function called fibrogranular clusters, or FGCs. Although the function of these nucleoplasmic structures remains to be elucidated, it is conceivable that the binding of BR RNP particles to large non-chromatin structures, such as CFs and FGCs, imposes restrictions on the movement of the BR particles. Thus, the BR RNP particles attached to CFs and FGCs cannot diffuse freely in the nucleoplasm, at least not at rates expected for free RNPs (Miralles et al. 2000).

**Composition of CFs and FGCs**

The CFs can be labeled by IEM with a monoclonal antibody raised against nucleoplasmic proteins from *C. tentans* salivary gland cells. This antibody enabled the identification of the RNA-binding protein Hrp65 as a component of the CFs (Miralles et al. 2000). We show in Paper II that Hrp65 can self-associate and form oligomers of multiple subunits, which may mediate the formation of CFs *in vivo*. Another nucleoplasmic protein in *C. tentans*, p2D10, has been shown by IEM to have a broad nuclear distribution and to locate to FGCs (Sabri et al. 2002). As discussed in Papers I and III, IEM analysis of HEL and Hrp59 suggests that these two
RNA-binding proteins are also part of CFs and/or FGCs. Identification of additional components of CFs and FGCs will be important in unravelling the function of these structures and the molecular basis for their interaction with the BR RNP particles.

Figure 5. 3D-reconstructions of nucleoplasmic BR particles. A) Electron micrograph showing a nucleoplasmic BR particle in a thin cryosection of a salivary gland observed in the TEM. The BR particle appears as a dense granule with a diameter of about 50 nm. The arrow points at a fiber that contacts the BR particle and extends into the surrounding nucleoplasm. The bar represents 100 nm. B) Three examples of nucleoplasmic BR RNP particles analyzed by electron tomography. The numbers on the images (1 to 4) refer to the domains of the BR particles (Skoglund et al. 1986). The BR particle shown in a is apparently free, not in contact with any other structure. The BR particles in b and c are in contact with connecting fibers (CFs) that extend into the nucleoplasm. The fiber in c is continuous with a fibrogranular cluster (FCG) located in the upper left corner of the image. The bar represents 25 nm. Reprinted with permission from Kiesler and Visa 2004. © Springer-Verlag.

The *C. tentans* Hrp65 protein

Hrp65 is a basic 65-kDa protein that has two classical RNA recognition motifs (RRMs) located between two auxiliary domains. Three isoforms of the protein have been characterized, generated by alternative splicing of one pre-mRNA (Figure 6). Two of the isoforms, Hrp65-1 and Hrp65-
2, are ubiquitously expressed during larval and adult development of *C. tentans*, while a third isoform, Hrp65-3, has so far only been detected in a tissue culture-derived cell line of embryonic origin (Miralles 2001).

**Figure 6.** Graphical depictions of the three Hrp65 isoforms. The sequence between amino acids 1-499 is common to all isoforms, but the C-terminal sequences are variable.

The Hrp65-2 isoform interacts with actin via its isoform-specific C-terminal sequence and plays a role in transcription in complex with actin (Percipalle et al. 2003). It is still not clear whether the suggested role in transcription is restricted to the Hrp65-2 isoform, and further work is required to establish the function(s) of each Hrp65 isoform.

Thus far, four different antibodies have been generated against Hrp65. Two of these, the monoclonal antibody 4E9 and a peptide-specific polyclonal rabbit antibody, recognize all isoforms of Hrp65. The two remaining antibodies are isoform-specific polyclonal antibodies recognizing the C-terminal sequences of Hrp65-1 and Hrp65-2, respectively. Although the 4E9 antibody has been useful for immunofluorescent labeling of *C. tentans* cells and chromosomes (Miralles et al. 2000 and Paper I), this antibody cannot be used for IEM as it gives poor labeling on cryosections. However, by IEM labeling of cryosections of *C. tentans* salivary glands with the polyclonal antibody that recognizes all isoforms, it was shown that Hrp65 is present in CFs associated with nucleoplasmic BR particles (Miralles et al. 2000). The presence of Hrp65 in CFs has later been confirmed by IEM labeling using the Hrp65-1 specific antibody (Paper III). The Hrp65-1 specific antibody also labels the nucleoplasmic BR particles to some extent. It is likely that this occasional detection of Hrp65-1 on BR particles results from overlap of CFs with BR particles in the cryosection. However, the presence of Hrp65 on the nucleoplasmic BR particle itself cannot be excluded, and further analysis by electron
tomography (ET) in combination with immunolabeling is needed to clarify this point.

Immunofluorescent staining using either of the two antibodies recognizing all Hrp65 isoforms gives very low labeling, or no labeling, of the BR1 and BR2 puffs on chromosome IV, although the protein is detected at other loci (Miralles et al. 2000 and Paper I). This suggests that Hrp65 is not a component of the BR RNP particle, but binds to other transcripts, directly or indirectly, in a locus-specific manner.

The Hrp65-2 specific antibody has very low immunoreactivity when used for cytology. Still, when this antibody was used for IEM at high concentrations, Hrp65-2 was detected on nascent BR particles (Percipalle et al. 2003). This work showed that the distribution of Hrp65-2 along the BR gene is opposite to that of typical hnRNP proteins: instead of being continuously incorporated during transcription elongation, the abundance of Hrp65-2 on BR particles decreases as the particles elongate at the gene. The absence of labeling at BR puffs observed with the two antibodies that recognize all isoforms, however, is not consistent with this observation. As Hrp65-2 has been shown to play a role in transcription elongation together with actin (Percipalle et al. 2003), it is possible that Hrp65-2 is recruited to the gene by a mechanism that is different from that by which many other RNA-binding proteins are incorporated into the pre-mRNA. Some epitopes may be concealed in this case. On the other hand, the absence of labeling may also be explained by the low abundance of Hrp65-2 on nascent transcripts.

