The PHD-finger protein 5 is a part of the spliceosome and acts as a DNA binding protein.

Dissertation

zur Erlangung des Doktorgrades
der Mathematisch-Naturwissenschaftlichen Fakultäten
der Georg-August-Universität zu Göttingen
vorgelegt von

Tomasz Marek Rzymski
aus Torun, Polen
Göttingen 2004

D7
Referent: Prof. Dr. W. Engel
Korreferentin: PD Dr. S. Hoyer-Fender
Tag der mündlichen Prüfungen:

CONTENTS

1 Introduction

1.1 Identification of the murine PHF5a

PHF5a (PHD-finger protein 5) was firstly identified among genes differentially expressed between the spermatogonia-derived cell line GC-1spg and the spermatocyte derived cell line GC-4spc (Tascou et al., 2001). Using a suppression subtractive hybridization (SSH) the complete cDNA (859bp in length) of mouse PHF5a was isolated. This differential display was performed in order to identify genes responsible for the malignant transformation of male germ cells and invasiveness of GC-1spg cells. The spermatogonia-derived GC-1spg cells were shown to have a significantly higher invasive activity and a higher potential to malignant transformation as compared to GC-4spc cells. On gross inspection GC-1spg cells display characteristics of germ cells at a stage between spermatogonia type B and primary spermatocytes (Hofmann et al., 1992). They represent a valuable model to study the origin of testicular germ cell tumors (TGCT) as they can be considered as precancerous. Therefore, genes overexpressed in GC-1spg are putative candidates involved in TGCT formation. A second immortalized germ cell line GC-4spc has been characterized as between the preleptotene spermatocyte and the early pachytene spermatocyte stage (Tascou et al., 2000).

1.2 PHF5a belongs to a multigene family containing PHD-fingers

A novel murine multigene family containing a putative PHD-finger (plant homeodomain finger) was described by Trappe et al. (2002). It was demonstrated that the active murine gene PHF5a, which is localized on chromosome 15E encodes a small evolutionary conserved protein of 110 amino acids (Figure 1). The syntenic region on human

chromosome 22q13.2 contains the corresponding human PHF5a gene. Interestingly, all other PHF5a related sequences which can be found in the murine genome display features of processed pseudogenes, which are generated by transposition in the genome via an RNA intermediate (Mighell et al., 2000; Rogers, 1985; Vanin, 1985).

The PHF5a protein has a predicted molecular weight of 12.4 kDa and its theoretical p*I* lies at 8.413. The N- and C- terminal parts of PHF5a are rich in basic amino acids and the p*I* for amino acids 1-21 and 85-110 lies at 10.2. A putative polyadenylation signal AATAAA is located at nucleotide position 833-838. By using the computer program PredictNLS a nuclear localization signal was proposed at the C-terminus of PHF5a (Figure 1). Moreover, by using a PHF5a-GFP fusion protein strong fluorescence was observed in the nuclear matrix of NIH3T3 cells while a deletion of 20 amino acids at the C-terminus of PHF5a resulted in a cytoplasmic distribution. Therefore, the deleted amino acid sequence of PHF5a (ERKKYGFKKR) may represent a possible nuclear localization signal (NLS). In addition, the conserved cysteine and histidine residues in the central part of PHF5a were proposed to form an imperfect PHD-finger domain as it did not display all characteristic features of this motif (Aasland et al., 1995). I*n vitro* PHD fingers are able to bind to DNA although Lyngso et al. (2000) showed that the PHD finger of the transcription factor SPBP is involved in chromatin-mediated transcriptional regulation acting as a domain of protein-protein interaction. PHD fingers are commonly found in transcriptional activators, repressors and cofactors and in proteins involved in chromatin modulation (Saha et al.,1992; Aasland et al., 1995; Moosmann et al., 1996). Therefore, it was hypothesized that PHF5a could act as a chromatin-associated protein (Trappe et al., 2002).

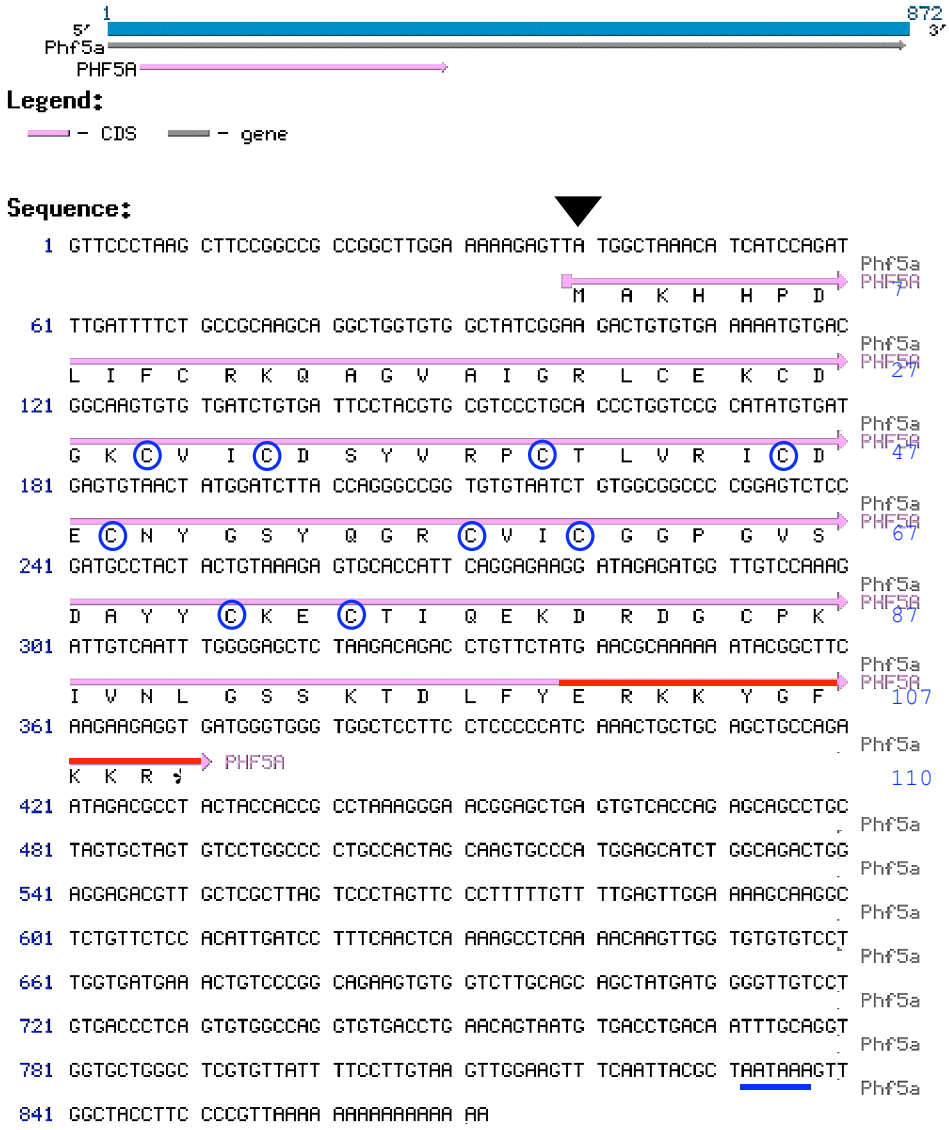


Figure 1. Nucleotide and deduced amino acid sequences of murine PHF5a. The nucleotide sequence representing the ORF is underlined in pink and the putative polyadenylation signal AATAAA is underlined in blue. The translation initiation codon ATG is marked by an arrow and the stop codon TGA is marked by an asterisk. The putative nuclear localization signal is underlined in red. Conserved cysteine residues of the PHD-finger are encircled in blue. Nucleotide residues numbers are shown at the left of each line, amino acid residue numbers are shown at the right of each line.

1.3 PHF5A is highly conserved during eukaryotic evolution

One of the most interesting features of PHF5a is the level of evolutionary conservation. PHF5a appears to be one of the most highly conserved proteins known to date.

At the protein level the PHF5a sequence is 100% identical in all vertebrates studied (Figure 2). If the comparison is limited to multicellular organisms, which includes rat, mouse, human, zebra fish, the fruit fly *Drosophila melanogaster*, the mosquito *Anopheles gambiae*, the nematode *C. elegans*, and the plant *Arabidopsis thaliana*, the degree of sequence identity is over 80%. Even the most distant PHF5a ortholog found in the budding yeast *Saccharomyces cerevisiae* is 55% identical to the vertebrate PHF5a. This suggests a very high selective pressure to conserve PHF5a sequences during evolution. Interestingly, as it was suggested by Oltra et al. (2003) that the PHF5a sequence was conserved without even a single amino acid change for a period of more than 530 million years, the date of the oldest fish fossil found. Furthermore, to date no ortholog of PHF5a could be found in the genomes of prokaryotic organisms, therefore, it is possible that its presence and function is restricted to eukaryotic organisms. The level of evolutionary conservation indicates that PHF5a may contribute to basic cellular function (Stanchi et al., 2001).

```
M.musculus gi:19263793     MAKHHPDLIFCRKQAGVAIGRLCEKCDGKCVICDSYVRPCTLVRICDECNYGSYQGRCVI 60
H.sapiens gi:46576678      MAKHHPDLIFCRKQAGVAIGRLCEKCDGKCVICDSYVRPCTLVRICDECNYGSYQGRCVI 60
R.norvegicus gi:20302071   MAKHHPDLIFCRKQAGVAIGRLCEKCDGKCVICDSYVRPCTLVRICDECNYGSYQGRCVI 60
Danio rerio gi:20977587    MAKHHPDLIFCRKQAGVAIGRLCEKCDGKCVICDSYVRPCTLVRICDECNYGSYQGRCVI 60
A.gambie gi:21301692       MAKHHPDLIFCRKQPGVAIGRLCEKCDGKCVICDSYVRPCTLVRICDECNYGSYQGRCVI 60
D.melanogaster gi:22945771 MAKHHPDLIFCRKQPGVAIGRLCEKDDGKCVICDSYVRPCTLVRICDECNYGSYQGRCVI 60
A.thaliana gi:3420051      MAKHHPDLIMCRKQPGIAIGRLCEKCDGKCVICDSYVRPCTLVRICDECNYGSFQGRCVI 60
C.elegans gi:32564382      MAKHHPDLIFCRKQPGIAIGRLCEKCDGRCVICDSHVRPCTLVRICEECNYGSYQGRCVI 60
Ssacch.pombe gi:6073754    MSKHHPDLVLCRRQPGITVGKLCERCDEKCPICDSHVRPTTLVRICDECAFGSSQDRCII 60
S.cerevisae gi:6325351     MSRHQFDLIMCLKQPGVQTGLLCEKCDGKCPICDSYVRPKRKVRVCENCSFGKQAKNCII 60
                           *::*: **::* :*.*:  * ***: * :* ****:***   **:*::* :*.  .*:*

Mus musculus               CGG-PGVSDAYYCKECTIQEKDRDGCPKIVNLGSSKTDLFYERKKYGFKKR------- 110
Homo sapiens               CGG-PGVSDAYYCKECTIQEKDRDGCPKIVNLGSSKTDLFYERKKYGFKKR------- 110
Rattus norvegicus          CGG-PGVSDAYYCKECTIQEKDRDGCPKIVNLGSSKTDLFYERKKYGFKKR------- 110
Danio rerio                CGG-PGVSDAYYCKECTIQEKDRDGCPKIVNLGSSKTDLFYERKKYGFKKR------- 110
Anopheles gambie           CGG-PGVSDAYYCKECTIQEKDRDGCPKIVNLGSSKTDLFYERKKYGFKR-------- 109
Drosophila melanogaster    CGG-PGVSDAYYCKSCTIQEKDRDGCPKIVNLGSSKTDLFYERKKYGFKQNY------ 111
Arabidopsis thaliana       CGG-VGISDAYYCKECTQQEKDRDGCPKIVNLGSAKTDLFYERKKYGFKKR------- 110
Caenorhabditis elegans     CGG-AGVSDAYYCKECTILEKDRDGCPKIVNLGSAKTDLFYERKKFGSKKS------- 110
Ssacch. pombe              CGA-PGVSDCYYCSECTRMEYDRDGCPRVINLGSSRTDWFYERKKFKNAGKEMPGATY 117
Saccharomyces cerevisae    CNLNVGVNDAFYCWECCRLGKDKDGCPRILNLGSNRLDRHFEKKKKV---------- 107
                           *.   *:.*.:** .*     *:****:::**** : * .:*:**
```

Figure 2. Multiple amino acid sequence alignment of murine PHF5a with homologous proteins. Multisequence alignment was performed using the ClustalW program (http://www.ebi.ac.uk/clustalw). The coloring of amino acid residues takes place according to the following physicochemical criteria: red- small and hydrophobic including aromatic Y; blue- acidic; magenta- basic; green- Hydroxyl + Amine + Basic – Q; gray- others. An alignment displays the following symbols denoting the degree of conservation observed in each column: "*" indicates that the residues or nucleotides in that column are identical in all sequences in the alignment; ":" indicates that conserved substitutions have been observed; "." indicates that semi-conserved substitutions are observed.

1.4 PHF5a is ubiquitously expressed in murine tissues and takes part in embryonic development of *C. elegans*

Previously, it was shown by using the RNase protection assay (RPA) that the active murine locus, PHF5a, is expressed ubiquitously in tissues of prenatal and postnatal mice. Furthermore, it was also demonstrated that PHF5a is strongly expressed in the spermatogonia-derived cell line GC-1spg, the spermatocyte-derived cell line GC-4spc, and in the Leydig cell-derived cell line MA-10 while it is only weakly expressed in the Sertoli cell-derived cell line 15P-1 (Trappe et al., 2002).

In addition, the expression of the PHF5a ortholog in *C. elegans* was investigated (Trappe et al., 2002). It is noteworthy, that in contrast to the murine orthologous gene PHF5a expression in *C. elegans* is time–dependent and tissue-specific. PHF5a expression starts within the morphogenetic phase of embryonic development, lasts to the stage of adult worms and is restricted to the developing pharynx, body wall muscular structures and to anal muscles. These results are inconsistent with the ubiquitous pattern of PHF5a expression in mouse tissues. Moreover, application of the RNAi depletion technology resulted in complete embryonic lethality in the early morphogenetic phase of development in *C. elegans* offspring while it had no effect on adult worms. These results provide valuable evidence that the PHF5a protein is essential in the morphogenesis of *C. elegans*.

1.5 The PHF5a protein is involved in transcriptional regulation

In the light of a recent report, PHF5a is also involved in modulating the response of the connexin 43 gene to estrogen (Oltra et al., 2003). Connexin 43 is a member of a family of transmembrane proteins forming cell-to-cell channels. An uterine cDNA expression library was screened for proteins that interact directly with the connexin 43 gene promoter. By using this screen the rat PHF5a orthologous protein named Ini was identified. Ini binds to a 38-nucleotide region of the connexin 43 promoter (Oltra et al., 2003; Oltra and Werner, 1998) and this sequence had been previously shown to function as a *cis*-activator in the transcription of the connexin 43 gene (Chen et al., 1995). The protein Ini is able to bind to the promoter of connexin 43 leading to an increase of connexin 43 mRNA levels in the myometrium after estrogen treatment. However, it is worth noting that PHF5a is expressed in several tissues which are not responsive to estrogen. Therefore, this expression pattern of the PHF5a gene indicates that PHF5a functions as a more general transcription cofactor.

1.6 PHF5a is implicated in pre-mRNA processing

The PHF5a orthologous gene ini1 in *Schizosaccharomyces pombe* was also investigated by gene disruption experiments (Oltra et al., 2003). On the one hand, deletions of ini1 resulted in cell death and in a cell cycle block. Previous reports indicate that PHF5a is capable of binding to DNA, it was discussed that it might be the result of disturbances in expression of some critical factors required for the G2 to M phase transition. On the other hand, ini1 cell mutants displayed also pre-mRNA processing defects and an accumulation of pre-mRNA was observed for six of the seven genes tested. The latter results strongly support a role for ini1 in general mRNA splicing processes.

Moreover, The PHF5a ortholog Rds3p of *Saccharomyces cerevisiae* was identified as a critical pre-mRNA splicing factor and an integral component of the yeast spliceosome (Wang and Rymond, 2003). Yeast extracts lacking Rds3p activity are splicing defective and exhibit an arrest in spliceosome assembly prior to stable U2 snRNP recruitment. In addition, it was shown that Rds3p interacts with at least five U2 snRNP proteins present in the SF3b subcomplex, namely Cus1p, Hsh155p, Hsh49p, Its3p/Snu17p, and Rse1p and with the Yra1p RNA export protein (Wang and Rymond, 2003).

In addition, the human PHF5a protein was also identified by mass spectrometry as a novel protein associated with the 17S U2 snRNP and one of its stable subunits, SF3b (Will et al., 2002). In conclusion, the PHF5a protein displays hallmarks of transcription-splicing factors (Zolotukhin et al., 2003; Kameoka et al., 2004).

1.7 Aims of the study

A. Several lines of evidences are providing new insights into a putative role of the PHF5a protein as a crucial part of the spliceosome (Will at al., 2003, Wang and Rymond, 2003, Oltra et al., 2003). PHF5a protein interaction studies in yeast cells showed that PHF5a binds to U2 snRNP proteins, however, pre-mRNA processing in mammals differs in many aspects. Therefore, we employed the yeast two-hybrid library screening of murine 11.5-days embryo library to reveal putative PHF5a interaction partners. In addition, isolation of PHF5a in association with proteins involved in chromatin activation may explain its dual function as a transcription-splicing factor.

The topics of the present study were:

- To further characterize PHF5a interacting partners and determine minimal fragments sufficient to form stable complexes with the PHF5a protein.

- To determine which PHF5a domains are responsible for the association with a set of putative interacting partners.
- To verify PHF5a yeast two-hybrid results with *in vitro* protein binding studies by using coimmunoprecipitation experiments.
- To determine if specific PHF5a protein interactions are detectable *in vivo* in mammalian cells.
- To further analyze the subcellular localization of the PHF5a protein and its potential colocalization with interacting proteins.
- To investigate if the PHF5a protein is dispensable for pre-mRNA splicing in mammalian cells.

B. Because PHF5a was initially described as a protein differentially expressed between spermatogonia-derived GC-1spg cells and spermatocyte-derived GC-4spc cells detailed PHF5a expression analyses during postnatal testicular development were employed. The following experiments were performed:

- Western blot analyses of testicular protein extracts from different days in postnatal development and mouse mutants with an arrest of spermatogenesis at specific stages.
- Immunohistochemical expression studies of PHF5a on testis sections and testicular cell suspensions.
- Time and spatial coexpression studies of the PHF5a protein and putative interacting partners.
- In addition, analysis of PHF5a expression during mouse embryonic development was investigated by using the whole mount in situ hybridization technique.

C. PHF5a was shown to be responsible for a cell cycle block in yeast. Therefore, we examined the proliferation rate of murine fibroblast cells in the absence of PHF5a.

D. Finally, DNA-binding studies were employed to obtain conclusive evidences that the PHF5a protein is able to recognize specific DNA motifs. The following studies were performed:

- *In vitro* genomic DNA-pull down assay by using the recombinant PHF5a protein,
- Systematic Evolution of Ligands by Exponential Enrichment (SELEX) and Whole-genome PCR (WG-PCR) to determine PHF5a DNA-binding consensus sequences.

2 Materials and methods

2.1 Chemicals and reagents

Chemicals which are not included in the list below were purchased from the companies Roth) and Merck, respectively.

Agar	Difco
Agarose	Invitrogen
Alkaline phosphatase	Boehringer
Ammonium acetate	Fluka
Ampicillin	Sigma
Ampuwa	Fresenius AG
Aprotinin	Sigma
Bacto-Tryptone	Difco
Chloroform	Baker
Dextran sulfate	Amersham
Diethylpyrocarbonate (DEPC)	Sigma
Dimethylsulfoxid (DMSO)	Merck
dNTPs (100 mM)	Boehringer
Dye Terminator Mix	Applied Biosystems; Amersham
Developer	Kodak

Ethanol	Baker
Ethidium bromide	Sigma
Ficoll 400	Amersham
FKS	Invitrogen
Formaldehyde	Invitrogen
Formamide	Fluka
Glycerol	Gibco
IPTG	Biomol
Isoamyl alcohol	Fluka
Klenow-DNA-Polymerase	Amersham
Kb Ladder	Invitrtogen
Leupetin	Sigma
Salmon sperm DNA	Sigma
ß-Mercaptoethanol	Serva
Orange-G	Sigma
PBS	Invitrogen
Phenol	Invitrogen
PMSF	Sigma
Proteinase K	Boehringer
Radiochemicals: [_ ^{32}P]-dCTP, [_^{35}S]-methionine	Amersham
Restriction enzymes	Invitrogen
Reverse Transcriptase	Invitrogen

RNase A	Invitrogen
RNase Inhibitor	Boehringer
RNA Ladder	Invitrogen
Sodium Dodecyl Sulfate (SDS)	Serva
T_4-DNA-Ligase	Boehringer
T_4-DNA-Polymerase	Boehringer
Taq-DNA-Polymerase	Amersham
Tris	Sigma
Tween 20	Fluka
Vecta Shield with DAPI	Vecta
X-Gal	Biomol

2.2. Solutions and buffers

All standard buffers and solutions were prepared according to Sambrook et al. (1989).

Coomassie Solution:	30% (v/v) Methanol
	10% (v/v) Acetic Acid
	0.5% (w/v) Coomassie Brilliant Blue R 250
Denaturing Solution:	1.5 M NaCl
	0.5 M NaOH
Denhardt´s Solution (50x)	1% BSA

	1% Polyvinylpyrrolidon
	1% Ficoll 400
	in 20 x SSC (pH 7.0)
dNTP-Mix (25 mM)	100 mM dATP
	100 mM dGTP
	100 mM dCTP
	100 mM dTTP
Glycin buffer	1.44% (w/v) Glycin
	0.3% (w/v) Tris
	0.1% (w/v) SDS
Hybridization Solution	5 x SSC
	5 x Denhardt´s Solution
	10% Dextran sulfate
	0.1% SDS
	Salmon Sperm DNA 100 µg/ml
LB-Agar	10 g Bacto-Trypton
	5 g Yeast extract
	10 g NaCl
	15 g Agar

	Per 1000 ml H_2O
LB-Medium	10 g Bacto-Trypton
	5 g Yeast extract
	10 g NaCl
	Per 1000 ml H_2O
E1 buffer (Mini prep)	50 mM Tris-Cl, pH 8.0
	10 mM EDTA
	100 µg/ml RNase A
E2 buffer (Mini prep)	200 mM NaOH, 1% SDS
E3 buffer (Mini prep)	3.0 M Natrium acetate (pH 5.5)
Lysis buffer for proteins	150 mM NaCl
	10 mM EDTA
	50 mM Tris-HCl, pH 7.6
	1% Triton X-100
	1% Proteinases inhibitors:
	Leupetin (1 µg/ml)
	Aprotinin (1 µg/ml)
	PMSF (1 µg/ml)
Methyl-blue solution	0.5 M NaAc

	0.04% Methyl-blue
MOPS buffer (10x)	50 mM NaAc
	10 mM EDTA
	200 mM MOPS (pH 7.0)
	Neutralization Solution 1.5 M NaCl
	1 M Tris-HCl (pH 7.0)
PBS buffer	130 mM NaCl
	7 mM Na_2HPO_4
	4 mM NaH_2HPO_4
PBT buffer	0.1% Tween 20 in PBS (1x)
Running buffer	25 mM Tris
	152 mM Glycin
	0.1% SDS
SSC (20x)	3 M NaCl
	0.3 M Tri-Natrium citrate
	pH 7.0 with NaOH
Stop Mix I	95% Formamid
	20 mM EDTA

	0.05% Brom-phenol-blue
	0.05% Xcylencyanol
Stop Mix II	15% Ficoll
	200 mM EDTA
	0.1% Orange G
TBE buffer (5x)	225 mM Tris pH 8.3
	225 mM Boric acid
	10 mM EDTA
TE buffer (10x)	10 mM EDTA
	100 mM Tris pH 8.0

2.3 Laboratory materials

The laboratory materials, which are not listed here, were bought from Schütt and Krannich, respectively.

Whatman blotting paper (GB 002, GB 003 and GB 004)	Schleicher and Schüll
Cell culture flask	Greiner
Culture slides	Falcon

Dialysis hoses	Serva
Disposable filter	Minisart NMI Sartorius
Filter paper 0858	Schleicher and Schüll
HPTLC Aluminum folio	Merck
Hybond C	Amersham
Hybond N	Amersham
Petri dishes	Greiner
Pipette tips	Eppendorf
Microcentrifuge tubes	Eppendorf
Transfection flask	Lab-Tek/Nalge, Nunc
X-ray films	Amersham
Superfrost Slides	Menze

2.4 Bacterial strains and media

Competent cells *E. coli* DH5_	(Hanahan, 1983). These cells were used for plasmid transformation.
Competent cells *E.coli* BL-21 DE3	(Novagen). These cells were used for expression of recombinant proteins.
LB-Medium:	1% bactotryptone; 0.5% yeast extract; 0.5 NaCl; pH 7.0 (supplemented with 50 µg/ml ampicillin to maintain selection pressure)

Agar plates:	LB-Medium + 1.5% agar (supplemented with 50 µg/ml ampicillin, 2% X-gal and 1 mM IPTG)

2.5 Yeast strains and media

AH109 (BD Clontech)	MATa, reporter genes: HIS3, ADE2, lacZ, MEL1; transformation markers: trp1, leu2
Y187 (BD Clontech)	MAT_, reporter genes: lacZ, MEL1 ; transformation markers: trp1, leu2

YPD medium:	20 g/L Difco peptone 10 g/L Yeast extract 20 g/L Agar (for plates only) 2% Glucose [Optional] For adenine-supplemented YPD (YPDA), add 15 ml of a 0.2% adenine hemisulphate solution per liter of medium (final concentration is 0.003%)

SD medium	6.7 g/L Yeast nitrogen base without amino acids 20 g Agar/L (for plates only) 100 ml/L of the appropriate sterile 10x dropout solution 2% Glucose

A combination of a minimal SD base and a dropout supplement will produce a synthetic, defined minimal medium lacking one or more specific nutrients. The specific nutrients omitted depend on the desired selection medium.

Nutrient	**10X Concentration**
L-Adenine hemisulfate salt	200 mg/l
L-Arginine HCl	200 mg/l
L-Histidine HCl monohydrate	200 mg/l
L-Isoleucine	300 mg/l
L-Leucine	1000 mg/l
L-Lysine HCl	300 mg/l
L-Methionine	200 mg/l
L-Phenylalanine	500 mg/l
L-Threonine	2000 mg/l
L-Tryptophan	200 mg/l
L-Tyrosine	300 mg/l
L-Uracil	200 mg/l
L-Valine	1500 mg/l

2.6 Eukaryotic cell lines

NIH3T3	Mouse embryonic fibroblast cell line, ATCC, Rockville, USA
GC-1spg	Mouse spermatogonia-specific cell line, ATCC, Rockville, USA
GC-4spc	Mouse spermatocyte-specific was established and described by Tascou and co-workers (2000).

2.7 Plasmids and cDNA fragments

pET41a, pET41b, pET41c	Novagen

pGBKT7	BD Clontech
pGADT7	BD Clonetch
pEGFP- C1	BD Clontech
pDsRed-C1	BD Clontech
pGEM-T Easy	Promega
pTKneo	Prof. Brose, MPI für Experimentelle Medizin, Göttingen
pBluescript SK+	Stratagen
pTrEX-1.1 neo	Novagen
pBridge	BD Clontech

2.8 Synthetic oligonucleotide primers

The synthetic oligonucleotide primers used either for sequencing of constructs or cDNA clones were ordered from the company Qiagen .

Standard sequencing primers:

SP6	5’ – AGG TGA CAC TAT AGA ATA C – 3’
T3	5’ – AAT TAA CCC TCA CTA AAG GG – 3’
T7	5’ –GTA ATA GCA CTC ACT ATA GGG C– 3’
T7 modified	5’ – TAA TAC GAC TCA CTA TAG GGA – 3’

1F	5’ – CTCGAGCTCAAGCTTCGAATTC– 3’ (PHF5a cloned into pEGFP-N1, clone pRT21, *EcoRI*)
2F	5’ – CAGAATTCCAAGTGTGTGATCTGTGATTCC– 3’ **(** PHF5a, gi:19263792, nucleotide position 123-143, *EcoRI***)**

2R	5’ – GTCTGCAGGAATCACAGATCACACACTTG– 3’
	(PHF5a, gi:19263792, nucleotide position 123-145, *PstI***)**
3F	5’ – GACCATGGCCTACTACTGTAAAGAGTGC– 3’
	(PHF5a, gi:19263792, nucleotide position 244-264, *BamHI*)
3R	5’ – CAGAATTCTCCTGAATGGTGCACTCTTTAC– 3’
	(PHF5a, gi:19263792, nucleotide position 254-278, *EcoRI*)
4R	5’ – GTCGACTGCAGAATTCGCCTC– 3’
	(PHF5a, gi:19263792, nucleotide position 366-369, pRT21, *EcoRI*)
2FINGERFP	5’ – GACCATGGGCTATGGATCTTACCAGGGCCG– 3’
	(PHF5a, gi:19263792, nucleotide position 189-209, *BamHI*)
RINGRP	5’ – GTCTGCAGCGGCCCTGGTAAGATCCATAG– 3’
	(PHF5a, gi:19263792, nucleotide position 189-209, *PstI*)
PRT21RP	5’ – CAGAATTCTCAGCCTCTTCTTGAAGCC– 3’
	(PHF5a, gi:19263792, nucleotide position 355-372, pRT21, *EcoRI*)
PRT21FP	5’ – CTCAAGCTTCGAATTCTTGATGGC– 3’
	(PHF5a, gi:19263792, nucleotide position 1-5, pRT21, *EcoRI*)
RS-U2AFP	5’ – GTGAATTCTTGAAGCCCATCTCAAGAGAGCTAC– 3’
	(U2AF1, gi:31340597, nucleotide position 540-624, *EcoRI*)
RS-Srp75FP	5’ – GTGAATTCAAGCCAGGTTCTAGGCGGCGCCGG– 3’
	(SRp75, gi:22268150, nucleotide position 1516-1540, *EcoRI*)
RS-Srp75RP	5’ – GTGGATCCTTAGGACCTTGAGTGGGACCTGG– 3’
	(SRp75, gi:22268150, nucleotide position 2439-2461, *BamHI*)
RS-Srp20FP	5’ – GTGAATTCCTGTCGAATGGTGAAAAGAGAAGTC– 3’
	(SRp20, gi:47940163, nucleotide position 329-353, *EcoRI*)
RS-Srp20RP	5’ – GTGGATCCTATTTCCTTTCATTTGACCTAGATC– 3’
	(SRp20, gi:47940163, nucleotide position 561-586, *EcoRI*)
SR-ASF/SF2FP	5’ – GTGAATTCGCCTACATCCCGCGTATAGTTGATAG– 3’
	(ASF/SF2, gi:28386235, nucleotide position 823-846, *EcoRI*)
SR-ASF/SF2RP	5’ – GTGGATCCTTATGTACGAGAGCGAGATCTGC– 3’
	(ASF/SF2, gi:28386235, nucleotide position 660-684, *BamHI*)
SR-Srp30cFP	5’ – GTGAATTCGGAGGTCGGGGTGGGTGGCCCCGTG– 3’
	(SRp30c, gi:15126568, nucleotide position 405-430, *EcoRI*)

SR-Srp30cRP	5’ – GTGGATCCTCAGTAGGGCCTGAAAGGAGAGAAG– 3’
	(SRp30c, gi:15126568, nucleotide position 777-803, *BamHI*)
U2AF1FP	5’ – ATGGCGGAATACTTGGCCTCCATCTTCG– 3’
	(U2AF1, gi:31340597, nucleotide position 21-48)
U2AF1RP	5’ – TCAAAATCGTCCAGATCTCTCCCGGTC– 3’
	(U2AF1, gi:31340597, nucleotide position 714-740)
globinFP	5’ – GCTCCTCACATTTGCTTCTGACATAGTTGTG– 3’
	(mus musculus beta globin gene exon1)
globinRP	5’ – CCTGAAGTTCTCAGGATCCACATGCAGCTTG– 3’
	(mus musculus beta globin gene exon3)
Y2HAD3	5’ – GTTGAAGTGAACTTGCGGGG– 3’
Y2H2	5’ – AATACCACTACAATGGATG– 3’
SELFP	5’ – CATCGATACAGTTCGATATC– 3’
SELRP	5’ – CTCCTATACTGAGTTCATG– 3’
WGPCR 1	5’ – GCACTAGTGGCCTATGCGG– 3’
WGPCR2	5’ – GTACCTTCGTTGCCGGATC– 3’
MboI linker 2	5’ – GATCCGGCAACGAAGGTACCATGGCCGCATAGGCCACTA GTGC– 3’
MboI linker 1	5’ – GCACTAGTGGCCTATGCGGCCATGGTACCTTCGTTGC CG–3’
MUT5’FP	5’ – GAATTCTTGATGGCTA**T**AC**TC**CATCCAGATTTG– 3’
MUT5’RP	5’ – CAAATCTGGATG**GA**GT**A**TAGCCATCAAGAATTC– 3’
	(PHF5a, gi:19263792, introduced mutations are in bold)
MUT3’FP	5’ – GACCTGTTCTATGAAC**CG**A**T**AAAATACGGCTTC– 3’
MUT3’RP	5’ – GAAGCCGTATTTT**AT****CG**GTTCATAGAACAGGTC– 3’
	(PHF5a, gi:19263792, introduced mutations are in bold)

2.9 Antibodies

Goat anti-rabbit IgG

Alkaline phosphatase conjugate Sigma

Goat anti-rabbit IgG HRP conjugate	Sigma
Rabbit anti-mouse IgG Alkaline phosphatase conjugate	Sigma
Rabbit anti-mouse IgG HRP conjugate	Sigma
Goat anti-rabbit IgG Cy^3 conjugate	Sigma
Rabbit anti-mouse IgG FITC conjugate	Sigma
Anti-_ tubulin	Sigma
Anti-HA	BD Clontech
Anti-c myc	BD Clontech
Anti- PHF5a	Raised against synthetic peptides in rabbit (Eurogentec)
Anti- $U2AF^{35}$	Provided by Prof. Tom Maniatis,
Anti- $U2AF^{65}$	Provided by Prof. Tom Maniatis (Department of Molecular and Cellular

	Biology, Harvard University, 7 Divinity Avenue, Cambridge, MA, USA)
Anti-HA	BD Clonetch
Anti-c-Myc	BD Clonetch

2.10 Isolation of nucleic acids

2.10.1 Small-scale preparation of plasmid DNA (Mini prep)

Plasmid DNA was extracted from bacterial cultures according to the alkaline lysis procedure which is described in Molecular Cloning: Laboratory Manual (Maniatis at al., 1989).

2.10.2 Large-scale preparation of plasmid DNA (Midi prep)

Ultra pure supercoiled plasmid DNA with high yield was prepared using the Qiagen Plasmid Midi Kit (Qiagen) according to the original instruction manual: Qiagen Plasmid Purification Handbook (2001). The isolated plasmid DNA was further used for transfection, sequencing, restriction analysis and subcloning.

2.10.3 Isolation of total RNA from eukaryotic cells

High quality total RNA from mammalian cells was extracted using the RNeasy Mini Kit (Qiagen), according to the original instruction manual.

2.10.4 Determination of nucleic acid concentration

The concentration of nucleic acids was determined photometrically by measuring the absorption of the samples at 260 nm (Spectrophotometer Ultrospec 300pro, Amersham). DNA and RNA quality, i.e. contamination with salt and protein was checked by measurements at 230, 280, and 320 nm. The concentration can be calculated according to the formula: $C = (E\ 260 - E\ 320) \times f \times c$

C = concentration of sample (µg/µl)

E 260 = absorption at 260 nm

E 320 = absorption at 320 nm

f = dilution factor

c = concentration (standard) / absorption (standard)

For double stranded DNA : c = 0.05 µg/µl

For single stranded DNA : c = 0.03 µg/µl

For RNA : c = 0.04 µg/µl

2.11 Cloning techniques

2.11.1 Restriction analysis of DNA

Restriction enzyme digestions were performed by incubating double-stranded DNA molecules with an appropriate amount of restriction enzyme in its respective buffer as recommended by the supplier, and at the optimal temperature for the specific enzyme. These reactions were usually incubated for 1-3 hours or over night to insure complete digestion at the optimal temperature for enzyme activity which was typically 37°C.

2.11.2 Isolation of DNA fragments from agarose gels

For the isolation of DNA fragments from agarose gels the QIAEX-II kit (Qiagen) was employed. After separation of DNA on an agarose gel, the DNA band which had to be isolated was excised with a razor blade, weighed and treated as described in the user manual. Isolated DNA fragments were checked on agarose gels and used for subcloning or as a probe for Northern blot experiments.

2.11.3 Dephosphorylation of 5’ ends of DNA

To prevent the recircularization of plasmids without insertion of foreign DNA, alkaline phosphatase treatment was performed. Alkaline phosphatase catalyses the hydrolysis of 5'-phosphate residues from DNA. The following items were mixed: 1-5 µg vector DNA, 5 µl 10x reaction buffer, 1 µl alkaline phosphatase (1U) in a total volume of 50 µl and incubated at 37°C for 30 min. Then the reaction was stopped by heating at 85°C for 15 min. The dephosphorylated DNA was purified by phenol/ chloroform extraction and ethanol precipitation.

2.11.4 Filling-up reaction of DNA ends

0.1-4 µg of digested DNA was mixed with 0.05 mM dNTPs and 1-5 U of Klenow fragment. The reaction was incubated at 37°C for 10 min, then stopped by heating at 75°C for 10 min. Subsequently, DNA was purified by phenol/ chloroform extraction and ethanol precipitation.

2.11.5 Ligation of DNA fragments

The ligation of an insert into a vector was carried out in the following reaction mix: 25-50 ng vector DNA, 50-100 ng insert DNA, 1µl ligation buffer (10x), 1 µl T4 DNA ligase (5 U/µl) in a total volume of 10 µl. Blunt-end ligations were carried out at 16°C overnight, whereas sticky-end ligations were carried out at room temperature for 2-4 hours.

2.11.6 Subcloning of PCR and RT-PCR products

Taq and other polymerases have a terminal transferase activity which results in the nontemplated addition of a single nucleotide to the 3'-ends of PCR products. In the presence of all four dNTPs, dATP is preferentially added. This terminal transferase activity is the basis of TA-cloning strategy. For the subcloning of PCR or RT-PCR products, the pGEM-T Easy Vector system which has 5' T overhangs was employed. The following items were mixed;

50 ng of pGEM-T or pGEM-T Easy Vector

PCR product (3:1, insert: vector ratio)

1 µl T4 DNA Ligase 10x buffer

1 µl T4 DNA Ligase

in a total volume of 10 µl

The contents were mixed and the reaction was incubated overnight at 16°C.