When taken together, these observations suggest that Hrp65 is involved in different gene expression events and that the different isoforms can participate in specific functions. At the BR1 and BR2 genes, Hrp65-2 is present but not enriched, while Hrp65-1 is probably absent. Nucleoplasmic BR particles that have been released after transcription do not seem to contain either Hrp65-1 or Hrp65-2, but Hrp65-1 (and maybe Hrp65-2) is present in CFs in the nucleoplasm. Besides, some cytoplasmic staining is also observed with the anti-Hrp65 antibodies, especially with the antibody specific for Hrp65-2 (Paper II).
Proteins related to Hrp65

Sequence analysis reveals that Hrp65 is highly similar to a group of RNA-binding proteins characterized by the presence of a “Drosophila behavior and human splicing” (DBHS) domain (Dong et al. 1993). This group of proteins includes mammalian PSP-1, PSF, p54nrp/NonO and Drosophila NonA/Bj6 (Fox et al. 2002, Patton et al. 1993, Dong et al. 1993, Yang et al. 1993, Jones and Rubin 1990, von Besser et al. 1990). The DBHS proteins appear to have diverse functions in gene expression.

The Drosophila NonA/Bj6 protein

The Drosophila protein NonA/Bj6 is most similar to Hrp65, having 53% sequence identity with Hrp65. The two proteins are structurally very similar, not only in the DBHS domain but also in the C-terminal domain. Moreover, the genes that code for Hrp65 and NonA in Chironomus and Drosophila, respectively, are similar in terms of exon organization and splicing patterns (Miralles and Visa 2001).

The function of NonA at the molecular level is not known. The protein was first identified as a chromatin protein that binds to developmentally regulated puffs in a transcription-dependent manner (Frasch and Saumweber 1989, von Besser et al. 1990, Saumweber et al. 1990). NonA binds in vitro to single-stranded DNA and RNA with moderate affinity (Reim et al. 1999b). Amorphic mutations of the nonA gene cause defects in the nervous system, resulting in a number of specific defects including deterioration of vision, behaviour and co-ordination of movement (Jones and Rubin 1990, Rendahl et al. 1992, Stanewsky et al. 1996). A complete loss of function in NonA has shown that the protein is cell lethal in early embryogenesis but not during larval and adult life, indicating that NonA is essential early in development (Stanewsky et al. 1993). It is possible that the nonA-like gene, which is highly homologous to nonA, can recover some of the vital functions in the absence of nonA during later developmental stages (Rendahl et al. 1996). It has been shown by site-directed mutagenesis that the RNP-2 and RNP-1 consensus sequences of the first RRM domain are essential for biological NonA function, although these mutations do not affect the location of the protein in chromosome puffs (Stanewsky 1996, Reim et al. 1999b).
In *Drosophila* cells, NonA is present in high molecular weight complexes and associates with a number of RNA-binding proteins (Reim et al. 1999a, Hovemann et al. 2000). The same has also been shown for Hrp65 (Miralles et al. 2000 and Papers I and III), suggesting that NonA and Hrp65 share the same binding partners *in vivo* and that they may be functionally equivalent.

**The mammalian DBHS proteins**

The mammalian poly-pyrimidine tract-binding protein (PTB)-associated splicing factor (PSF) was originally identified and characterized in a complex with PTB (Patton et al. 1993). However, later studies showed that a major part of PSF in the nucleus is associated with the nuclear matrix, where it does not interact with PTB (Meissner et al. 2000). Immunofluorescent images of PSF reveal a punctuate distribution in nuclei of mammalian cells, with PSF spots at the same locations as splicing speckles (Shav-Tal et al. 2001a, 2001b, Dye and Patton 2001).

A protein named p54
\(^{\text{mb}}\) in human and NonO in mouse (Dong et al. 1993, Yang et al. 1993), and the human paraspeckle protein 1 (PSP-1) (Fox et al. 2002), are closely related to PSF. PSP-1 and p54
\(^{\text{mb}}\)/NonO, together with another nuclear factor PSP-2, are components of nuclear compartments of unknown function termed paraspeckles (see page 18) (Fox et al. 2002). In addition, proteomic investigations have identified PSF, p54
\(^{\text{mb}}\)/NonO, and PSP-1 in the nucleolus upon inhibition of transcription (Andersen et al. 2002), and PSF and p54
\(^{\text{mb}}\)/NonO in association with the nuclear membrane (Dreger et al. 2001).

PSF and p54
\(^{\text{mb}}\)/NonO are commonly found associated with each other *in vivo* (Straub et al. 1998, Mathur et al. 2001, Zhang and Carmichael 2001, Emili et al. 2002, Peng et al. 2002) and they both participate in a variety of nuclear processes. Originally identified as a splicing factor, PSF binds to the poly-pyrimidine tract of mammalian introns and is essential for splicing *in vitro* (Patton et al. 1993). Other studies suggested that PSF participates in the second catalytical step of pre-mRNA splicing, where two exons are joined and an intron lariat is released (Gozani et al. 1994, Lindsay et al. 1995). More recently, p54
\(^{\text{mb}}\)/NonO has been implicated in splicing, together with PSF. The two proteins have been reported to bind to the U5 snRNA and to associate with spliceosomes (Peng et al. 2002). In addition, PSF and p54
\(^{\text{mb}}\)/NonO bind to the CTD of both RNA
pol IIa and RNA pol IIo (Emili et al. 2002). Moreover, p54nrbinds directly to the 5’ splice site of introns as part of large protein complexes that contain RNA pol IIo, snRNPs and several factors involved in transcription elongation and splicing, including PSF (Kameoka et al. 2004). Taken together, these latter findings suggest a role for PSF and p54nrb/NonO in mediating contact between the transcription- and splicing machineries.