2.11.7 Transformation of competent cells with plasmid DNA

(Hanahan, 1983)

Transformation of bacteria was done by gently mixing one aliquot of competent bacteria cells (100 µl) with 10 µl of the ligation reaction. After incubation for 30 min on ice, bacteria were heat shocked for 45 sec at 42°C, cooled down for 2 min on ice. After adding 450 µl of LB medium, bacteria cells were incubated at 37°C in a shaker with a speed of 160 rpm for 1 hour to allow recovery of heat shocked bacteria and were plated out on LB-agar plates containing appropriate antibiotic (50 µg/ml) and whenever required, 1 mM IPTG and X-Gal 40 mg/ml.

2.12 Gel electrophoresis and blotting techniques

2.12.1 Agarose gel electrophoresis of DNA and Southern blot analysis

(Southern, 1975)

DNA fragments were separated on an agarose 0.5x TBE gel, denatured, transferred and immobilized onto a nylon Hybond C membrane (Amersham) according to the standard protocol.

2.12.2 Agarose gel electrophoresis of RNA and Northern blot analysis

(Maniatis and Efstratiadis, 1980)

RNA samples were resolved on a denaturing agarose MOPS gel, transferred and immobilized onto a nylon membrane according to the original protocol.

2.12.3 DNA and RNA molecular weight ladders

To determine the size of nucleic acids fragments on agarose gels, molecular weight ladders were loaded in parallel.

1 kb DNA Ladder	Invitrogen
100 bp DNA Ladder	Invitrogen
0.24-9.5 RNA Ladder	Invitrogen

2.13 Sequencing

The non-radioactive sequencing was achieved with the Dye Terminator Cycle Sequencing-Kit (Amersham). The principle of this procedure is based on the method described by Sanger et al. (1992). The sequencing reaction was carried out in a total volume of 10 µl containing 1 µg plasmid DNA or 100-200 ng purified PCR product, 10 pmol primer and 4 µl reaction mix (containing dNTPs, dideoxy dye terminators and *Taq* DNA

polymerase). Elongation and chain termination take place during the following program in a thermocycler:

5 min denaturation followed by 25 cycles: 95°C, 30 sec, denaturing; 55°C, 15 sec, annealing; 70°C, 4 min, elongation. Samples were purified, loaded on the sequence gel and analyzed by the Sequence Unit at the Institute of Human Genetics, Goettingen.

2.14 One-step RT-PCR

To obtain specific RT-PCR products, the QIAGEN OneStep RT-PCR kit (Qiagen) was employed which contains optimized components that allow both reverse transcription and PCR amplification to take place in what is commonly referred to as a ”one-step” reaction.

2.15 Labeling and hybridization of nucleic acids

2.15.1 Generation of ^{32}P labeled cDNA by using the random prime method

The Ready Prime kit II (Amersham) was employed for labeling DNA fragments with [_-^{32}P] isotopes. The method relies on the random priming principle developed by Feinberg and Vogelstein (1989).

2.15.2 Northern blot hybridization with radioactive by labeled cDNA-probes

The membrane to be hybridized was equilibrated in 2x SSC and transferred to a hybridization bottle. After adding 10 ml of hybridization solution and sheared salmon DNA, the membrane was incubated for 2 hours in the hybridization oven at temperature 65°C. Then, the labeled probe was heat denatured and added to the hybridization solution. The

hybridization was carried out overnight in the oven. Next day, the filter was washed for 10 min with 2x SSC at room temperature. Finally the filter was washed with 0.2x SSC containing 0.1% SDS at the hybridization temperature. After drying the filter, it was sealed in Saran wrap and exposed to autoradiography overnight at -80°C or to a Phosphoimager screen (Bio-Rad) for 1-4 hours. The film was developed in X-Ray Automatic Processor Curix 60 or the screen was scanned in a Phosphoimager (Bio-Rad).

2.16 Protein techniques

2.16.1 Isolation of total proteins from eukaryotic cells

Eukaryotic cells were cultured at 37°C in a humidified incubator with 5% CO_2 and grown to 80–90% confluence before isolation of total proteins. Cells were washed with PBS and 300 µl of lysis buffer per big (75 cm^2) cell culture flask (Greiner Nunc) was added. The whole cell lysate was collected with a cell scraper and transferred to a separate tube. The sample was homogenized by using ultrasound (Sonifier 250, Branson Ultrasonic) and centrifuged for 5 min at 10000 x g. The supernatant contains the whole cell lysate which was immediately used for Western blotting or stored at -20°C.

2.16.2 Determination of protein concentration

(Bradford, 1976)

To determine the protein concentration, the Bio-Rad protein assay was employed which is a dye binding assay based on the differential color change of a dye in response to

various concentrations of protein. The assay is based on the observation that the absorbance maximum for an acidic solution of Coomasie Blue G-250 shifts from 494 to 595 nm when the binding to the protein occurs. The absorption of the color reaction was measured at 595 nm in a microplate reader (Microplate Reader 450, Bio-Rad).

2.16.3 Electrophoresis of proteins

The NuPAGE® Pre-Cast Gel System (Invitrogen) is a polyacrylamide gel system for high performance gel electrophoresis and is based on SDS-PAGE gel chemistry (Laemmli, 1970). It consists of NuPAGE® Bis-Tris Pre-Cast Gels and specially optimized buffers which have an operating pH of 7.0, giving the system several advantages over existing polyacrylamide gel systems with an operating pH of 8.0. A neutral pH increased the stability in both proteins and gels, providing with increased confidence in electrophoretic results. Protein samples (50 µg) with NuPAGE LDS (Sample buffer) were heat denaturated at 70°C for 10 min, chilled at room temperature for 5 min, and loaded onto NuPage 10% Bis-Tris pre-cast (Invitrogen). To determine the size of the proteins on a gel, 10 µl of a pre-stained molecular weight standard (See Blue Plus2, Invitrogen) was loaded. The gel was run at 150 V for 2 hours at room temperature with 1x MES buffer.

2.16.4 Western blotting of proteins onto nitrocellulose filters

After the electrophoresis of proteins on a polyacrylamide gel, proteins were transfered on a nylon membrane by the semi-dry system using an electro-blotter (Biometra). The transfer was carried out at 100 mA at room temperature for 1-2 hours in the transfer buffer: 25 mM Tris pH 8.3, 150 mM glycine, 10% methanol.

2.16.5 Staining of polyacrylamide gels

To assess transfer efficiency of proteins onto nitrocellulose membranes, the gel was incubated overnight in Coomassie blue solution and washed in water for 2-3 hours at room temperature.

2.16.6 Incubation of protein-bound membranes with antibodies

The membrane was first incubated in a wash solution with 5% non-fat dry milk for 1 hour at RT followed by an incubation step with a primary antibody at the recommended antibody dilution in wash solution with 2% non-fat dry milk for 1 hour at room temperature. Then, the membrane was washed five times in wash solution with 2% dry milk for 5 to 10 min. and incubated with the appropriate secondary antibody at 1:10000 dilution in wash solution with 2% non-fat dry milk for 1 hour at room temperature. After an 1-hour incubation step the membrane was washed five times in wash solution with 2% dry milk, one time in wash solution without dry milk and one time in P3 buffer for 5 min at RT. Finally, the proteins on the membrane were visualized by an incubation step in the dark with 10 ml of staining solution (alkaline phosphatase substrate solution) for 15 min and rinsed with water to stop the reaction.

Wash solution:	150 mM NaCl
	100 mM Tris/HCl, pH 7.5
P3 buffer:	100 mM NaCl
	50 mM $MgCl_2$

100 mM Tris/HCl, pH 9.5

Staining Solution: 45 µl NBT (75 mg/ml in DMF)

35 µl BCIP (50 mg/ml in DMF) in 10 ml of P3 buffer

(Carl Roth & Co., Karlsruhe)

Alternatively secondary anti-mouse or anti-rabbit HRP-coupled antibody was used at a working dilution between 1:20,000 to 1:100,000. Detection of antibodies was carried out with SuperSignal® West Pico Chemiluminescent Substrate (Pierce). The blot stained with antibodies was incubated with Working Solution for 5 min and exposed to an X-ray film.

2.16.7 Expression of recombinant proteins in the pET system (Novagen)

Plasmids with pET41 constructs were transformed to expression in the host bacterial strain *E.coli* BL-21(DE3). A single bacterial colony containing the vector with the fusion construct was picked from a freshly streaked plate and 50 ml LB culture with kanamycin was inoculated. Bacterial culture was incubated with shaking at 37°C until an OD600 of 0.4–1 was reached. A noninduced sample was removed as a control. To the remainder, IPTG from a 100 mM stock was added to a final concentration of 0.4 mM and the incubation step was continued for 2–3 hours. Then, the induced sample was removed and flasks were placed on ice for 5 min. Cells were harvested by centrifugation at 5000 x g for 5 min at 4°C, resuspended in 0.25 culture volume of cold 20 mM Tris-HCl, pH 8.0, and centrifuged as above. Finally, the supernatant was removed and cells were stored as a frozen pellet at –70°C or used directly for purification.

2.16.8 Purification of GST fusion proteins

GST-fusion proteins were purified from bacterial cell extracts using the GST-binding kit (Novagen) according to the manufacturer's instruction. Integrity of the resulting proteins was checked by SDS-PAGE analysis.

2.16.9 Peptide analysis

Different computational tools were applied to select potential antigenic peptides. Before synthesis of the peptide, a hydrophilicity/hydrophobicity profile analysis was carried out and for further confirmation antigenicity prediction was performed. In the next step, predictions of secondary structures such as _-turns and _-helices in combination with the surface probability of the protein region were the parameters which enabled us to select the most appropriate peptides. In the last step, the primary sequence of the PHF5a protein was compared with sequences at the international data bank (ExPASy) to select unique sequences for antibody generation. Two peptides were selected and synthesized. The sequences of peptides for PHF5a are as follows:

EP010897 (PHF5a)	H2N	GSS KTD LFY ER(SpacerC6) C	$CONH_2$
EP010898 (PHF5a)	H2N	CTD LFY ERK KYG FKK R	COOH

2.16.10 Immunization of rabbits

The company Eurogentec performed immunization using the DOUBLE X program. Two peptides were selected and synthesized instead of one. Using modern algorithms for peptide selection, the success rate for peptide immunization can be as high as 75%. This still means a 25% chance of failure. By using the DOUBLE X program the success rate is increased to 93.75%. 5 mg of each peptide was conjugated with carrier protein molecules (KLH) and mixed together before immunization. Two rabbits were immunized with 100 μg of antigen mixed with Freund's complete adjuvant in a 1:1 ratio. Before injection, pre- immune

sera were collected from the animals. After 14 days a second booster immunization was performed with a 1:1 ratio of antigen with Freund's incomplete adjuvant. A third booster was given after 28 days and final bleeding of the animals was done after 36 days. The antiserum was aliquoted and stored at -80^{o}C.

2.16.11 Determination of polyclonal antibody titers

After SDS-PAGE and electrotransfer of total mouse proteins to a nylon membrane, the membrane was blocked for 1 hour at room temperature. Then, the filter was cut and each lane was incubated with different dilutions of antiserum (1:25, 1:100, 1:500) in washing buffer for overnight at 40^{o}C. Thereafter, the unbound antiserum was removed by washing the membrane 3 times for 20 min with washing buffer. The secondary antibody coupled with alkaline phosphatase was diluted 1:10000 in washing buffer, added to the blot and incubated for 1 hour. Again the unbound antibodies were removed by washing 4 times for 15 min with washing buffer. The chromogenic reaction was performed with 33 µl of NBT and 66 µl of BCIP solution in buffer AP until chromogenic precipitate developed. The reaction was stopped by washing the membrane several times with water. The membrane was air-dried and stored in the dark.

2.16.12 Affinity purification of polyclonal antibodies

For antibody purification, SulfoLink Coupling Gel (Pierce) was used. The gel consists of immobilized iodoacetyl on a crosslinked agarose support. SulfoLink support binds specifically to sulfhydryl groups. The 12-atom spacer arm makes binding more efficient. This longer arm is designed for conjugating small peptides to the support.

2.16.13 Immobilization of peptides

Sample preparation buffer: 0.1 M sodium phosphate

5 mM EDTA-Na, pH 6.0

Coupling buffer: 50 mM Tris

5 mM EDTA-Na, pH 8.5

The peptide (10 mg) was dissolved in 1 ml of sample preparation buffer. The solution was added to a vial containing 6 mg of 2-mercaptoethylamine (2-MEA) and incubated at 37°C for 1.5 hours. The mixture was cooled to room temperature and applied to the 5 ml desalting column which was equilibrated with 30 ml of coupling buffer to remove excess 2-MEA from the reduced sample. 1 ml fractions were collected and fractions 4, 5 and 6 were pooled for gel coupling.

2.16.14 Gel coupling and blocking of nonspecific binding sites

The reduced protein mixture (3 ml) was added to 2 ml SulfoLink Coupling Gel column after equilibrating with 12 ml of coupling buffer. The column was mixed at room temperature for 15 min, and then incubated for 30 min without mixing. Subsequently, the column was washed with 6 ml of coupling buffer, then 2 ml of 0.05 M cysteine in coupling buffer was applied to the column. The column was mixed for 15 min at room temperature, and then incubated for 30 min without mixing

2.16.15 Washing and deactivation of the SulfoLink column

Washing buffer A: 1.0 M NaCl

Washing buffer B: 1.0 M NaCl

0.05% sodium azide in PBS

A series of alternate washings with buffer A (injection 4x4 ml) and buffer B (injection 3x4 ml) were performed. Finally, 4 ml of 0.05% degassed sodium azide in PBS was injected, and then the top porous disc was inserted to the column.

2.16.16 Purification of antibodies

The column was equilibrated with 6 ml of PBS. The antiserum (8 ml) was applied onto the column. The column was incubated at room temperature for 1 hour. During pumping a constant flow rate of 0.5 ml/min was maintained. The column was washed with 16 ml of PBS. Elution was done with 8 ml of glycine buffer (100 mM, pH 2.5-3.0) and subsequently 1 ml fractions were collected and monitored by absorption at 280 nm. Fractions 3 and 4 were pooled and 0.05% sodium azide was added. Purified antiserum was stored at -20°C and the column was re-equilibrated with 10 volumes of PBS.

2.17 Immunoprecipitation

2.17.1 *In vitro* transcription and translation

In vitro transcription/ translation was performed using Promega's TNT® T7 Coupled Reticulocyte Lysate System (Promega) to prepare ^{35}S-Met-labeled bait and library proteins. The coupled transcription/translation system greatly simplifies the process and reduces the experimental time. Circular plasmid DNA (pGBKT7 and pGADT7 vectors) with the inserts under the control of the T7 promoter was used as a template. Reaction components were assembled in the microcentrifuge tube as follows:

TNT® Rabbit Reticulocyte Lysate	25 µl
TNT® Reaction Buffer	2 µl
TNT® RNA Polymerase T7	1 µl
Amino Acid Mixture, minus Methionine 1mM	1 µl
[^{35}S]methionine (>1,000 Ci/mmol at 10 mCi/ml)	2 µl
RNasin® Ribonuclease Inhibitor (40 U/µl)	1 µl
DNA template(s) (0.5 µg/µl)	2 µl
Nuclease-Free Water	to a final volume of 50µl

The reaction was incubated for 90 min at 30^{o}C and immediately used in the immunoprecipitation assay

2.17.2 Coimmunoprecipitation experiments

Following reagents were combined in a 1.5 ml microcentrifuge tube on ice.

10 µl in vitro translated (^{35}S-methionine-labeled) bait protein

10 µl in vitro translated (^{35}S-methionine-labeled) library protein

The mixture was incubated at room temperature for 1 hour and 10 µl (i.e., 1 µg) of c-Myc monoclonal antibody **or** HA-Tag polyclonal antibody was added to the reaction tube with appropriate proteins. The mixture was incubated for an additional 1 hour. Meanwhile, Protein A beads were prepared as follows:

Beads were mixed gently by inverting several times. A sufficient volume of beads was transferred a to a clean 1.5 ml microcentrifuge tube. The beads were washed twice with 200 µl of PBS in a microcentrifuge tube and centrifuged at 5,000 x g for 30 sec. The supernatant was removed by aspiration with a micropipette. Steps were repeated and finally, the beads were resuspended to their original volume (i.e., the original bead volume was transferred to microcentrifuge tube) by adding fresh PBS.

3 µl of Protein A beads were added to the reaction tube. To ensure adequate mixing, the reaction tube was rotated at room temperature for 1 hour. The tube was centrifuged at 5,000 x g for 10 sec and the supernatant was discarded. Beads were washed 5 times with buffer I (150 mM NaCl, 50 mM Tris, pH 7.4, 2 mM EDTA, pH 8.0, 1% Triton X-100/ 10% glycerol/ 2 mM DTT) and 3 times with buffer II (150 mM NaCl/ 50 mM Tris, pH 7.4). Finally the beads were resuspended in 20 _l SDS-PAGE-Loading Buffer, denaturated and loaded onto a SDS-PAGE minigel.

After electrophoresis gel was fixed in a fixation solution (isopropanol:water:acetic acid, 25:65:10) for 45 min, subsequently rinsed in H_2O and incubated in Amplify Fluorographic Reagent (Amersham) according to the manufacturer's instructions. The gel was placed onto pre-wetted Whatman 3MM paper, covered with Saran wrap and dried at 80^{o}C under constant vacuum. The Saran wrap was removed and gel was exposed to an X-ray film overnight at room temperature.

2.18 Yeast-hybrid techniques

Yeast two-hybrid experiments were performed using the MATCHMAKER System 3 (BD, Clontech). In the MATCHMAKER System 3 a bait gene is expressed as a fusion protein with the GAL4 DNA-binding domain (DNA-BD), while another gene or cDNA is expressed as a fusion protein with the GAL4 activation domain (AD); (Fields and Song, 1989; Chien et al., 1991). If bait and library fusion proteins interact, the DNA-BD and AD are brought into proximity, thus activating transcription of four reporter genes ADE2, HIS3, and MEL1 (or lacZ)- under the control of distinct GAL4 upstream activating sequences (UASs) and TATA boxes. This technology can be used to identify novel protein interactions, confirm suspected interactions, and define interacting domains.

2.18.1 Transformation of yeast cells

(Ito et al., 1983)

Lithium Acetate-mediated yeast transformation was used according to the Clontech protocol based on the original method (Ito et al., 1983) and modified by Schiestl and Gietz (1989), Hill et al. (1991), and Gietz et al. (1992). In the LiAc transformation method, competent yeast cells are prepared and suspended in a LiAc solution with the plasmid DNA to be transformed, along with excess carrier DNA. Polyethylene glycol (PEG) with the

appropriate amount of LiAc is then added and the mixture of DNA and yeast is incubated at 30°C. After the incubation steps, DMSO is added and the cells are heat shocked, which allows the DNA to enter the cells. The cells are then plated on the appropriate medium to select for transformants containing the introduced plasmid(s)

2.18.2 Yeast two-hybrid library screening

To screen for interaction partners, the pretransformed murine 11.5-days embryo Matchmaker library (BD Clontech) was mixed with the yeast strain containing the bait construct. Mating of both strains occurred during a 24-hours incubation step at 30^{o}C in 2x YPDA medium. Mating cultures were plated on SD minimal medium lacking leucine, tryptophan, histidine, adenine (-Leu, -Try, -His, -Ade) and supplemented with X- αGal. After 4 days, blue colonies were selected and library inserts were amplified by colony PCR and sequenced from both ends. The identity of individual clones was determined using the BLAST program to screen the GenBank database (NCBI). Interactions of RS domain proteins with PHF5a and domain mapping experiments were investigated by cotransformation of appropriate BK and AD constructs into the yeast strain AH109. Selection of transformants was performed as described above.

2.18.3 Quantitative α-Gal assay

Yeast cells were cultured in selective SD medium (-His, -Try, -Leu) at 30^{o}C for 16-18 hours with shaking. After the incubation step, the OD600 was measured and 1.0 ml of the culture was centrifuged to pellet the cells completely. The supernatant was transferred and 48 µl assay buffer (0.5 M NaOAc, pH 4.5, 100 mM PNP α-Gal, 2:1 (v/v) ratio) was added to 16 µl of the sample. The reaction was stopped after 30 min by adding 136 µl of 1 M Na_2CO_3. Each experiment was performed in triplicate. Optical density was recorded by using microplate reader (Bio-Rad) at 410 nm and the amount of α-galactosidase was calculated.

2.18.4 Yeast three-hybrid

The more complex protein interactions (ternary protein complex formation) can be investigated with the three-hybrid system. The third protein in this system can participate in the interaction in several ways: as a "bridge", interacting with two proteins that do not directly interact with each other; to stabilize a weak interaction between two proteins; or as an inhibitor or modifier (e.g., kinase) of one or both of the proteins. Alternatively, expression of the third protein may inhibit the two-hybrid interaction.

In the yeast three-hybrid assay (BD Clontech) cells were transformed with the bicistronic construct pBridge and pGADT7. The pBridge vector contains two distinct multiple cloning sites to allow expression of the DNA-binding domain (BD) fusion as well as a third protein. The third protein in this system can participate in the interaction as a "bridge", interacting with two proteins that do not directly interact with each other (Figure 3). Selection of positive clones was performed as described above (2.18.2). Expression of the bridge protein was controlled by a conditional (P^{Met25}) promoter, such that expression could be repressed by the addition of 1 mM methionine on the selective plate.

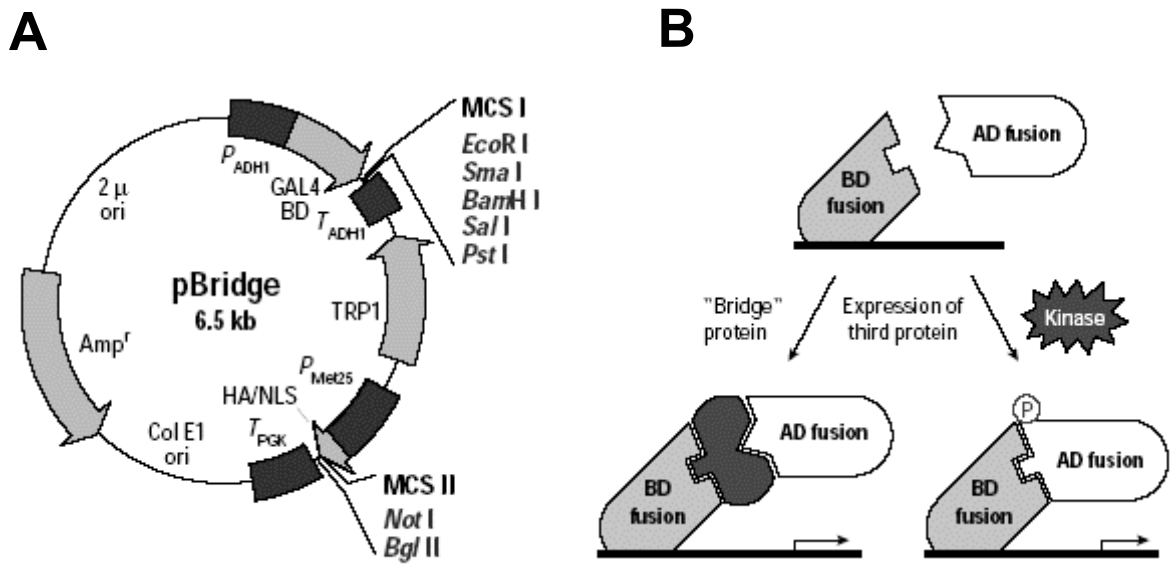


Figure 3. **A**. Schematic representation of the pBridge vector used in the yeast-three hybrid assay. A protein of interest is expressed as a fusion to the DNA-BD, while an additional protein of interest is conditionally expressed from the P^{Met25} promoter. Interactions dependent on three separate proteins can be investigated when pBridge is combined with an AD fusion vector from a GAL4 based two-hybrid system. **B.** Schematic representation of the three-hybrid system strategies. pBridge expresses both the DNA-BD fusion and the third protein. The AD fusion is expressed by a two-hybrid system vector. The conditionally expressed third protein can play a structural (left) or modifying (right) role in the interaction that restores reporter gene expression (according to the MATCHMAKER yeast three-hybrid vector, CLONTECHniques April, 1998)

2.19 Site directed mutagenesis

Site directed mutagenesis was used to introduce mutations in the sequence of PHF5a. Using QuikChange® Site-Directed Mutagenesis Kit from Stratagene basic amino acids from terminal parts of PHF5a were substituted by neutral aminoacids using specific primers. In the procedure supercoiled vector DNA with the PHF5a as an insert and complementary forward and reverse primers with desired mutations were used. PCR extension generated mutated plasmid containing staggered nicks. Subsequently DNA product is digested with *Dpn*I endonuclease specific for methylated parental DNA and therefore mutation-containing plasmids are selected. DNA isolated from DH-5_ *E. coli* strains is dam methylated and therefore susceptible to *Dpn*I digestion. The nicked vector DNA containing the desired mutations is then transformed into XL1-Blue supercompetent cells.

2.20 SELEX and Whole-Genome PCR

Two different approaches were used for PHF5a binding site selection. The SELEX (Systematic Evolution of Ligands by Exponential Enrichment) process consists of three basic steps: selection for binding from a library of random DNA oligonucleotides; amplification of

those oligonucleotides which have bind during the selection round; repetition of these two steps with increasing stringent conditions during the selection round to find the best binders. This process leads to “aptamers” which would bind with high affinity to the protein of interest.

Recombinant GST fusion PHF5a protein was incubated with the library of random 12 bp oligonucleotides with primer adaptors at both ends allowing for amplification of selected molecules.

2.20.1 SELEX

A PHF5a-GST protein was prepared and coupled to the GST-resin according to section 2.16.7, 2.16.8. Approximately 100 ng of GST-coupled PHF5a protein was mixed with 350 ng of library in buffer A (20 mM HEPES, pH 7.9, 40 mM KCl, 50 uM $ZnSO_4$, $MgCl_2$, 0.1 mM EGTA, 5% glycerol, 1 mg/ml BSA, 1x protease inhibitor cocktail, 2 ug/ml poly-dI-dC, 5 mM DTT, 1 mM PMSF) and incubated for 20 min at room temperature with rocking. The GST-resin was washed 3 times with buffer A, suspended in 30 _l of H_2O and boiled for 3 min. The supernatant from the boiling step was used for PCR amplification. The amplified product was resolved on a 3% agarose gel, purified and used for other rounds of exponential selection and amplification.

2.20.2 WG-PCR

The Whole Genome PCR method (WG-PCR) was developed to identify direct target gene promoter/enhancer sequences for DNA binding proteins (Kinzler and Vogelstein, 1989). Genomic DNA fragments are selected by their binding to the protein of interest and amplified by PCR. After repeated cycles of selection and amplification, target sequences were sequenced and analyzed

WG-PCR Protocol:

15 _g of human genomic DNA was digested with *Mbo*I and ligated to *Mbo*I cutted primer adapters. Ligated DNA was separated from free linkers by electrophoresis in an agarose gel. Subsequently, 1 _g of linker ligated DNA was incubated with 5 _g of recombinant GST fusion protein for 20 min on ice in 50 _l of 1 x binding buffer. (10x 200 mM Tris, pH 7.6, 500 mM NaCl, 10 mM $MgCl_2$, 2 mM EDTA, 50% glycerol, 5 mM DTT, 0.5 mM PMSF).

The GST-fusion protein in the complex with the DNA target sequence was coupled with GST-resin and purified by 3 times washing in buffer TN (10 mM Tris-HCl, pH 7.5, 150 mM NaCl). Finally, beads were dissolved in 30 _l of H_2O and boiled. Pulled down linker ligated DNA fragments were amplified by PCR and used for additional rounds of selection. After 5 rounds of selection and amplification DNA fragments were cloned into the pGEM-T easy vector, sequenced and analyzed.

2.21 Whole mount in situ hybridization

(Wilkinson, 1992; modified)

”Whole mount” hybridization was employed for the hybridization of entire mouse embryos, in order to determine the pattern of expression of genes during embryogenesis.

2.21.1 Embryo dissection

Mice were breeded and the females checked the following day for the presence of a vaginal plug, indicating 0.5-day p.c. embryos. On the 10.5 and 11.5-days post coitum (p.c), the pregnant female mice were dissected and the embryos removed and placed in ice-cold

PBS, where all the extraembryonic membranes were removed. The embryos were then fixed overnight rocking at 4°C in 4% paraformaldehyde (in PBS).

The embryos were washed twice in ice-cold PBT for 10 min and then gradually dehydrated by 2 min incubation steps in 25%, 50%, 75% and 100% methanol (in PBT), respectively. Embryos can be stored at this point.

2.21.2 DNA preparation and probe synthesis

The fragment to be transcribed must first be cloned into one of the transcription vectors and the orientation of the insert has to be determined. Transcription vectors, such as pBluescript (Stratagene, USA) or pGEM (Promega), possess transcription start points for bacteriophage RNA polymerases (T3, T7, M13) situated on either of the polylinker sites of the plasmid. The plasmid is then linearised to avoid transcription continuing over the insert of the plasmid itself.

The clone PHF5a in the pGEM T easy vector was linearised with *Sst*II for the sense probe and with *Sal*I for the antisense probe. The linearisation of the plasmid was tested on an agarose gel and the DNA was precipitated.

The transcription was carried out as decribed:

13 µl H_2O

2 µl 10x transcription buffer

1 µl 0.2 M DTT

2 µl dNTP mix (10 mM ATP, CTP, GTP, 6.5 mM UTP; 3.5 mM DIG-11-UTP)

1 µl linearised plasmid (1 µg/µl)

0.5 µl Placental ribonuclease inhibitor (100 U/µl)

1 µl T7 RNA polymerase (10 U/µl) for the sense and T3 for the antisense probe

Incubation was performed for 2 hours at 37°C. The amount and integrity of the RNA synthesized was estimated by loading 1 µl aliquot of reaction on 1% agarose gel with 10% formaldehyde. Subsequently, 2 µl DNase I (ribonuclease free) was added and the reaction was incubated at 37°C for 15 min. The purification of the product was carried out as follows:

100 µl H_2O

10 µl 4M LiCl

300 µl ethanol

The mixture was incubated for 30 min at -80°C, centrifuged for 10 min and washed with 70% ethanol. The dried pellet was dissolved in H_2O to a concentration of 0.1 µg/µl and stored at -20°C. Then, 10 µl probe was used for each ml of hybridization mix.

2.21.3 Pre-hybridisation treatment and hybridization of the embryos

Rehydratation was performed by taking the embryos through this MeOH/PBT series in reverse. They were washed twice with PBT and bleached with 6% hydrogen peroxide in PBT, fresh prepared, for 1 hour. This treatment has been found to decrease considerably the amount of nonspecific background (Wilkinson, 1992). The embryos were then washed three times with PBT and treated with 10 µg/µl proteinase K in PBT for 20 min. Proteinase K is used to improve the signal by increasing the accessibility of target RNA (Wilkinson, 1992). After treatment with proteinase K, the embryos were washed with freshly prepared 2 mg/ml glycine in PBT, then twice with PBT. The embryos were refixed with fresh 0.2% glutaraldehyde and 4% paraformaldehyde in PBT for 20 min and then washed twice with PBT.

The hybridization steps were carried out in a 6 well cell culture dish and the solutions were exchanged using a fine syringe. Prehybridization was carried out, with hybridization mix (PreHyb; 50% Formamid, 5 x SSC, pH 4.5, 1% SDS, 50 µg/ml yeast tRNA) at 70°C for 1 hour. For the hybridization, the labelled probe was added to the solution and incubated at the

same conditions, overnight. The DIG labelled probe was diluted to 1 µg/µl and 10 µl was used for the hybridization.

2.21.4 Detection of Hybridization Signals

Posthybridization washes were performed next to remove the nonspecifically bound probe, followed by the detection of hybridization signals as described:

Solution I: 50% Formamide, 5x SSC, pH 4.5, 1% SDS, in non-DEPC water

Solution II: 50% Formamide, 2x SSC, pH 4.5, 0.2% SDS, in non-DEPC

MAB: 100 mM Maleic acid, 150 mM NaCl, 2 mM Levamisole, 0.1% Tween-20. NaOH pellets were used to bring the pH to 7.5 and the solution was freshly prepared on the day of use.

2x	30 min wash solution 1	70°C
	10 min wash solution 1:2 (1:1)	70°C
3x	5 min wash solution 2	room temperature
2x	30 min wash solution 2 with 100 µg/µl Rnase A	37°C
	5 min wash solution 2	room temperature
	5 min wash solution 3	room temperature
2x	30 min wash solution 3	65°C
3x	5 min. Maleic acid buffer (MAB)	room temperature

Maleic acid with 2% blocking solution was added to the embryos and incubated at room temperature for 15 min, then 20% FCS was added and further 60 min of incubation was carried out. The incubation continued at 4°C for 30min. After that, the antibody was added to the reaction which remained rocking at 4°C overnight.

Post-antibody washes were carried out to remove unbound antibody as described below:

NTMT: 100 mM Tris-HCl, pH 9.5, 50 mM $MgCl_2$, 100 mM NaCl, 0.1% Tween-20, 2 mM levamisole.

2x	5 min. MAB	room temperature
2x	4 hours TBST	4°C
1x	overnight TBST	4°C
3x	10 min. NTMT	4°C

2.21.5 Staining of embryos

In order to detect the alkaline phosphatase activity of the anti-DIG antibody, the embryos were incubated by rocking at 4°C in colouring solution (NTMT containing 4.5ul NBT, 3.5 _l BCIP per ml). The staining reaction was kept in the dark and the colour's development was monitored. When the desired extent was achieved, the reaction was stopped by washing the embryos twice with PBT (5 min). They were transfered through 30%, 50% and then 80% glycerol, each step lasting 1day and stored in the last solution at 4°C.

2.22 Eukaryotic cell culture methods

2.22.1 Cell culture conditions

NIH3T3 cells and GC-4spc were grown in DMEM medium (PAN) containing 15% fetal bovine serum (PAN) and 1% penicillin/streptomycin solution (PAN). The cells were cultured at 37°C in a humidified incubator with 5% CO_2.

2.22.2 Trypsinization of eukaryotic cells

Cells were washed twice with sterile PBS and incubated in minimal amount trypsin-EDTA (PAN) (0.5 g/l trypsin, 0.2 g/l EDTA) at 37°C until they had detached from the dish. The process was controlled under an inverted microscope. Trypsin was inhibited by addition of growth medium in which the cells were subsequently resuspended. Cell counting was performed, when necessary, using an improved Neubauer chamber, and the cells were plated out or harvested for cryoconservation.

2.22.3 Cryoconservation and thawing of eukaryotic cells

Resuspended cells were centrifuged (1000 x g for 5 min at 4°C) in 4 ml growth medium. The supernatant was aspirated and the cells resuspended (1-5 x 10^7 cells/ml) in ice-cold freezing medium (DMEM, 20% FCS, 10% DMSO). Cells were kept for 7 days at -80°C and then stored in liquid nitrogen. For revitalization, frozen cells were quickly thawed, gently transferred to disposable Falcon tubes containing 4 ml cold growth medium and centrifuged as described above. The supernatant was discarded by aspiration, and cells were plated out after being resuspended in a suitable amount of growth medium.

2.22.4 Transient transfection of cells

NIH3T3 cells (5 x 10^4) were plated on coverslips 24 hours before transfection, and washed with PBS. The construct DNA pEGFP-C1/pDSRed1-C1 (1.5 _g per well) was then introduced into the cells using the PolyFect transfection reagent (Qiagen) according to the manufacturer's instructions. Twenty-four hours later, transiently transfected cells were fixed on chamber slides cells with 4% paraformaldehyde for 15 min.

2.22.5 Cytochemical analysis

Cells attached to the cover slips were rinsed twice with PBS and fixed with 4% formaldehyde in PBS for 6 min at room temperature, then rinsed with PBS. Fixed cells were permeabilized with 0.2% Triton X-100 in PBS for 6 min. Alternatively, cells were fixed in -20^{o}C methanol. Cells were stained with antibody by blocking in 2% BSA solution in PBS buffer, subsequently incubated overnight at 4^{o}C with primary antibodies and 2 hours at room temperature with secondary, either FITC or Cy^{3}-coupled antibodies. After washing in PBT buffer the stained and transiently transfected cells were mounted with Vectashield mounting medium with DAPI (Vector Laboratories Inc.) and observed under a fluorescence microscope (BX-60, Olympus).

2.22.6 Gene silencing by RNA interference

The RNAi –based method can be used to target specific mRNAs for degradation, thereby silencing their expression. It was shown, that sequence-specific single strand RNA oligonuclotides, both in sense and antisense orientation, can be employed as a negative control for this specific gene silencing, as it has no effect on specific mRNA degradation (Tuschl et al., 1999). Transfection of NIH3T3 cells was accomplished using the Oligofectamine Reagent (Invitrogen) which forms stable complexes between the lipid and oligonucleotides permitting efficient delivery of the RNA molecules into mammalian cells. The transfection procedure was performed according to the supplier's instructions with PHF5a gene-specific siRNA duplexes (Eurogentec) at 0.66 µg oligonucleotides per 0.5 ml transfection medium (150 nM).

2.22.7 Cell proliferation assay

Cell proliferation was determined by using the EZ4U Kit (Biozol) as described in the user manual. This method is based on the finding that living cells are capable to reduce uncoloured tetrazolium salts into intensively coloured formazan derivatives. This reduction process requires functional mitochondria, which are inactivated within a few minutes after cell death. This method therefore provides an excellent tool for the discrimination of living and death cells. Approximately 3000 cells/well were cultivated in 200 µl cell culture medium in a 96-well plate as described above. After 2 hours, 24 hours and 48 hours of incubation at 37°C in a humidified incubator with 5% CO_2, 20 µl of substrate (EZ4U Kit) per well was added and after 2 hours the optical density at 450 nm and 620 nm (references) was measured. All experiments were assayed in triplicates.