PSF and p54nrb/NonO bind to RNA and to DNA (Zhang et al. 1993, Patton et al. 1993, Yang et al. 1993, Basu et al. 1997) and they participate in transcriptional regulation. For example, PSF down-regulates the transcriptional activity of the porcine P-450 cholesterol side-chain cleavage gene (P450scc) by directly binding to an insulin-like growth factor (IGF) response element within this gene through its N-terminal domain (Urban et al. 2000, 2002). Also, PSF represses transcription by binding to the DNA-binding domain of thyroid hormone receptors (TRs) and retinoid X receptors (RXRs) in complex with p54nrb/NonO, thereby recruiting the Sin3A protein, which mediates gene silencing by assembling histone deacetylases (HDACs) (Mathur et al. 2001). Similarly, PSF and p54nrb/NonO, in complex with Sin3A and steroidogenic factor (SF-1), repress the transcription of human CYP17, a gene involved in steroidogenesis (Sewer et al. 2001).

Another interesting function that has been attributed to p54nrb/NonO and PSF is that of mediating nuclear retention of inosine-containing double-stranded RNAs (dsRNA) (Zhang and Carmichael 2001). Long dsRNA structures generated by sense-antisense RNA interactions are unselectively hyperedited by dsRNA-dependent deaminases (ADARs) in the nucleus, where a large part of the adenosines in the RNA are deaminated into inosines (reviewed by Bass 2002). Such extensively edited RNAs are retained in the nucleus by a protein complex containing p54nrb, PSF and matrin-3, a structural protein of the nuclear matrix. p54nrb has a high affinity for inosine-containing RNA and it has been suggested that the hyperedited RNAs are retained in the nucleus by anchoring to the nuclear matrix via the p54nrb/PSF/matin3 complex.

PSF and p54nrb/NonO have also been implicated in other processes, including tumorigenesis, DNA unwinding and DNA pairing (reviewed by Shav-Tal and Zipori 2002).
To conclude, Hrp65 belongs to a group of evolutionarily conserved proteins that are involved in a variety of nuclear processes that associate with a large variety of factors. Thus, it is difficult to infer the function of Hrp65 from what is known about the mammalian homologues, and further investigation of Hrp65 and Hrp65-containing complexes is required to gain insight into the function of this protein in pre-mRNA biogenesis.
Present investigation

Aim and experimental approach

Two main objectives underlie the investigations that are presented in this thesis. One objective was to further investigate the composition of CFs and FGCs, nucleoplasmic structures that associate with BR (pre)-mRNP particles in salivary gland cells of *C. tentans*. The second objective was to further examine the Hrp65 protein that composes the CFs, and to investigate its role in (pre)-mRNA biogenesis. For both purposes, our initial strategy has been to identify and characterize proteins that interact with Hrp65.

We have used the yeast two-hybrid screening method to identify proteins in *C. tentans* that interact with Hrp65. The Hrp65-binding proteins have been cloned and characterized at the molecular level, and their distributions at the cellular and ultrastructural levels have been analyzed in situ. Hrp65-containing complexes have been analyzed using a combination of biochemical methods. Questions related to the functional aspects of Hrp65 and its associated proteins have been addressed by taking advantage of different eukaryotic model systems.

Results and discussion

We have identified four different Hrp65-interacting proteins using the yeast two-hybrid screening method (Fields and Song 1989, Chien et al. 1991) (Table 2). The complete open reading frame of the Hrp65-1 isoform was fused to the GAL4 DNA-binding domain (DBD), and used as a bait to screen an expression library obtained by fusion of the GAL4 trans-activation domain (AD) to cDNA obtained from *C. tentans* tissue culture cells. The interactions between bait and GAL4 AD-fusion proteins were analyzed in a yeast strain containing three reporter genes: *ade2, his3* and
lacZ/mel1. Initially, we screened 4x10⁶ clones by selecting for the expression of his3 and lacZ/mel1. Approximately 300 positive clones were selected in this initial screen, and these clones were subsequently analyzed for expression of ade2, his3 and lacZ/mel1. This triple-selection resulted in the survival of 52 clones that were further analyzed by control experiments in the yeast two-hybrid system and sequenced. In one of the sequenced clones, the GAL4 AD was found to be out of frame with the open reading frame (ORF) encoded by the cDNA (not shown), and the rest of the clones were found to encode three different proteins: Hrp65 itself, Hrp59 and Y9 (Table 2).

The HEL/UAP56 protein was initially found in a separate yeast two-hybrid screen using the C-terminal part of Hrp65-2. Interaction of HEL with full-length Hrp65-1 was then demonstrated by control experiments in the two-hybrid system (Paper III).

<table>
<thead>
<tr>
<th>Protein</th>
<th>Known domains/motifs</th>
<th>Number of isolated clones</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hrp65¹</td>
<td>RRM (2)/DBHS</td>
<td>21</td>
</tr>
<tr>
<td>Hrp59¹</td>
<td>RRM (3)</td>
<td>5</td>
</tr>
<tr>
<td>Y9¹</td>
<td>RING, PDZ</td>
<td>25</td>
</tr>
<tr>
<td>HEL/UAP56²</td>
<td>DECD box</td>
<td>1</td>
</tr>
</tbody>
</table>

*Table 2.* Hrp65-interacting proteins identified by yeast two-hybrid screening.
¹ Identified using the complete ORF of Hrp65-1 as bait
² Identified using the amino acids 306-517 of Hrp65-2 as bait

**Studies on Hrp59 (Paper I)**

In the yeast two-hybrid screen with Hrp65-1, we have isolated a novel hnRNP protein in *C. tentans* named Hrp59. The cDNA sequence in the isolated clones contains three RRMs and is 48% identical to a previously uncharacterized protein in *Drosophila*. There are two homologues of Hrp59 in mammals: the hnRNP M protein, which has previously been implicated in splicing regulation, and Myef-2, which has been implicated in transcriptional repression. We have verified the interaction of Hrp59 and Hrp65 by *in vitro* binding experiments, and we have also shown by
immunoprecipitation experiments that the proteins are associated in the same complexes in vivo. This association is in part mediated by binding of the two proteins to the same RNAs as shown by ribonuclease (RNase) digestion of immunoprecipitated complexes, and in part by direct protein-protein interaction as shown by immunoprecipitation experiments performed after covalent cross-linking of protein complexes in vivo.