2.22.8 *In vivo* coimmunoprecipitation

In vivo coimmunoprecipitation allows the detection of protein-protein interactions occurring under physiological protein concentration and conditions in living cells. The method is limited by the accessibility to antibodies, specific for interacting proteins. Immunoprecipitation of PHF5a, $U2AF^{35}$ and $U2AF^{65}$ was performed as follows: GC-4spc cells were washed, harvested and lysed for 30 min in the lysis buffer (50 mM Tris, pH 7.4, 150 mM NaCl, 2 mM EDTA, 0.1% NP40, 2.5% glycerol, 2 mM DTT, 25 µg/ml aprotinin, 10 µg/ml pepstatin, 10 µg/ml hymostatin, 10 µg/ml leupeptin, 1 mM PMSF). Lysates were gently incubated overnight at 4^oC with appropriate antibodies (anti $U2AF^{35}$ and anti-$U2AF^{65}$ antibodies at dilution 1:50 and anti- PHF5a at dilution 1:20 in PBS buffer) and with 10 _l of protein A agarose beads (ICN). Unbound material was washed out from the beads by rocking

the suspension 5 times, 10 min with cold lysis buffer and centrifugation at 600 x g, 2 min, and final washing in PBS. Bound proteins were suspended in SDS loading buffer, denaturated at 70^{o}C for 5 min and subsequently separated on 4-12% SDS polyacrylamide gel and detected by Western blotting.

2.22.9 Preparation of cell suspension for immunostaining

The cell suspensions were prepared from testes of adult mice and stained with anti-PHF5a antibody. This two step procedure includes preparation of cell suspension by digestion of testes with collagenase/trypsin and mechanical dispersion, followed by seeding of cells on laminin-coated dishes and staining with antibodies.

Preparation of cell suspension from murine testes

Testes were isolated from mice 5 to 60 days old. The *Tunica albuginea* was removed and testes were incubated in 10 volumes of Hank's balanced salt solution without Ca^{2+} and Mg^{2+} (HBSS), containing 1mg/ml collagenase (Type IV, Sigma) at 37^{o}C with gentle agitation for 15 min until tubules separated (optional 200-500 _g/ml DNAse). Tubules were dispersed by careful dissection and washed 2-4 times in 10 volumes of HBSS and incubated in 1mM EDTA/ 0.25% trypsin for 5 min at 37^{o}C. Separation and dispersion of tubular cells was done by pipetting and gentle agitation. Subsequently 10-20% of fetal bovine serum (FBS) was added and large pieces of undigested material were removed and cells were filtered through a nylon mesh with 60 _m pore size to remove large clumps of cells. The filtrate was centrifuged at 600 x g for 5 min at 16^{o}C and the supernatant was removed carefully from the pellet.

Finally, the cells were resuspended in the sperm cell medium or Dulbecco's modified Eagle's medium, containing 10% FBS.

Preparation of laminin-coated dishes

Falcon Petri dishes were incubated overnight at 37°C with laminin (Sigma) at the concentration of 20 _g/ml in PBS. Solution was removed and dishes were washed 3 times with 2 ml of PBS. Dishes were blocked with 0.5 mg/ml BSA for 1 hour at 37°C. Before adding cells, dishes were washed 3 times with PBS. Finally, cells were suspended (4 x 10^7 cells in 8 ml DMEM + FBS), divided among four coated dishes and incubated at 32°C, followed by immunostaining according to the protocol given in section 2.21.5.

2.23 Antibody staining of testis sections

The localization PHF5a and U2AF^{35} proteins localization within the testis was performed by immunostaining of testis sections with specific antibodies as follows:

Testes were fixed overnight in 10% formalin at 4°C. The fixed tissue was paraffin embedded and sections were mounted on slides. Paraffin was cleared by 2 times incubation in RotiClean for 10 min. Slides were rinsed twice for 2 min in 100% (18:1:1 ethanol:methanol:isopropanol) and rehydrated in a gradient of alcohols at 95%, 80%, 50% for 2 min each followed by incubation in deionized H_2O for 5 min. Slides were placed face-up and covered with 1% SDS in PBS and incubated for 5 min at room temperature. Slides were washed 3 times with PBS and sections were blocked in blocking buffer (serum from host species of secondary antibody to be used, diluted 1:10 in PBS). Tissues were covered with

primary antibody (1.0-10 _g/ml or 1:20- 1:50 dilution) diluted in PBS and incubated 2 hours at 37^{o}C.

To wash out unbound antibody, slides were rinsed 3 times with PBS and incubated with secondary antibody (1:100-1:200), diluted in blocking buffer for 1 hour at 37^{o}C. Subsequently, slides were rinsed 3 times with PBS, counterstained with DAPI, mounted and visualized.

3 Results

3.1 Selection of PHF5a interaction partners

3.1.1 Yeast two-hybrid library screening

In order to determine PHF5a binding partners the yeast-two hybrid library screening was performed. The Matchmaker murine embryo 11.5-days library (BD Clontech) was screened by using the PHF5a protein as a bait. The BK/PHFa construct used for screening of the yeast two-hybrid library was generated by fusing the complete ORF of PHF5a (Accession No.: NM_026737) in frame into the pGBKT7 vector (Figure 4)

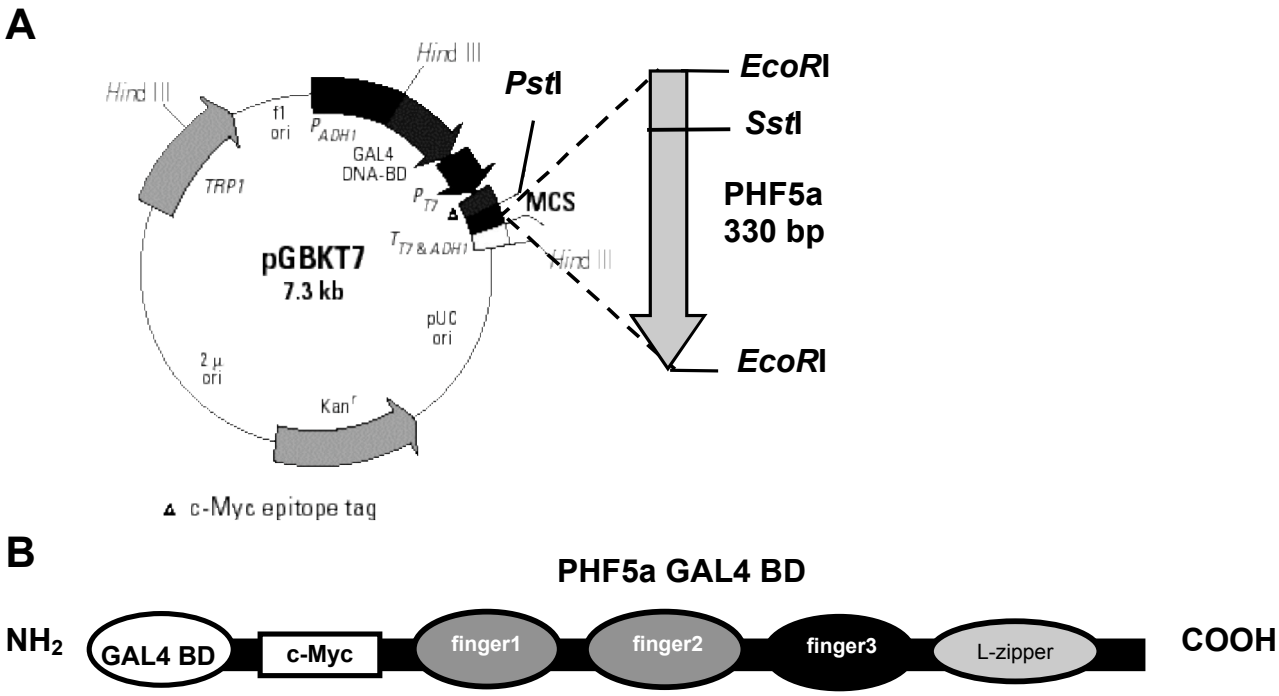


Figure 4. **A**. Schematic representation of the yeast two-hybrid bait construct BK/PHFA. The PHF5a fragment was amplified with primers pRT-21 FP and pRT-21 RP, purified, digested with *EcoR*I and cloned in frame to *EcoRI* site of pGBKT7 vector. The correct orientation was determined by double test digestions of resulting DNA clones with *Sst*I/*Pst*I. P_{ADH1} -truncated *S. cerevisiae* ADH1 promoter, GAL4 DNA-BD- GAL4 DNA binding domain, (P_{T7})- T7 RNA polymerase promoter, ($P_{T7\ ADH1}$)- transcription termination signals, (MCS)- Multiple Cloning Site, (pUC ori)- pUC plasmid replication origin, (Kan^r)- Kanamycin resistance gene, (2 µ ori)-

Yeast 2 μ replication origin, (TRP1)- TRP1 coding sequences, (f1 ori)- f1 bacteriophage origin of replication. Circular pGBKT7 vector map was adapted from pGBKT7 Vector Information, Protocol # PT3248-5, CLONTECH Laboratories, Inc. (1999), **B**. Schematic representation of the PHF5a GAL4 DNA-BD fusion protein. (GAL4 BD)- GAL4 DNA binding domain, (c-Myc)- c-Myc epitope tag, (finger1)- zinc finger, (finger2)- zinc finger 2, (l-zipper)- leucine zipper.

Additional tests were performed to exclude possible autoactivation of the reporter genes by the PHF5a-BD fusion protein alone. The *Saccharomyces cerevisiae* strain AH109 was transformed with the bait construct BK/PHF5a and served as a mating partner for the host strain Y187 pretransformed with the Matchmaker 11.5-days mouse embryo cDNA library. Both strains were combined and incubated for 24 hours before plating on the medium stringency selective SD plated (-Leu, -Try, - His). Following selection on medium stringency conditions, 200 resulting clones were tested on the SD medium (-Leu , -Try, -His, -Ade) and supplemented in X α-Gal (Figure 5). Approximately 80% of clones growing on high stringency plates turned blue as a result of MEL1 reporter activation. Its product _-galactosidase is a secreted enzyme and could by assayed directly on selective SD plates. About 100 positive clones (robust > 2 mm blue colonies) were picked and tested for the presence of a cDNA library insert by the colony PCR using pACT2 vector-specific primers Y2H2 and Y2HAD3'. Subsequently, DNA sequencing of 60 of these clones revealed that cDNAs from four genes were present multiple times as independent overlapping clones, namely U2AF^{35}, SRp40, Ddx1 and mDomino.

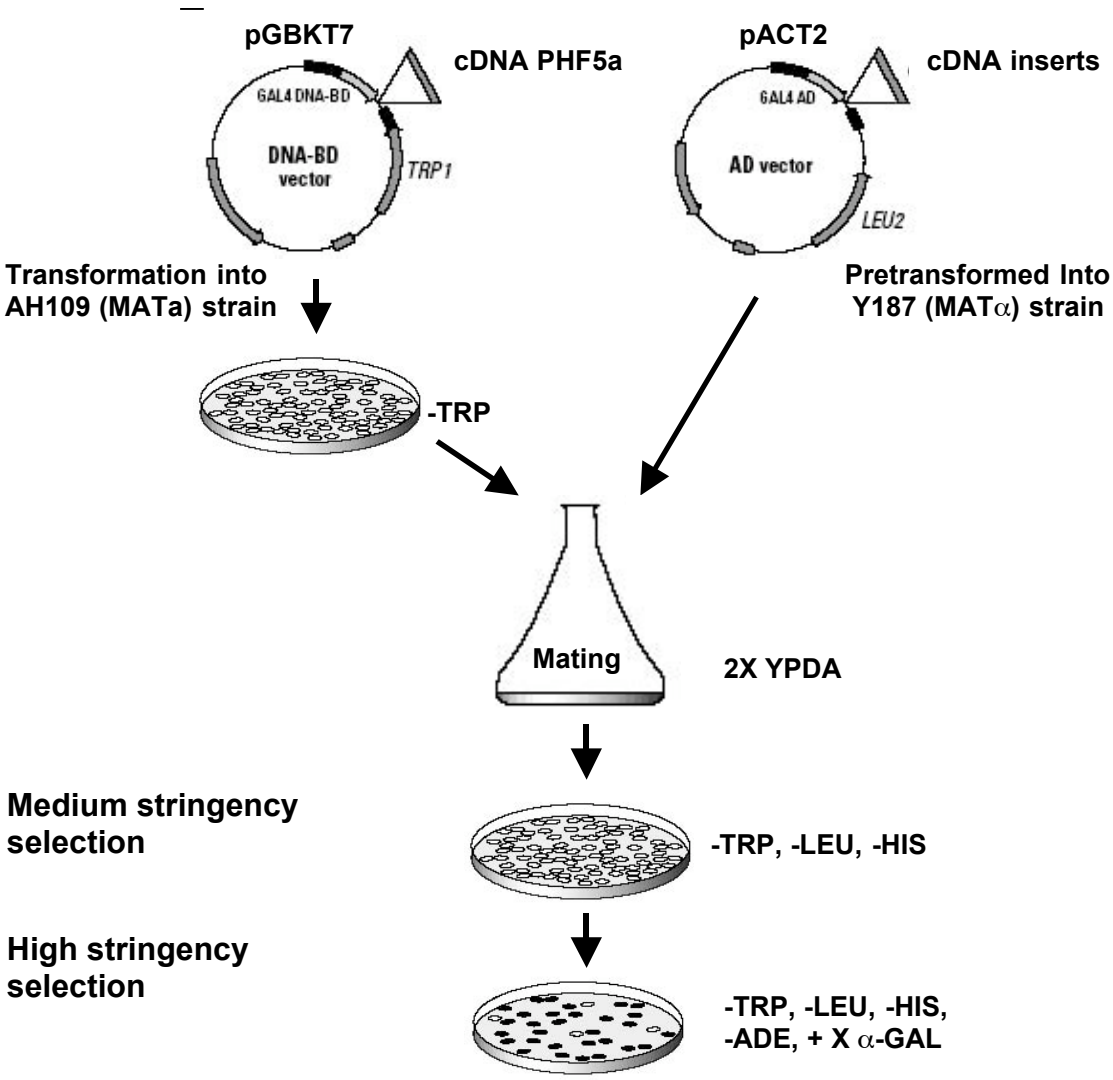
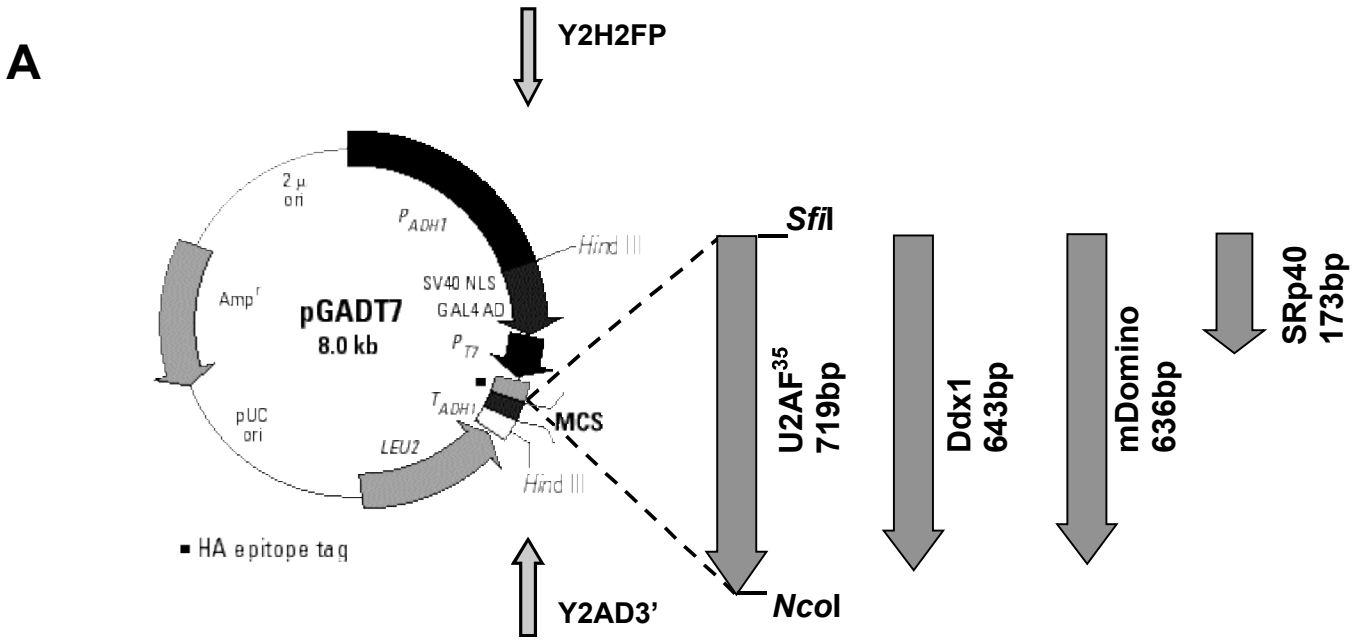
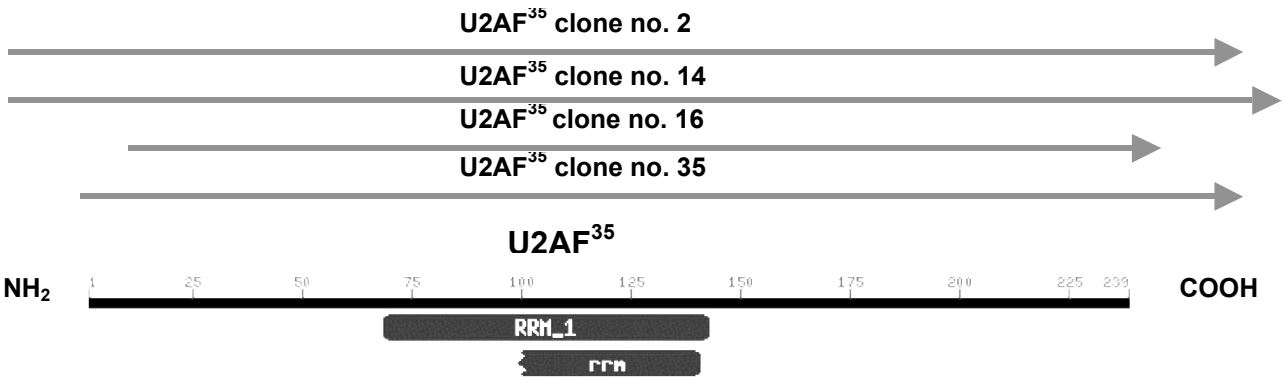


Figure 5. Schematic representation of the strategy used for the selection of PHF5a interacting proteins by the yeast two-hybrid embryo 11.5-days library screening. Picture partially adopted form the BD Matchmaker Pretransformed Libraries Users Manual, BD Biosciences Clontech, (1999), modified.

The following cDNA clones were selected for further work: clone no.35 containing the cDNA fragment of the splicing factor U2AF^{35} from nucleotide position 21-740 (Accession No.:NM_024187) corresponding to the complete ORF of U2AF^{35}; no. 94 with the shortest interacting fragment of the splicing factor SRp40 from nucleotide position 659-832 (Accession No.:NM_009159) spanning of the C-terminal region rich in arginine-serine (RS domain); no. 91 with the shortest interacting cDNA fragment of the ATP-dependent RNA helicase DEAD/H (Ddx1) from nucleotide position 1660- 2303 (Accession No.:NM_134040)

and no. 100 containing the shortest interacting cDNA fragment of the SWI2/SNF2-type ATPase/helicase protein mDomino-s from nucleotide position 1880-2516 (Accession No.: AB092695). The constructs: AD/SRp40-RS, AD/U2AF^{35}, AD/mDomino, and AD/Ddx1 were prepared by digestion with *Nco*I/*Sfi*I of colony PCR products obtained from yeast colonies and cloning in frame into the pGADT7 vector (Figure 6).





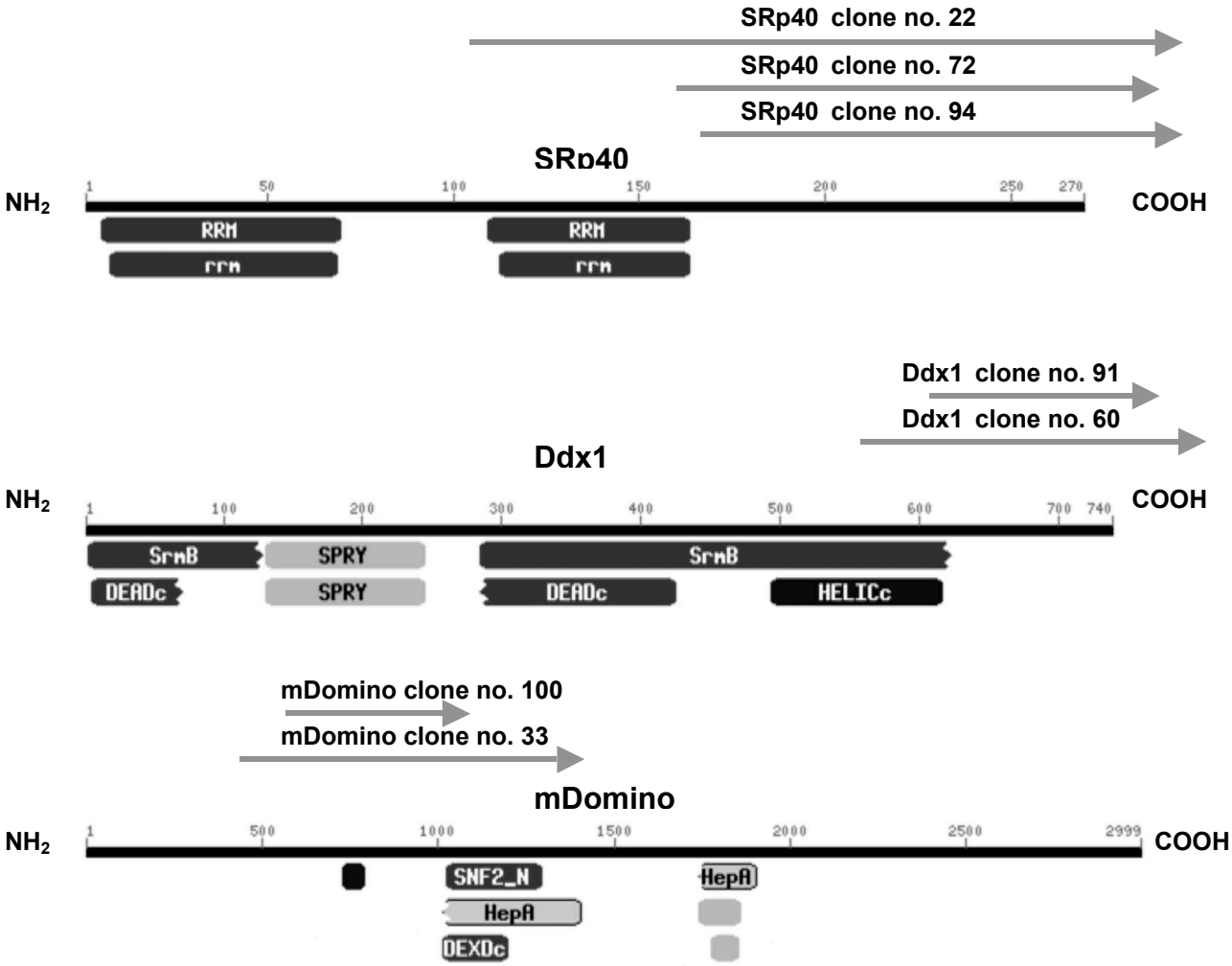


Figure 6. **A**. Schematic representation of pGADT7 constructs containing cDNA fragments isolated by the yeast two-hybrid assay. The yeast colony-PCR products generated with Y2H2 FP and Y2AD3' primers from the yeast clone 35 (U2AF^{35}), clone 91(Ddx1), clone 100 (mDomino) and clone 94 (SRp40) were purified, double digested with enzymes *Nco*I/*Sfi*I and cloned in frame into the pGADT7 vector. (P_{ADH1})- Full-length S. cerevisiae ADH1 promoter, (SV40 NLS GAL4 AD)- GAL4 acceptor domain polypeptide with SV40 Nuclear Localization Signal, (P_{T7})- T7 RNA polymerase promoter, (HA)- HA epitope tag, (MCS)- Multiple Cloning Sites, (T_{ADH1})- Transcription termination signal, (LEU2)- LEU2 coding sequences, (pUC ori)- pUC plasmid replication origin, (Amp^r)- Ampicillin resistance gene, (2 μ ori)- Yeast 2 μ replication origin. Circular pADT7 vector map adapted from pGADT7 Vector Information, Protocol # PT3249-5, CLONTECH Laboratories, Inc., 1999. **B.** Schematic representations of domain structures of U2AF^{35}, SRp40, Ddx1 and mDomino with annotated corresponding cDNA clones found in the yeast two-hybrid library screening (according to the NCBI Conserved Domain Search v. 2.00). (RRM)- RNA recognition motif, (SrmB)- Superfamily II DNA and RNA helicases, (DEADc)- DEAD-box helicases, (SPRY)- Domain in SPla and the RYanodine Receptor, (HELICc)- Helicase superfamily c-

terminal domain, (DEXH)- DEXH-box helicases, (SNF2_N)- SNF2 family N-terminal domain, (HepA)- Superfamily II DNA/RNA helicases

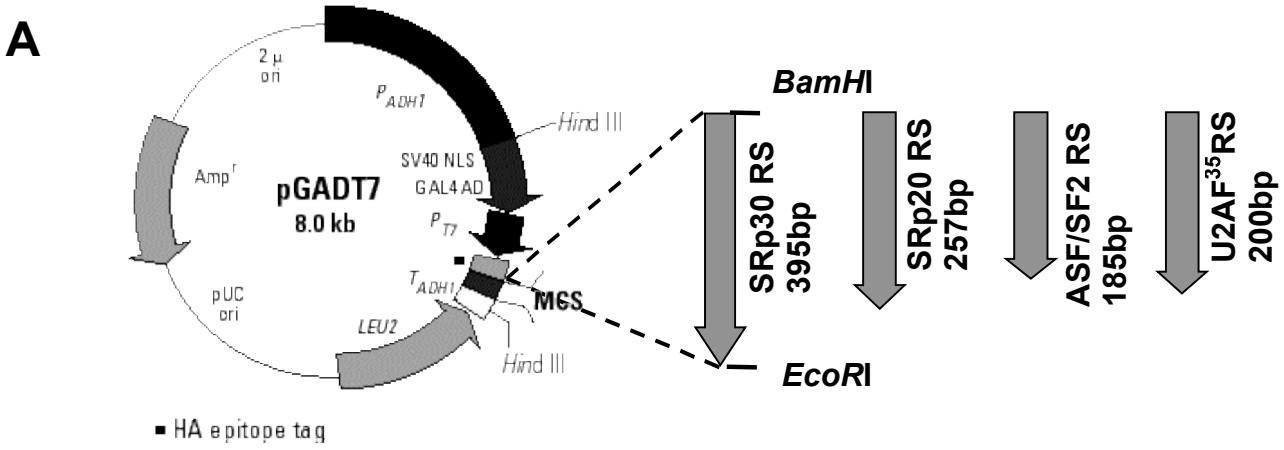
Three of the identified interaction partners ($U2AF^{35}$, SRp40 and Ddx1) consist of either domains characteristic for RNA processing proteins or were previously described to be involved in this process (Zhang et al., 1992; Screaton et al., 1995; Bléoo et al., 2001). Non-snRNP protein splicing factors SRp40 and $U2AF^{35}$ contain domains rich in alternating serine and arginine residues (RS domains). Clone no. 94 encoding the C-terminal RS domain of the SRp40 protein was sufficient to maintain the interaction with the PHF5a protein, therefore, construct encoding the RS domain of $U2AF^{35}$ was generated. For this purpose $U2AF^{35}$ cDNA fragment from nucleotide position 540-740 was amplified by RT-PCR on total RNA extracted from murine 11.5-days embryos using primers U2AF-RS FR, U2AF-RS RP and cloned in frame into the pGADT7 vector.

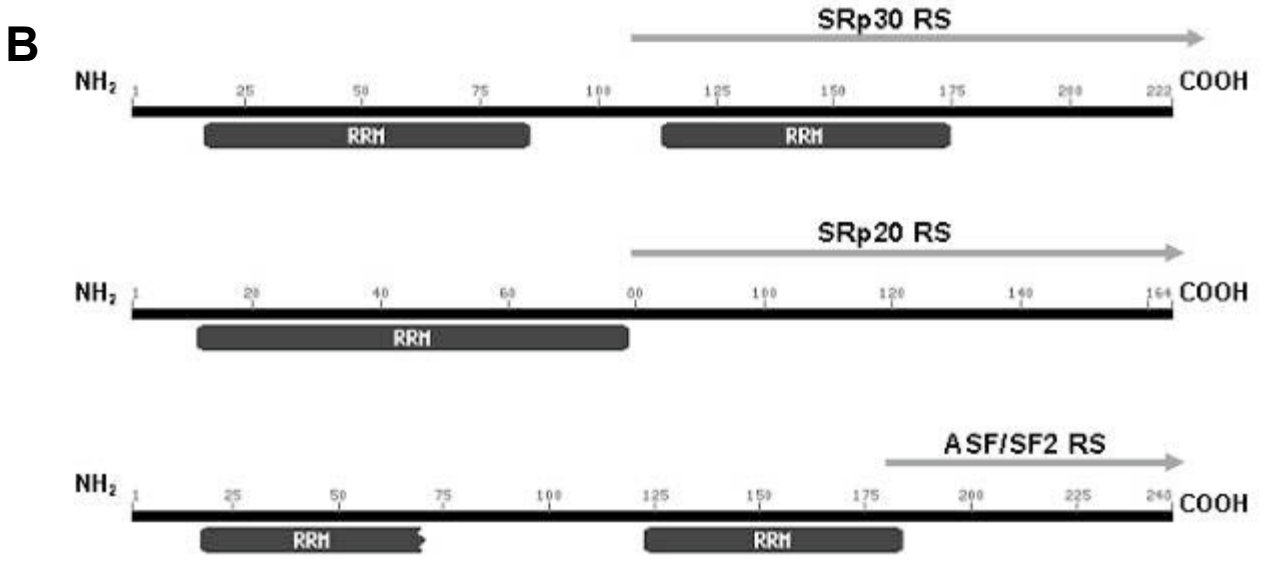
Furthermore, yeast cotransformants with the PHF5a protein as a pray and the RS domain of $U2AF^{35}$ as a bait could grow on selective SD plates (-Leu, -Try, -His, -Ade). This result suggests that PHF5a would act as an arginine-serine rich (RS) domain binding protein. Ddx1 and mDomino are ATP-dependent helicases and Ddx1 is believed to be involved in RNA processing (Godbout et al., 1993; Bléoo et al., 2001)

3.1.2 Determination of PHF5a binding affinity to different RS domains by the quantitative _-galactosidase assay

Double transformants of the bait construct and constructs containing library founders resulted in blue colonies as a result of MEL1 activation. The product of the MEL1 gene, namely α-galactosidase, could be quantitatively measured in the yeast liquid medium and therefore, it allows for directly comparing the relative strength of PHF5a interactions with different proteins.

In order to answer the question, whether PHF5a interacts unspecifically with all RS domain-containing proteins or if PHF5a binding is rather restricted to proteins found in the yeast two-hybrid library screening, fragments containing RS domains of three other SR proteins were cloned and tested. For that purpose, constructs encoding RS domains of SRp20 (Accession No.: NM_013663) containing a cDNA fragment from nucleotide position 355-612; SRp30c (Accession No.:NM_025573) containing a cDNA fragment from nucleotide position 406-801; ASF/SF2 (Accession No.: NM_173374) containing a cDNA fragment from nucleotide position 697-882 and U2AF^{35} containing a cDNA fragment from nucleotide position 540-740 were generated by RT-PCR, and cloned in frame into pGADT7 vector (Figure 7).





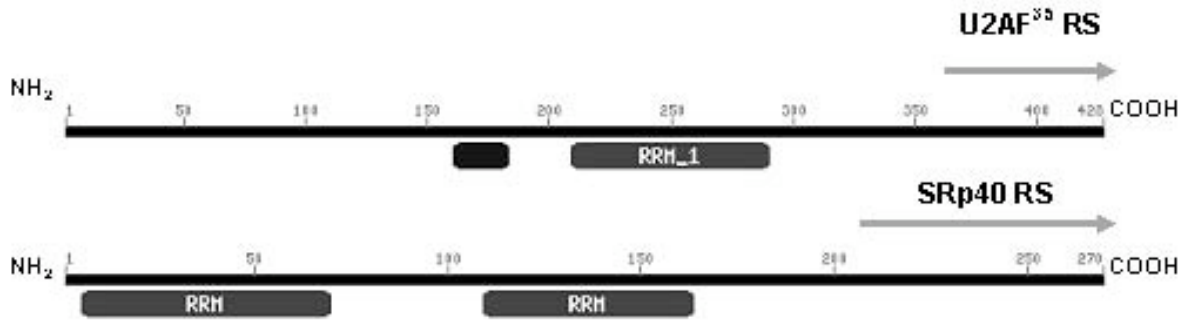


Figure 7. **A**. Schematic representation of pGADT7 constructs containing RS domains of splicing factors SRp30, SRp20, ASF/SF2 and U2AF^{35} used in the quantitative _-galactosidase assay. cDNA fragments corresponding to the RS domain of SRp30, SRp20, ASF/SF2 and U2AF^{35} were amplified by the one-step RT-PCR using primers RS-SRp30 FP, RS-SRp30 RP; RS-SRp20 FP, RS-SRp20 RP; RS-ASF/SF2 FP, RS-ASF/SF2 RP and RS-U2AF FP, RS-U2AF RP, respectively. RT-PCR products were purified, digested with *EcoR*I/*BamH*I and cloned in frame into the pGADT7 vector. **. B.** Schematic representation of domain structure of proteins containing RS domains namely: SRp30, SRp20, ASF/SF2, U2AF^{35} and SRp40. Localisations of the corresponding pGADT7 clones containing RS domains are annotated as grey lines. Domain architecture according to the NCBI Conserved Domain Search v. 2.00.

The BK/PHFa bait construct was cotransformed together with constructs containing RS domains of following splicing factors: ASF/SF2, SRp30 and SRp20. Subsequently, yeast cotransformants were tested on selective SD plates (-Leu, -Try, -His, -Ade). Yeast cells containing the bait PHF5a-BD construct and the SRp30-AD construct could not grow on these plates. Cotransformants bearing the bait PHF5a-BD fusion protein and ASF/SF2-AD or SRp20-AD fusion constructs could form colonies, but they did not turn blue in the presence of x-α Gal.

The quantitative α-galactosidase assay was implemented to compare the strength of PHF5a interaction with different RS domains (Figure 8). The galactosidase activities in the yeast clones bearing the PHF5a bait and RS domains of ASF/SF2 and SRp20 were almost undetectable, whereas the strongest activity was measured for the DEAD/H (Ddx1) helicase

and U2AF^{35}. Taken together, these results indicate that protein interactions between PHF5a and ASF/SF2, SRp20, SRp30 are either very weak or are not detectable in this assay.

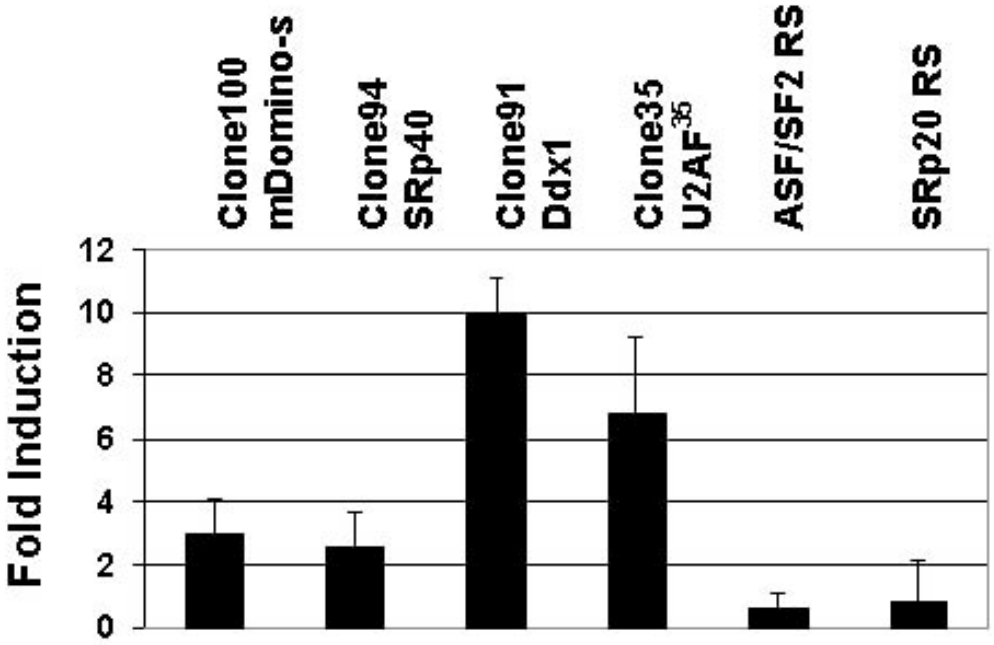


Figure 8. The α-Gal quantitative assay shows the relative strength of interactions between PHF5a and splicing proteins. The bars show the fold induction of α-galactosidase activity as compared to empty vectors.

3.1.3 Coimmunoprecipitation of the PHF5a protein with interaction partners

The coimmunoprecipitation assay was performed to answer the question whether in vitro protein-protein interactions would corroborate the interactions found by the yeast two-hybrid assay. For this experiment, pGADT7 constructs coding hemagglutinin (HA) tagged versions of SRp40, U2AF^{35}, Ddx1 and mDomino were used. Recombinant [^{35}S] labelled proteins were produced by the *in vitro* transcription-translation experiment as described in Materials and Methods. These proteins and their putative binding partner, i.e. the c-Myc tagged PHF5a protein were used in the immunoprecipitation assay.

After 1 hour incubation at room temperature the c- Myc tagged PHF5a protein and all four proteins found by the yeast two-hybrid assay and expressed as a fusion protein with a HA tag could be efficiently coimmunopreciptated. This interaction was detectable in all cases using both combinations, i.e. by using either c-Myc or HA antibodies recognizing adequate epitopes. There was no unspecific precipitation of the HA tagged proteins by using the c-Myc antibody and the c-Myc tagged proteins by using the HA antibody (Figure 9 and Figure 10).