Immunofluorescent analysis of Hrp59 in C. tentans salivary gland cells revealed that the protein is predominantly located in the nucleus, where it is abundant in the nucleoplasm and in many bands on the polytene chromosomes. Hrp59 co-localizes with Hrp65 in the majority of these bands. However, the two proteins are present in different proportions in many bands, suggesting that they are independently recruited to the same target loci. We have also observed that the distribution of Hrp59 and Hrp65 on the chromosomes is distinct from that of other major hnRNP proteins.

We have shown that treatment of C. tentans salivary glands with RNase prior to fixation and immunostaining displaced both proteins from the chromosomes without significantly altering their nucleoplasmic distribution, indicating that Hrp59 and Hrp65 associate with chromosomes by binding to nascent RNAs. After labeling of newly synthesized RNA in C. tentans salivary gland cells by BrUTP incorporation, we could compare the distribution of Hrp59 with that of active transcription sites by immunofluorescent analysis. This revealed that Hrp59 is present in bands on polytene chromosomes that overlap with sites of ongoing transcription. Not all transcription sites were labeled with the anti-Hrp59 antibody, and in the majority of loci, the intensity of Hrp59 labeling did not parallel the level of BrUTP incorporation. These observations indicated that Hrp59 has specific preferences for certain pre-mRNAs, and that its recruitment to most loci is not proportional to the transcriptional activity in these loci. Furthermore, when analyzing the distribution of Hrp59 and Hrp65 on polytene chromosomes after heat-shock, we observed that both proteins redistribute and become associated with major heat-shock puffs, showing that they can be recruited to genes upon transcriptional induction.

We have applied IEM to analyze the association of Hrp59 with nucleoplasmic BR RNP particles that have been released after transcription. A significant fraction of the nucleoplasmic BR particles were
labeled with the anti-Hrp59 antibody, including BR particles that are in transit through the NPC. On these translocating BR particles, the labeling was always detected on the nuclear side of the particle. Our IEM data suggest that Hrp59 remains associated with the BR particle after transcription until the BR particle translocates through the nuclear pore, and that the protein dissociates from the mRNP during the translocation process. Furthermore, the IEM analysis revealed that the majority of Hrp59 protein in the nucleus is not located on BR particles but on other nucleoplasmic structures, including CFs adjacent to BR particles. In some cases the antibody appeared to label the part of the CF that is proximal to the BR particle, and in other cases the labeling was detected on CFs further away from the particle or in larger structures that we believe are FGCs. It is interesting to note that the human hnRNP M protein associates with insoluble nucleoplasmic structures upon thermal stress (Mähl et al. 1989, Gattoni et al. 1996). These structures have subsequently been identified as nuclear stress bodies (see page 18) (Chiodi et al. 2000). The functional significance of this stress-induced redistribution of hnRNP M is not known, but it has been proposed that hnRNP M mediates heat-shock induced splicing arrest (Gattoni et al. 1996), and that recruitment of RNA-binding proteins to nuclear stress bodies can affect the alternative splicing pattern of selected genes (Biamonti 2004). In view of these observations on hnRNP M, it would be interesting to investigate whether the association of Hrp59 with CFs and FGCs changes in response to heat-shock.

The hypothesis derived from the immunofluorescence studies that Hrp59 associates preferentially with a subset of pre-mRNAs prompted us to identify the target transcripts that bind to Hrp59 in vivo. By virtue of the evolutionary conservation of Hrp59, we could perform a genome-wide search of Hrp59-targets in Drosophila. This was done by isolation of protein-RNA complexes from Drosophila S2 cells by immunoprecipitation, followed by hybridization of the reverse-transcribed RNA within these complexes to oligonucleotide microarrays covering the Drosophila genome. By selecting for genes that showed multifold enrichment of RNA in the immunoprecipitated samples, we identified 32 target transcripts, that corresponded to the most enriched mRNAs in the immunoprecipitated samples. Among these targets was the Hrp59 mRNA itself. To confirm this, we have visualized the Hrp59 locus on polytene chromosomes of C. tentans salivary gland cells by fluorescent in situ hybridization (FISH), and demonstrated the presence of Hrp59 protein in this locus.
by immunofluorescent labeling. The fact that Hrp59 binds to its own pre-mRNA suggests that it can regulate its own expression.

Sequence comparison of the identified target transcripts revealed the presence of a conserved purine-rich sequence element in the majority of Hrp59 targets. Using an in vitro assay, we have shown that Hrp59 binds preferentially to an oligonucleotide corresponding to this purine-rich motif. Interestingly, the purine-rich Hrp59 target motif is found in exonic sequences and it resembles a common exonic splicing enhancer (ESE) sequence (see page 22). Hrp59 binds to intron-containing pre-mRNAs as well as to spliced transcripts, as shown by RT-PCR analysis of Hrp59-containing RNP complexes. In Paper I, we discuss the possibility that Hrp59 regulates splicing of its target transcripts by binding to ESEs.

Studies on Hrp65 self-interaction (Paper II)

About 40% of the positive clones that we isolated in the yeast two-hybrid screen using Hrp65-1 as bait encoded different isoforms of the Hrp65 protein itself. Using an in vitro binding assay with recombinant protein, we have demonstrated that the different Hrp65 isoforms can interact with each other in all possible combinations. We have then mapped the region of the Hrp65-Hrp65 interaction by expressing deletion constructs of the protein and examining the ability of these constructs to bind to the full-length protein in the yeast two-hybrid system. The interaction domain is located to a region within the conserved DBHS domain, downstream of the two RRM s, that we have named the protein-binding domain (PBD). By co-expressing the PBDs of either Drosophila NonA or human PSF with Hrp65 in the yeast two-hybrid system, we have found these domains to be sufficient to interact with Hrp65. This observation strongly suggests that the PBD is conserved throughout evolution and that it mediates heterodimerization of DBHS proteins. This is also corroborated by another study, which demonstrated that human p54nrb binds to PSF through a region that corresponds to the PBD of p54nrb (Peng et al. 2002). In view of these findings, the DBHS domain can be described as a compound domain, constituted by an upstream region, containing the two RRM s, which is presumably involved in protein-RNA interaction; and the downstream PBD, which can mediate homodimerization or heterodimerization. Given the diversity of functions that have been attributed to
proteins of the DBHS group (see pages 40-43), one may speculate that these proteins serve as platforms for the assembly of multimolecular complexes in different cellular contexts, and that the property of oligomerization would provide a mechanism by which the DBHS proteins can promote multimolecular assembly.

The observed self-interaction of Hrp65 raises the possibility that the protein can form oligomers that correspond \textit{in vivo} to the CFs that have been observed in association with BR pre-mRNP particles, and therefore it was interesting to analyze whether Hrp65 can assemble into larger structures \textit{in vitro}. Recombinant Hrp65 protein expressed in \textit{E. coli} was shown to form high molecular-weight complexes that were resistant to RNase treatment, indicating that Hrp65 oligomerization occurs by direct protein-protein interaction. By fractionation of purified recombinant Hrp65 protein on a gel filtration column, we have demonstrated that Hrp65 can oligomerize into complexes of approximately 200-400 kDa, corresponding to oligomers of three to six molecules. It is noteworthy that these fractionation experiments do not provide information about the size of complexes that formed \textit{in vivo}, and as discussed in Paper II, it is likely that Hrp65 can form larger insoluble aggregates under native conditions, that are not formed in our experimental conditions. Nevertheless, our examination of recombinant Hrp65 demonstrates that the protein is capable of building oligomeric complexes, and suggests that Hrp65 is a structural component of the CFs. We have tried to investigate whether oligomerization of Hrp65 contributes to formation of CFs \textit{in vivo} by microinjecting recombinant PBD into live salivary gland cells. However, the PBD polymerized very efficiently under physiological conditions, and could therefore not be used for microinjection. Further investigations are needed to determine the size of Hrp65 complexes that are formed \textit{in vivo}, and to establish whether oligomerization of Hrp65 can contribute to the assembly of CFs.

Interestingly, we have found that Hrp65 oligomerization is important for the intracellular location of different Hrp65 isoforms. The endogenous Hrp65-2 protein shows a clear nuclear localization in \textit{C. tentans} cells when analyzed using Hrp65-2 specific antibodies, in spite of the fact that a GFP-Hrp65-2 fusion construct does not enter the nucleus in transient transfection assays. However, by co-transfection assays in HeLa cells, we found that Hrp65-2 enters the nucleus when co-expressed with Hrp65-1, but not when co-expressed with an Hrp65-1 construct where the NLS has
been removed. From these experiments we concluded that import of Hrp65-2 into the nucleus depends on the NLS of Hrp65-1.

It is conceivable that the described nuclear import mechanism for Hrp65-2 provides a strategy for the cell to regulate the nucleo-cytoplasmic distribution of the Hrp65-2 isoform. Indeed, the fact that Hrp65-2 participates in transcription of class II genes in complex with actin in the nucleus (Percipalle et al. 2003) suggests that variations in the nuclear concentration of Hrp65-2 has broad effects on gene activity. It would be interesting to discover if Hrp65-2 is involved in any processes in the cytoplasm, perhaps also in complex with actin.

Studies on HEL/UAP56 (Papers III and IV)

We have isolated and sequenced a cDNA that encodes the *C. tentans* homologue of *Drosophila* HEL or human UAP56, a putative RNA helicase involved in pre-mRNA splicing and mRNA export (see page 25), using a truncated form of the Hrp65-2 isoform as bait in a yeast two-hybrid screening. This protein is strikingly conserved among eukaryotes, and HEL in *C. tentans* (Ct-HEL) is 90% identical to HEL in *Drosophila* and 85% identical to human UAP56. We have shown that the binding of HEL to Hrp65 is not restricted to the Hrp65-2 isoform by co-expressing the cDNA clone encoding HEL in yeast together with Hrp65-1. We have also shown that HEL binds to Hrp65-1 and Hrp65-2 *in vitro*. From these experiments we have concluded that a region within the C-terminal part of Hrp65 mediates the binding of Hrp65 to HEL.

Antibodies raised against HEL in *Drosophila* cross-react with the protein in *C. tentans* (Paper IV), and this allowed us to investigate Ct-HEL. We have shown by immunoprecipitation experiments that HEL and Hrp65 associate *in vivo*, and we have demonstrated that this interaction is direct by performing immunoprecipitation experiments after *in vivo* cross-linking of protein complexes.

As shown in Papers I and IV, both HEL and Hrp65 are RNA-binding proteins that associate with pre-mRNAs cotranscriptionally. However, immunofluorescent analysis of HEL and Hrp65 on polytene chromosomes of *C. tentans* salivary gland cells revealed that the two proteins are differently distributed on the chromosomes, suggesting that they are not
co-recruited to the same population of nascent pre-mRNAs. In particular, HEL is highly abundant in the actively transcribing BR1 and BR2 loci, while the Hrp65 labeling is very low at BR1 and BR2, being enriched at other specific loci. When analyzing by IEM the association of HEL with nucleoplasmic BR RNP particles that have been released after transcription from the gene, we have observed that the protein is present on the BR particle itself, and in CFs and possibly in FGCs. Given the fact that HEL and Hrp65 co-localize in CFs but not on polytene chromosomes, it is likely that the two proteins interact during post-transcriptional steps of gene expression.