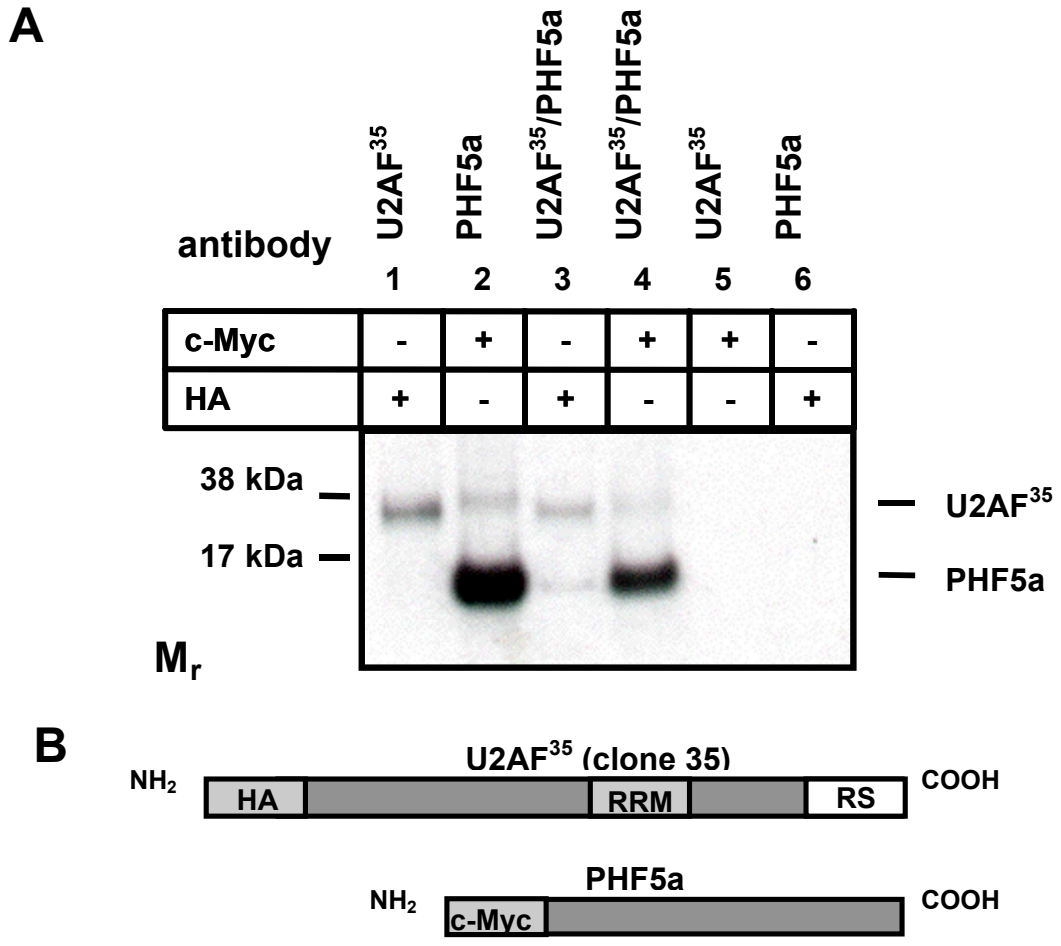


Figure 9. A. SDS-PAGE analysis showing that the murine PHF5a protein coimmunoprecipitates with the U2AF^{35} protein. The proteins were *in vitro* transcribed/translated from the constructs BK/PHF5a (c-Myc epitope) and AD/U2AF^{35} (HA epitope). PHF5a-c-Myc and U2AF^{35}-HA radioactively ^{35}S labeled fusion products were generated. After immunoprecipitation and elution from the Protein A beads, 10µl of the immunoprecipitate was loaded onto 4-12% SDS polyacrylamide gel. **Lane 1**: U2AF^{35}+ HA antibody; **Lane 2**: PHF5a + c-Myc antibody (note unspecific band at ~38 kDa); **Lane 3**: U2AF^{35}+ PHF5a + HA antbody; **Lane 4**: U2AF^{35} + PHF5a

+ c-Myc antibody; **Lane 5**: negative control U2AF^{35}+ c-Myc; **Lane 6**: negative control PHF5a + HA antibody. **B.** Schematic view of recombinant fusion proteins used in the coimmunoprecipitation assay.

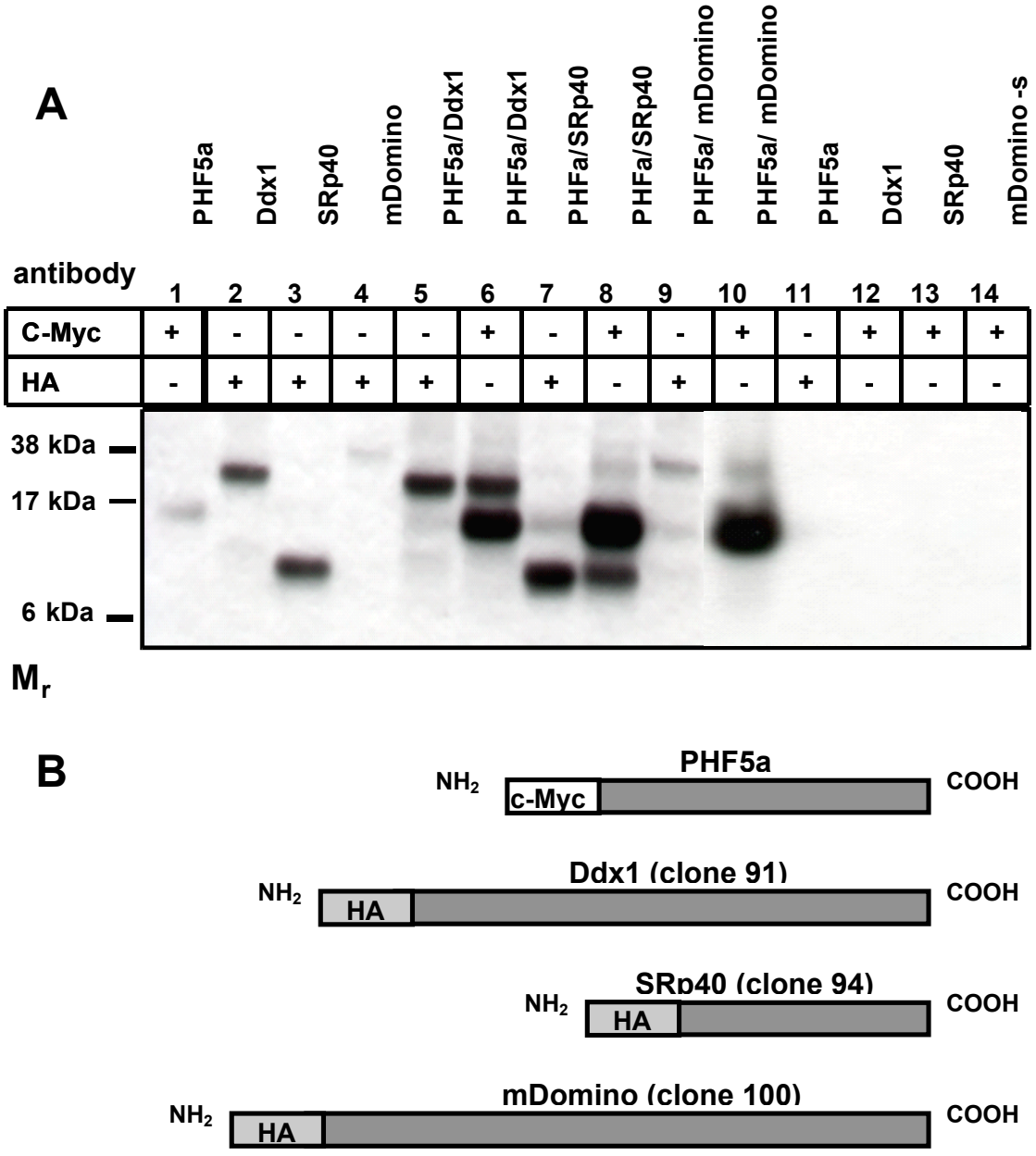


Figure 10. **A.** SDS-PAGE analysis showing that the murine PHF5a protein coimmunoprecipitates with Ddx1, SRp40 and mDomino-s**.** The ^{35}S labeled proteins with a HA epitope were produced from the constructs AD/Ddx1-RS, AD/SRp40-RS, and AD/mDomino-RS, respectively. After immunoprecipitation and elution from Protein A beads, 10 µl of the immunoprecipitate was loaded onto 4-12% SDS polyacrylamide gel. **Lane 1**: PHF5a + c-Myc antibody; **Lane 2**: Ddx1-RS + HA anibody; **Lane 3**: SRp40-RS + HA antibody; **Lane 4**: mDomino-RS + HA antibody; **Lane 5**: PHF5a + Ddx1-RS + HA antibody; **Lane 6**: PHF5a + Ddx1-RS + c-Myc antibody; **Lane 7**: PHF5a + SRp40-RS + HA antibody; **Lane 8**: PHF5a + SRp40-RS + c-Myc antibody; **Lane 9**:

PHF5a + mDomino-RS + HA antibody; **Lane 10**: PHF5a + mDomino-RS + c-Myc antibody; **Lane 11**: negative control PHF5a + HA antibody; **Lane 12**: negative control Ddx1-RS + c-Myc antibody; **Lane 13**: negative control SRp40-RS + c-Myc antibody; **Lane 14**: negative control mDomino-RS + c-Myc antibody. **B.** Schematic view of recombinant fusion proteins used in the coimmunoprecipitation assay.

In order to determine the binding specificity of PHF5a and U2AF^{35}, coimmunoprecipitation assay in buffers containing different ionic strength was performed. The complex of the PHF5a protein with U2AF^{35} was even stable after extensive washing in a wide range of increasing NaCl concentrations in the washing buffer, suggesting a rather hydrophobic than ionic interaction between the PHF5a protein and the RS domain containing protein U2AF^{35} (Figure 11).

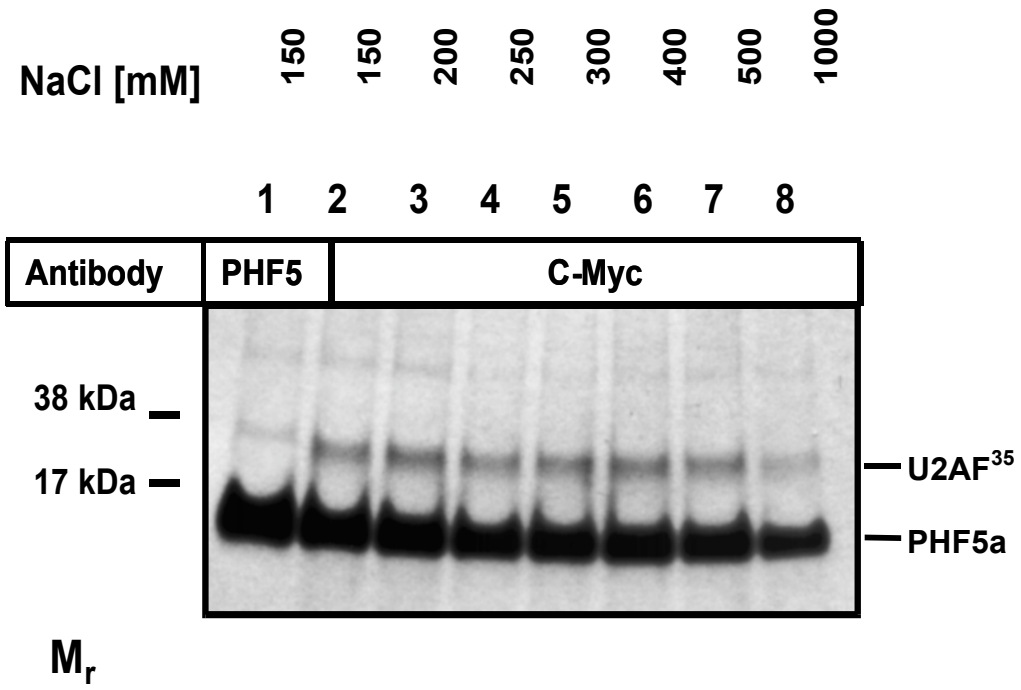


Figure 11. SDS-PAGE analysis demonstrating that the interaction between the PHF5a protein and the U2AF^{35} protein examined by the immunoprecipitation assay is salt-independent. **Lane 1**: PHF5a-c-Myc protein immunoprecipitated with PHF5a antibody; **Lane 2- 8**: PHF5a + U2AF^{35} + c-Myc antibody together with the increasing gradient of NaCl in the washing buffer.

3.2 Mapping of PHF5a interaction domains

In order to map interaction domains of the PHF5a protein, several PHF5a deletion constructs were prepared and tested for the ability to bind to the U2AF^{35} protein in the immunoprecipitation assay. Both the C-terminal part of U2AF^{35}Δ1-172 (fragment from amino acid position 173 to 239) containing the RS domain and the HA epitope were expressed using the construct AD/U2AF^{35}- RS. From different segments of PHF5a, only the truncated protein PHF5aΔ81-110, but not PHF5aΔ1-28,81-110 and PHFaΔ1-28 could be effectively coprecipitated with the RS domain of U2AF^{35} by using HA antibodies (Figure 12). These results suggest that the N-terminal part of PHF5a is responsible for the interaction with the RS domain of the U2AF^{35} protein. Two putative zinc fingers are localized in this region of the PHF5a protein and possibly folds in the ring finger domain (Oltra et al., 2003) responsible for protein-protein interactions.

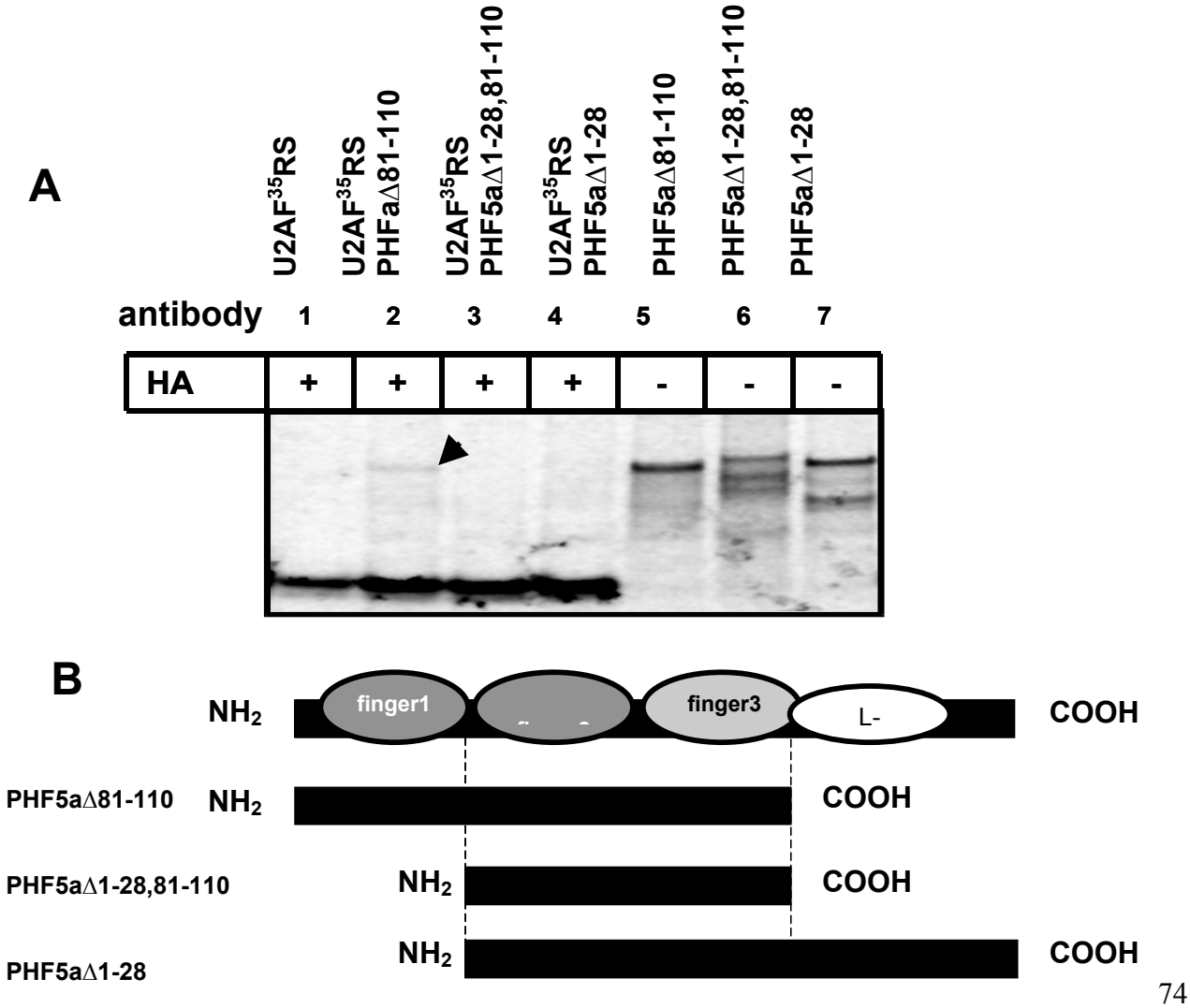


Figure 12. **A.** SDS-PAGE analysis demonstrating that truncated murine PHF5a protein containing the N-terminal part interacts with the arginine-serine rich domain of the $U2AF^{35}$ protein. *In vitro* translated proteins $U2AF^{35}$RS-HA, PHFaΔ81-110, PHF5aΔ1-28,81-110, PHF5aΔ1-28, were generated from constructs AD/$U2AF^{35}$-RS, GST/PHF5aΔ1-28, GST/PHF5aΔ1-28,81-110, GST/PHF5aΔ1-28, respectively. 10µl of the immunoprecipitate was loaded onto a 4-12% SDS polyacrylamide gel. **Lane 1**: $U2AF^{35}$RS-HA + HA antibody; **Lane 2**: $U2AF^{35}$RS-HA + PHFaΔ81-110 + HA antibody; **Lane 3**: $U2AF^{35}$RS-HA + PHF5aΔ1-28,81-110 + HA antibody; **Lane 4**: $U2AF^{35}$RS-HA + PHF5aΔ1-28 + HA; 10µl of the In vitro translated proteins 1-3 PHF5a-HA, 2-3 PHF5a-HA, 2-4 PHF5a-HA were loaded in **lanes 5-7**, respectively. **B.** Schematic view of the truncated fragments of the PHF5a protein used for mapping of protein interaction domains in the immunoprecipitation experiment.

To determine the PHF5a interaction domains more precisely, a directed yeast two-hybrid assay was employed (Figure 13). For this purpose, minimal interacting fragments of $U2AF^{35}$, SRp40, Ddx1 and mDomino-s were used for cotransformation experiments with different segments of PHF5a protein. Plasmids coding different peptide segments of PHF5a: PHFaΔ81-110 (from nucleotide position 40-284); PHF5aΔ1-28,81-110 (from nucleotide position 129-284); PHF5aΔ1-28 (from nucleotide position 129-371); PHF5aΔ38-110 (from nucleotide position 40-149); PHF5aΔ58-110 (from nucleotide position 40-215); PHF5aΔ1-50 (from nucleotide position 195-371); PHF5aΔ1-80 (from nucleotide position 285-371) were generated by the PCR and cloned into the pGBKT7 and pET 41 vector (Novagen), respectively.