Paper IV describes experiments in which we analyzed the interaction of HEL with the BR pre-mRNA in C. tentans by quantitative IEM. The anti-HEL antibody labeled BR pre-mRNP particles along the entire transcription unit, including the proximal part which contains early transcribing BR pre-mRNAs, indicating that HEL is incorporated into the newly synthesized pre-mRNA immediately after transcription initiation. The fact that multiple immuno-gold particles labeled each BR particle indicated that each BR pre-mRNA incorporates several molecules of HEL. We also analyzed the association of Aly/REF with the BR pre-mRNA. This protein was also detected in BR particles along the entire transcription unit, showing that Aly/REF associates cotranscriptionally with the BR pre-mRNA in the same way as HEL does.

HEL/UAP56/Sub2p was originally characterized as a splicing factor, and it has been proposed that it associates with the pre-mRNA during splicing and recruits the export factor Aly/REF/Yra1 to the spliced mRNA (see page 27). Yet, the fact that HEL/UAP56/Sub2p also participates in the transport of intronless mRNAs suggests that its role in export is separate from its role in splicing. We have shown that HEL is incorporated progressively into the BR pre-mRNA as transcription proceeds, in a manner that is independent of the position of introns in the pre-mRNA, by quantifying in IEM gold markers that label the BR particles. Our results are not compatible with the suggestion that HEL is deposited on the pre-mRNAs as a consequence of splicing, and they suggest that the protein can interact with the pre-mRNA by different means, i.e., by a mechanism that is related to the function of HEL in splicing, as well as by a different mechanism that is independent of splicing. In contrast to HEL, the abundance of Aly/REF on BR pre-mRNAs did not appear to increase along the BR transcription unit, suggesting that Aly/REF does
not associate progressively with the pre-mRNA throughout transcription elongation. This is in agreement with the proposal that REF is added to the pre-mRNA during splicing, as part of the EJC (see page 27).

We also analyzed the dissociation of HEL from the BR mRNP by comparing the relative amounts of HEL molecules detected on BR particles that are located in the nucleoplasm, docked at the NPC, or translocating through the nuclear pore. We found that HEL stays associated with the BR (pre-)mRNA until the BR RNP particle reaches the nuclear pore, and is released as the particle translocates through the pore and out to the cytoplasm. Although we observed a similar dissociation of Aly/REF from the BR mRNP during the translocation step, the relative decrease in labeling was lower than that observed for HEL, and we have therefore proposed that Aly/REF leaves the mRNP at a later stage in the translocation process.

Studies in yeast have shown that binding of Mex67, the functional homologue of the mRNA export receptor TAP/NXF1, to Yra1 (the homologue of Aly/REF) competes with the binding of Sub2 to Yra1 in vitro (Strässer and Hurt 2001). Based on this, a model has been proposed in which Mex67/TAP/NXF1 displaces Sub2/UAP56/HEL on the mRNP by binding to Yra1/Aly/REF, and consequently triggers translocation of the mRNP through the nuclear pore (Strässer and Hurt 2001, Linder and Stutz 2001). In consideration of this model it is conceivable that the dissociation of HEL from the BR mRNP early during the translocation process reflects binding of TAP/NXF1 to the BR mRNP. In this respect, it would be interesting to investigate the TAP/NXF1 protein in C. tentans and to analyze its association with the BR (pre)-mRNA in situ.

Studies on Y9

We have identified a novel protein in C. tentans, here referred to as Y9, in the yeast two-hybrid screen with Hrp65. About half of the isolated two-hybrid clones encoded the Y9 sequence (Table 2), and the two-hybrid interaction between Hrp65 and Y9 was confirmed by additional control experiments in the two-hybrid system (Table 3).
Table 3. Interaction between Hrp65 and Y9 in the yeast two-hybrid system. Yeast cells were transformed with a plasmid encoding the GAL4 AD-Y9 fusion together with either the GAL4 DBD-Hrp65-1 plasmid or with negative control plasmids encoding the GAL4 DBD alone, or a GAL4 DBD-p53 fusion. As a positive control, cells were transformed with a plasmid encoding the complete GAL4 factor. The protein-protein interactions were detected by plating the co-transformants onto double selective -His -Ade medium.

<table>
<thead>
<tr>
<th>plasmid 1</th>
<th>plasmid 2</th>
<th>Growth on -His -Ade</th>
</tr>
</thead>
<tbody>
<tr>
<td>pGAL4</td>
<td>-</td>
<td>+++</td>
</tr>
<tr>
<td>pDBD-GAL4-Hrp65-1</td>
<td>pAD-GAL4-Y9</td>
<td>++</td>
</tr>
<tr>
<td>pDBD-GAL4</td>
<td>AD-GAL4-Y9</td>
<td>-</td>
</tr>
<tr>
<td>pDBD-GAL4-p53</td>
<td>AD-GAL4-Y9</td>
<td>-</td>
</tr>
</tbody>
</table>

Sequencing of the complete ORF of the Y9 cDNA revealed a protein of 536 amino acids in which three conserved protein motifs could be predicted using Pfam (www.sanger.ac.uk/software/pfam): a zinc-finger domain of the RING type, a PDZ domain, and a bipartite NLS (Figure 7). The PDZ domain and the RING domain are both common protein-protein interaction modules in eukaryotes.

Figure 7. Graphical depiction of the Y9 protein structure. The conserved domains that are encoded by the 536 amino acid sequence include a PDZ domain, a zinc-finger domain of the RING type, and a bipartite nuclear localization signal (NLS).

A polyclonal serum was raised against a peptide (CKSQHDRHTTSTATP) in the Y9 sequence. This peptide was chosen based on predictions of antigenic properties (Peptide Structure, GCG package, Wisconsin Univ.). The anti-Y9 antibody was found by Western blot to recognize a protein in nuclear extracts of C. tentans cells that migrates at the expected size of Y9 (60 kDa) in SDS-PAGE (Figure 8, lane 1). This allowed us to confirm the interaction between Hrp65 and Y9 that had been observed in the two-hybrid system by immunoprecipitation experiments.
Figure 8. Detection of Y9 in protein extracts and immunoprecipitates. Lane 1: Nuclear extracts of C. tentans tissue culture cells were prepared as previously described (Paper I), separated on 10% SDS-PAGE and blotted onto a nitrocellulose membrane. The membrane was probed with polyclonal serum directed against a peptide in the Y9 amino acid sequence. The secondary antibodies were conjugated to alkaline phosphatase and developed with NBT/BCIP according to standard procedures. The solid arrow indicates migration of the Y9 protein at 60 kDa, the open arrowhead indicates the presence of a more slowly migrating protein that cross-reacts with the serum. The band marked by an asterisk most likely results from penetration of rabbit antibodies from the adjacent lane. Lanes 2-4: Protein complexes were immunoprecipitated from nuclear extracts of C. tentans tissue culture cells using either a polyclonal anti-Hrp59 antibody (lane 2), the 4E9 monoclonal antibody against Hrp65 (lane 3), or in the absence of antibodies (lane 3), as previously described (Paper I). The immunoprecipitated complexes were analyzed for the presence of Y9 as described above. The asterisk indicates cross-reactivity of the secondary antibody with the immunoglobulin heavy chain of the anti-Hrp59 antibody.

Antibodies against Hrp65 can precipitate Y9 from C. tentans nuclear extracts (Figure 8, lane 3), demonstrating that the two proteins are associated in the same protein complex(es) in vivo. We have also found that Y9 associates with Hrp59 (Figure 8, lane 2). Given that Hrp59 interacts directly with Hrp65 (Paper I), it is possible that Hrp65, Hrp59 and Y9 are associated in the same complex. We found the Y9-Hrp59 interaction to be completely resistant to RNase digestion by treating the immunoprecipitated protein complexes with RNase A, while this treatment partially disrupted the association of Y9 and Hrp65 (data not shown). This suggests that Y9 interacts with Hrp59 directly, while the interaction between Y9 and Hrp65 occurs by a combination of protein-RNA and protein-protein interaction. Further analysis is required to characterize the protein-RNA complexes containing Hrp65, Hrp59 and Y9, and to investigate the possibility that these proteins form a ternary complex.
The presence of two conserved domains in Y9 provides certain clues about the properties and function of the protein. Interestingly, both the PDZ domain and the RING domain are protein-protein interaction domains that play a role in the assembly of large multimolecular complexes, often by means of mediating protein oligomerization (reviewed by Borden et al. 2000, Harris and Lim 2001). Typically, PDZ domains are involved in the clustering of signaling molecules and play an important role in organizing protein networks on membranes (reviewed by Ponting et al. 1997, Ranganathan and Ross 1997). The RING domain (also referred to as the C3HC4 zinc-finger), a conserved cysteine-rich domain that binds two atoms of zinc, is present in many eukaryotic and viral proteins. Remarkably, several nuclear RING domain proteins, including the promyolytic leukemia (PML) protein and the breast cancer susceptibility gene product 1 (BRCA1), form subnuclear structures in a manner that depends on the integrity of the RING domain (Melnick and Licht 1999, Chen et al. 1995). Moreover, RING domains from various proteins self-assemble in vitro into spherical structures that can act as scaffolds for multimolecular assemblies (Kentsis et al. 2002). Although it is tempting to consider a role for Y9 in the assembly and integrity of CFs and FGCs in C. tentans salivary gland cells, conceivably in complex with Hrp65/Hrp59, the presence of Y9 in these structures has not been demonstrated.

Ubiquitination is a post-translational modification that targets proteins for degradation and is an effective regulator of cellular protein function. A large number of proteins containing a RING domain are known to act as E3 ubiquitin protein ligases (reviewed by Freemont 2000, Jackson et al. 2000, Joazeiro and Weissman 2000), enzymes that function in targeting specific proteins for ubiquitination. Database searches revealed that the closest human homologue of Y9 is SIAH1, a RING-domain protein belonging to the Seven-in-absentia (SINA) family that constitutes a group of proteins that are conserved from plants to mammals. The human SIAH-proteins are components of E3-ubiquitin ligase complexes that catalyze ubiquitination of a number of proteins with important cell regulatory functions. The SINA protein in Drosophila is involved in R7 photoreceptor cell differentiation in the developing eye, by a well-characterized mechanism that involves ubiquitin/proteasome-dependent degradation of Tramtrack, a negative regulator of neuronal differentiation (Carthew and Rubin 1990, Tang et al. 1997, Li et al. 1997). However,
there are three uncharacterized RING-domain proteins in *Drosophila* that are more similar to Y9 than SINA is. The *Drosophila* sequence with highest sequence similarity to Y9 (annotated with accession number CG6688) is 27% identical to Y9, and it has a domain organization similar to that of Y9, with a central RING domain and an N-terminal PDZ domain. Even if Y9 does not appear to be the closest homologue of the *Drosophila* SINA protein, the presence of a RING domain in Y9 and its similarity to other SINA-like proteins suggest that Y9 is an E3-ubiquitin ligase.

**Ongoing projects and future perspectives**

*Investigation of CFs and FGCs*

We have identified two novel proteins located in CFs: Hrp59 and HEL. These proteins appear to have a broader nucleoplasmic distribution than Hrp65, as they are also present on the BR particle itself and in larger structures that we believe are FGCs. Some putative components of CFs and FGCs must also be investigated, including p2D10 (Sabri and Visa 2002) and Y9. Moreover, our finding that at least two isoforms of Hrp65, Hrp65-1 and Hrp65-2, are present in the nucleus raises the question whether both of them are present in CFs. So far, our ability to describe the localization of nucleoplasmic components in detail has been constrained by limitations in the conventional IEM method. In particular, the FGCs cannot be clearly defined by conventional TEM, as it is difficult to distinguish in the cryosections structures that are in contact with BR particles from separate structures that overlap with BR particles. Therefore, we have recently started to analyze immunolabeled cryosections in 3D using ET. With this method, we hope to be able to investigate the composition of CFs and FGCs in further detail, and to establish the presence of different factors on BR particles, CFs and FGCs, respectively. For some components, this analysis will also require the production of antibodies that are useful for IEM.

Another application for the immuno-ET method will be to compare the CFs and FGCs, which have been described in the polytene salivary gland cells of *C. tentans*, with nucleoplasmic structures that exist in diploid cells. We have previously suggested that the IGCs are diploid.
equivalents of FGCs (Kiesler and Visa 2004). An initial step to investigate this possibility would be to compare the structure and the protein composition of FGCs in *C. tentans* salivary gland cells with those of IGCs in mammalian cells.

*Functional studies on Hrp65/NonA and Hrp59*

On account of the high similarity between Hrp65 and its *Drosophila* homologue NonA, the two proteins are probably functionally equivalent in many aspects. Several projects to study the function of NonA are now in progress. For instance, an investigation of pre-mRNAs binding to NonA was recently initiated, in which an RNA co-immunoprecipitation approach is applied that is similar to the one that has been used for the isolation of Hrp59-substrates (Paper I). This analysis will be repeated after silencing of NonA using RNA interference (RNAi) to investigate whether binding of NonA has any effect on the expression of these pre-mRNAs.

It would also be interesting to investigate whether Hrp65 oligomerization contributes to the formation of CFs in vivo. It may be feasible to interfere with Hrp65 oligomerization in *C. tentans* salivary gland cells by microinjecting a peptide or an antibody that binds to the PBD of Hrp65.

We are planning to investigate whether Hrp59 has any effect on splicing in vitro, using nuclear extracts that are either depleted of endogenous Hrp59, or enriched with recombinant Hrp59, in in vitro splicing assays. Silencing the expression of Hrp59 in *Drosophila* cells by RNAi would make it possible to investigate whether Hrp59 affects the expression of certain genes, and whether Hrp59 has any effect on the splicing of its natural substrates in vivo. Furthermore, it would be interesting to examine whether Hrp59 associates with its RNA substrates by binding to functional ESEs, and whether this binding has any effect on the assembly of spliceosomal components.
Concluding remarks

In a search for Hrp65-interacting proteins using the yeast two-hybrid screening system, we have identified four different proteins and we have shown that these proteins interact with Hrp65 in vivo. Hrp65 appears to associate with a variety of factors that are involved in different gene expression processes. One of these factors, the Hrp59 protein, associates with pre-mRNAs that contain an ESE sequence, suggesting that it may play a role in regulating pre-mRNA splicing events. This is in line with the proposal that hnRNP M, a human homologue of Hrp59, is an inhibitor of splicing. Another Hrp65-associated protein is the putative RNA helicase HEL, also referred to as UAP56 in mammals and Sub2 in yeast, which is involved in pre-mRNA splicing and in mRNA export. The function of the protein Y9 is currently unknown, but it is likely that it modifies proteins by ubiquitination, since it is similar to human SIAH1 and has a RING domain in its sequence. Over the last couple of years, it has become clear that the function of the mammalian homologues of Hrp65 is remarkably complex, as these proteins participate in numerous nuclear processes and associate with a large variety of factors. Accordingly, Hrp65 and its homologues can be considered to be “multi-task” factors in the nucleus. The discovery that Hrp65 can self-associate through a domain that is evolutionarily conserved, and that this domain also mediates heterodimerization between human homologues of Hrp65, is interesting in this respect as oligomerization may represent a key feature which makes these proteins suitable for mediating their various functions.

It is also interesting to note that Hrp65, Hrp59 and HEL, all of which are proteins that bind to pre-mRNAs cotranscriptionally and function in various aspects of mRNA biogenesis, are also present in CFs and/or FGCs. In recent years, nuclear structures have been described as highly dynamic assemblies that are formed and maintained by transient interactions of their protein components. In view of this concept, the numerous factors that mediate different nuclear activities may be self-sufficient in organizing themselves into larger structures. A future challenge will be to understand how the intricate and dynamic interplay of factors in nucleoplasmic complexes can regulate nuclear processes.
Acknowledgements

This study was carried out at the Department of Molecular Biology and Functional Genomics, Stockholm University. I would like to thank all of those who have contributed to the formation of this thesis, and especially:

My supervisor Neus Visa, for five years of excellent guidance, for sharing of knowledge and experience as well as of thoughts, ideas, and an enthusiasm for science that is severely infectious.

My co-authors and co-workers Anki Östlund-Farrants, Christos Samakovlis, Cisco Miralles, Lars-Göran Öfverstedt, Manuela Hase, and Ulf Skoglund for good collaborations, help and advice.

Mona Tibäck and Inger Granell for skilled technical assistance in the experimental work.

George Farrants for language editing; Manuela Hase and Soumya Chakravarty for critical reading of the thesis.

All the members of the lab, former and present. A special thanks to Cisco, Nafiseh and Petra for invaluable help and encouragement, and to my room-mates Emilia, Mona and Victoria for your kindness and patience.

My colleagues at the Department of Molecular Biology and Functional Genomics for your friendliness, and for providing a good atmosphere (research-wise and otherwise).

All my dear friends, especially Anna A, Anna K, Emilia, Emma, Janne, Josefin, Manuela, Monika, Peter, and Petra, for your genuine support and for your beautiful friendship.

Soumya, for your enduring love and glowing hopefulness, and for always being so close.

My wonderful parents for your endless encouragement and love, and for always being there for me.
References