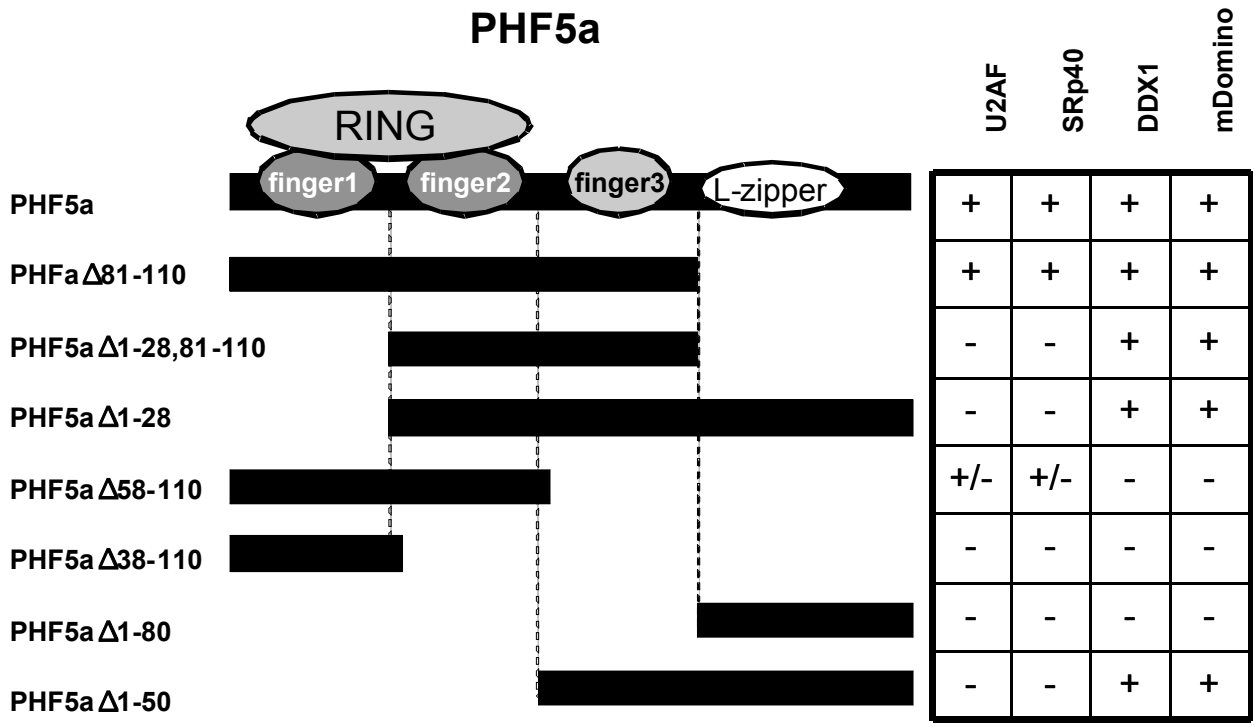


Figure 13. Schematic view of the truncated fragments of the PHF5a protein used for mapping of protein interaction domains in the directed yeast two-hybrid experiments. All interactions between proteins U2AF[35], SRp40, Ddx1, mDomino and different segments of the PHF5a protein were detected by the yeast-two hybrid assay and are summarized in the table.

PHF5a deletion mutants Δ81-110 and possibly PHF5a Δ57-110 could associate only with RS domains of splicing proteins U2AF[35] and SRp40. Cotransformants bearing RS domains with isolated first two zinc fingers of PHF5a (PHF5aΔ57-110) did not turn blue in the presence of x _-Gal, suggesting that the interaction with the first three zinc fingers is preferred under these conditions.

In contrast, the mapping results of PHF5a with both Ddx1 and mDomino indicate, that these helicases associate only with the C-terminal part of PHF5a. The minimal PHF5a fragment interacting with helicases can be restricted to the third zinc finger, which is present in all PHF5a fragments interacting with ATP-dependent helicases.

Trappe et al. (2001) suggested another model of folding for the PHF5a protein. In this model the deduced protein sequence of PHF5a shows a central PHD-finger like motif flanked by two highly basic domains. In order to exclude the possibility that protein interactions with PHF5a are mediated by ionic attraction of basic domains, mutagenesis studies were performed using the quick change kit (Stratagene) and subsequently, the coimmunoprecipitation assay. Two basic amino acids from the N-terminal part of PHF5a protein were substituted using primers MUT5' FP and MUT5'RP and two from the C-terminal part were substituted using primers MUT3'FP, MUT3'RP. The resulting amino acid substitutions were as follows: lysine at amino acid position 3 to isoleucine (Lys^{3}_Ile), histidine at amino acid position 4 to leucine (His^{4}_Leu), arginine at amino acid position 102 to proline (Arg^{102}_Pro) and lysine at amino acid position 103 to isoleucine (Lys^{103}_Ile) (Figure 14).

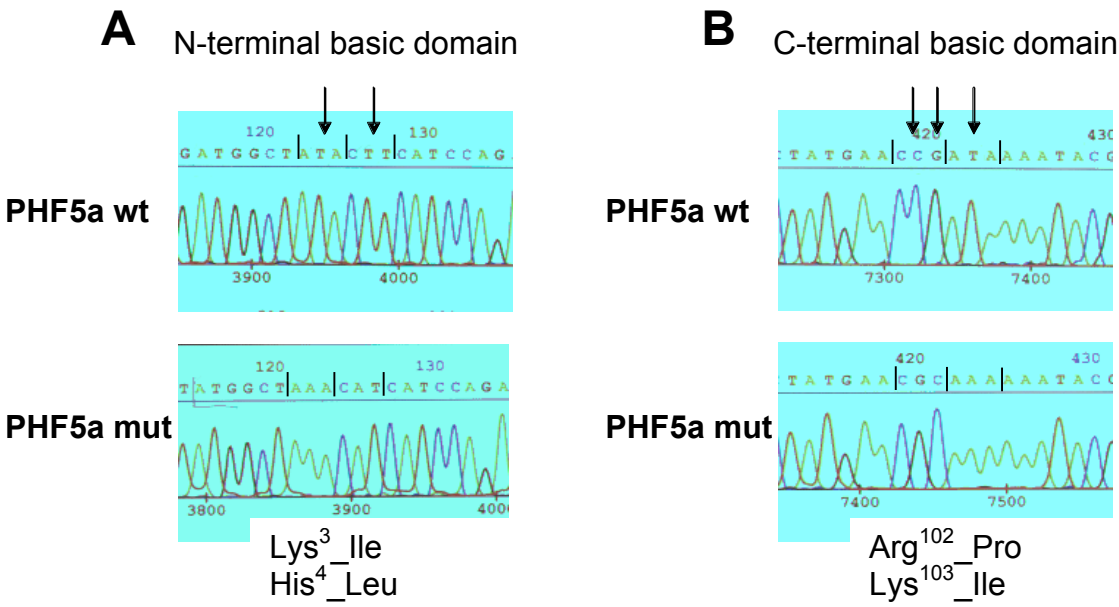


Figure 14. Sequencing results of the pGBKT7 clones containing newly introduced mutations changing terminal basic regions of the PHF5a protein. **A**. Sequence of the N-terminal part of the PHF5a clone with introduced substitutions changing Lys^{3}_Ile and His^{4}_Leu. **B**. Sequence of the C-terminal part of the PHF5a with substitutions Arg^{102}_Pro and Lys^{103}_Ile.

Recombinant mutated PHF5a proteins were produced and subsequently tested in the *in vitro* precipitation assay if they are able to bind effectively to the recombinant U2AF^{35} protein containing a long stretch of acidic glutamines at the C-terminal end. Both PHF5a proteins with mutated basic domains could be efficiently coprecipitated with the U2AF^{35} protein (Figure 15) suggesting that the basic amino acid regions are not essential in maintaining this interaction.

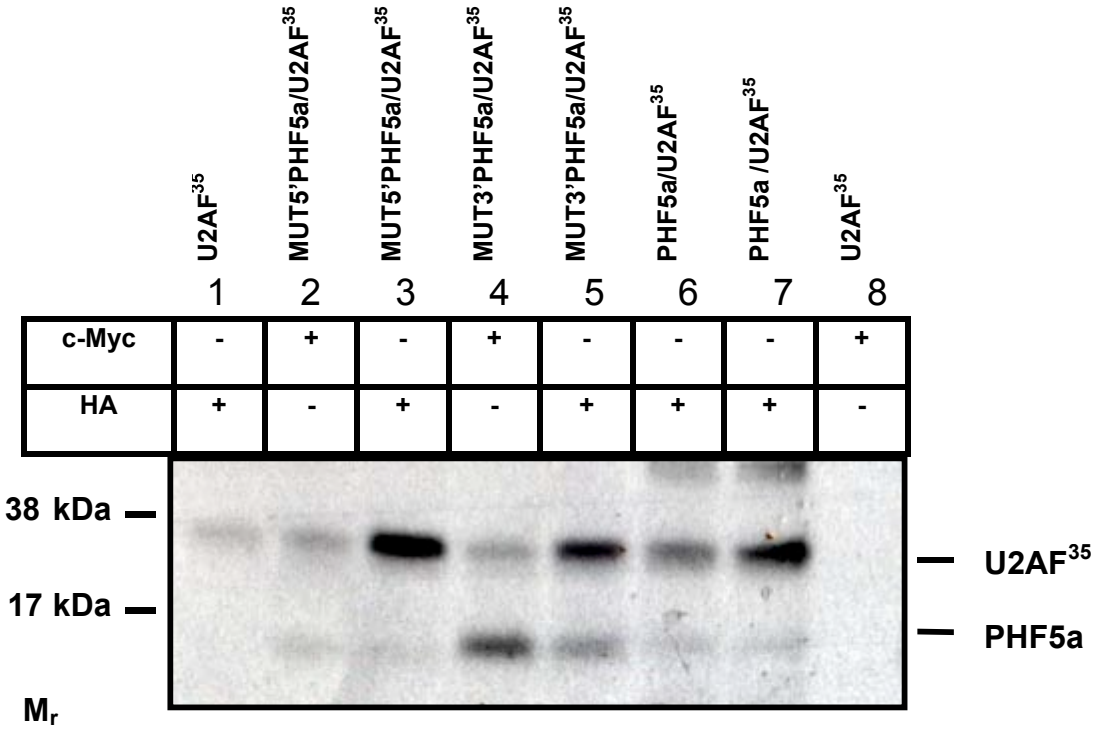


Figure 15. SDS-PAGE analysis showing that the terminal basic regions of the PHF5a protein do not influence binding to the U2AF^{35} protein. *In vitro* translated proteins U2AF^{35} –HA, MUT5'PHF5a-cMyc, MUT3'PHF5a-cMyc and PHF5a –cMyc were produced from the constructs AD/U2AF^{35}, BK/MUT5'PHF5a, BK/MUT3'PHF5a and BK/PHF5a, respectively. The immunoprecipitate was loaded onto a 4-12% SDS polyacrylamide gel: **Lane 1**. U2AF^{35} –HA+ HA antibody; **Lane 2**. MUT5'PHF5a-cMyc + U2AF^{35} –HA + c-Myc antibody; **Lane 3.** MUT5'PHF5a-cMyc + U2AF^{35} –HA + HA antibody; **Lane 4.** MUT3'PHF5a-cMyc + U2AF^{35} –HA + c-Myc antibody; **Lane 5.** MUT3'PHF5a-cMyc + U2AF^{35} –HA + HA antibody; **Lane 6, 7**. Positive control PHF5a-cMyc + U2AF^{35}-HA + HA antibody. Lane 8. Negative control U2AF^{35}-HA + c-Myc antibody.

3.3 PHF5a functions as a bridge protein between helicases and RS proteins

The N-terminal part of the PHF5a protein was identified as the binding site to RS domains of splicing factors U2AF^{35} and SRp40 and the C-terminal zinc finger was determined as the binding site responsible for binding to ATP-dependent helicases Ddx1 and mDomino. In order to answer the question, if the PHF5a protein is able to bind RS proteins and helicases simultaneously, the yeast three-hybrid assay was performed.

For this experiment yeast cells were transformed with the bicistronic pBridge vectors expressing constitutively U2AF^{35} or SRp40 as a binding domain (BD) fusion protein (Figure 16). The second protein, PHF5a was expressed under the control of the conditional promoter P^{Met25}. Subsequently, pBridge constructs were used in conjunction with pGADT7 constructs encoding either Ddx1 or mDomino as an activation domain (AD) fusion protein. PHF5a was conditionally expressed from the P^{Met25} promoter only in yeast cells which were growing on selective SD plates in the absence of methionine (Figure 17).

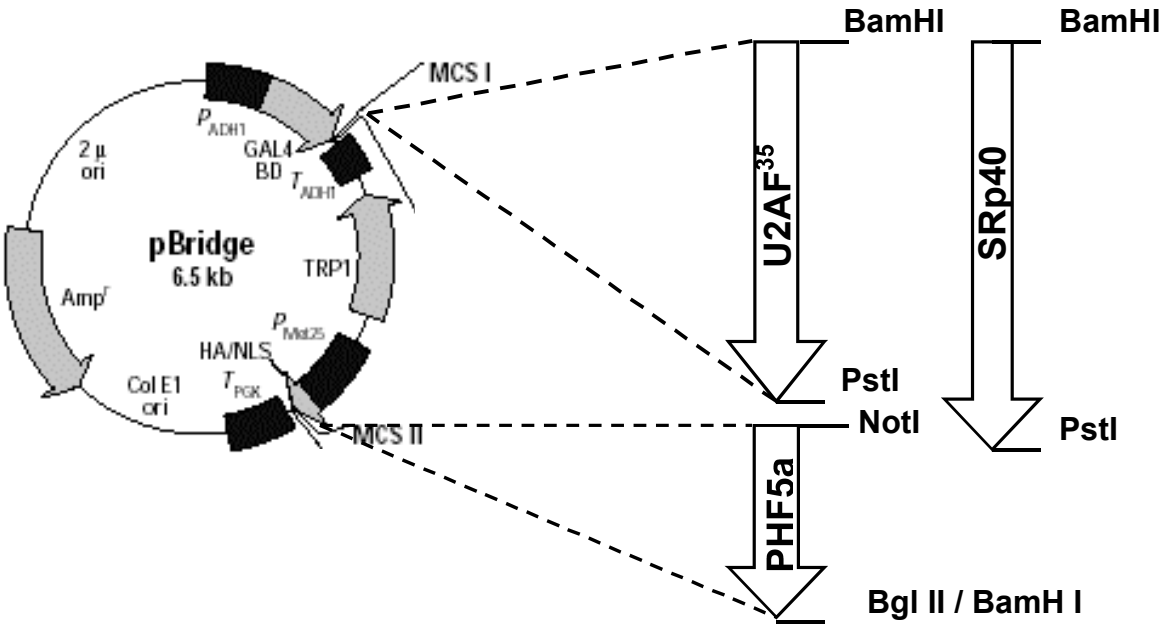
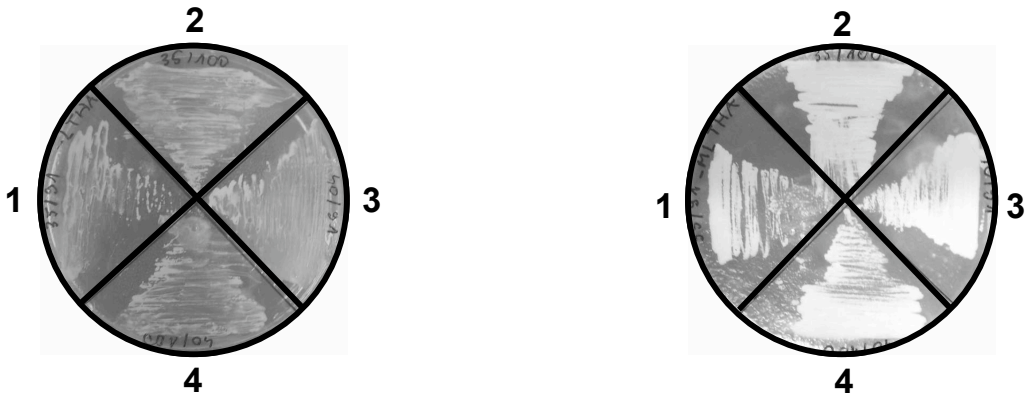


Figure 16. Schematic representation of bicistronic pBridge constructs used in the yeast three-hybrid assay. Two different bridge constructs were prepared and used in conjunction with pGADT7 constructs. The pBridge

construct bearing both open reading frames (ORFs) of U2AF^{35} and PHF5a was generated as follows: the ORF of U2AF^{35} was amplified by RT-PCR using primers U2AF^{35} FP, U2AF^{35} RP digested with *BamH*I and *Pst*I and cloned in frame to the MCSI of the pBridge vector. The ORF of PHF5a was prepared from the pEGFP-N1 construct by double digestion (*Not*I/*BamH*I) and cloning into the *Not*I/*Bgl*II sites of MCSII. Note that the *Bgl*II site was lost during the cloning of the *BamH*I digested PHF5a fragment. The second pBridge construct bearing SRp40 and PHF5a was prepared as follows: the ORF of SRp40 was generated by RT-PCR using primers SRp40 FP and SRp40 RP and cloned in frame into the *BamH*I/*Pst*I sites of MCSI. The PHF5a fragment was generated and cloned as described above for the first pBridge construct. (P_{ADH})- fragment containing the *S. cerevisiae* ADH1 promoter, (GAL4 BD)- GAL4 DNA-binding domain polypeptide, (MCSI)- Multi Cloning Site I, (T_{ADHI})- Transcription termination signal, (TRP1)- TRP1 coding sequence, P_{MET25}- *S. cerevisiae* MET25 promoter, (HA/NLS)- HA epitope and nuclear localization sequence, (MCS II)- Multicloning Site II, (T_{PGK})- *S. cerevisiae* PGK terminator, (Col E1 ori)- Col E1 origin of replication, (Amp^r)- Ampicillin resistance gene (b-lactamase), (2 μ ori)– Fragment containing the 2 μ origin of replication. The pBridge circular vector map adopted from pBridge Vector Information, CLONTECH Laboratories, Inc., 1999.

Continuous growth of double yeast transformants was observed only on selective SD plates lacking methionine, tryptophan, leucine, adenine and histidine, while addition of methionine at concentration of 1mM could arrest the yeast growth (Figure 17). This strongly indicates that PHF5a may act as a bridge protein between U2AF^{35} and SRp40 from the one site and ATP-dependent helicases Ddx1 and mDomino-s from the other.



1mM methionine

	BD fusion	P^{Met25} PHF5a	AD fusion
1	$U2AF^{35}$	-	Ddx1
2	$U2AF^{35}$	-	mDomino
3	SRp40	-	Ddx1
4	SRp40	-	mDomino

	BD fusion	P^{Met25} PHF5a	AD fusion
1	$U2AF^{35}$	+	Ddx1
2	$U2AF^{35}$	+	mDomino
3	SRp40	+	Ddx1
4	SRp40	+	mDomino

Figure 17. PHF5a can act as a bridge protein between RS domain-containing proteins: $U2AF^{35}$, SRp40 and ATP-dependent helicases: Ddx1, mDomino in the yeast three-hybrid assay. Yeasts cells transformed with pBridge and pGADT7 constructs (summarized in the table) were streaked on selective SD plates (1mM methionine) –Leu, -Try, -Ade, -His (left) and -Met -Leu, -Try, -Ade, -His, (right). Continuous growth of yeast transformants bearing splicing proteins BD-fusions and helicases AD-fusions was observed only in the absence of methionine which allows expression of the PHF5a "bridge" protein (controlled by the methionine repressor P^{Met25}).

3.4 Subcellular localization of the PHF5a

In order to determine subcellular localization of PHF5a protein NIH3T3 mouse fibroblast cells were stained with an affinity purified anti-PHF5a polyclonal antibody (Figure 18). Signals for PHF5a protein were detected predominantly within the nucleus of stained cells. Very weak staining for PHF5a protein was visible also in the cytoplasm adjacent to the nuclear envelope of NIH3T3 cells. Pattern present in the nucleus is characterized by large

speckles covering the whole nucleoplasm, interconnected by a fine fluorescent network. The chromosome region of mitotic cells arrested in the metaphase is not stained (Figure 19).

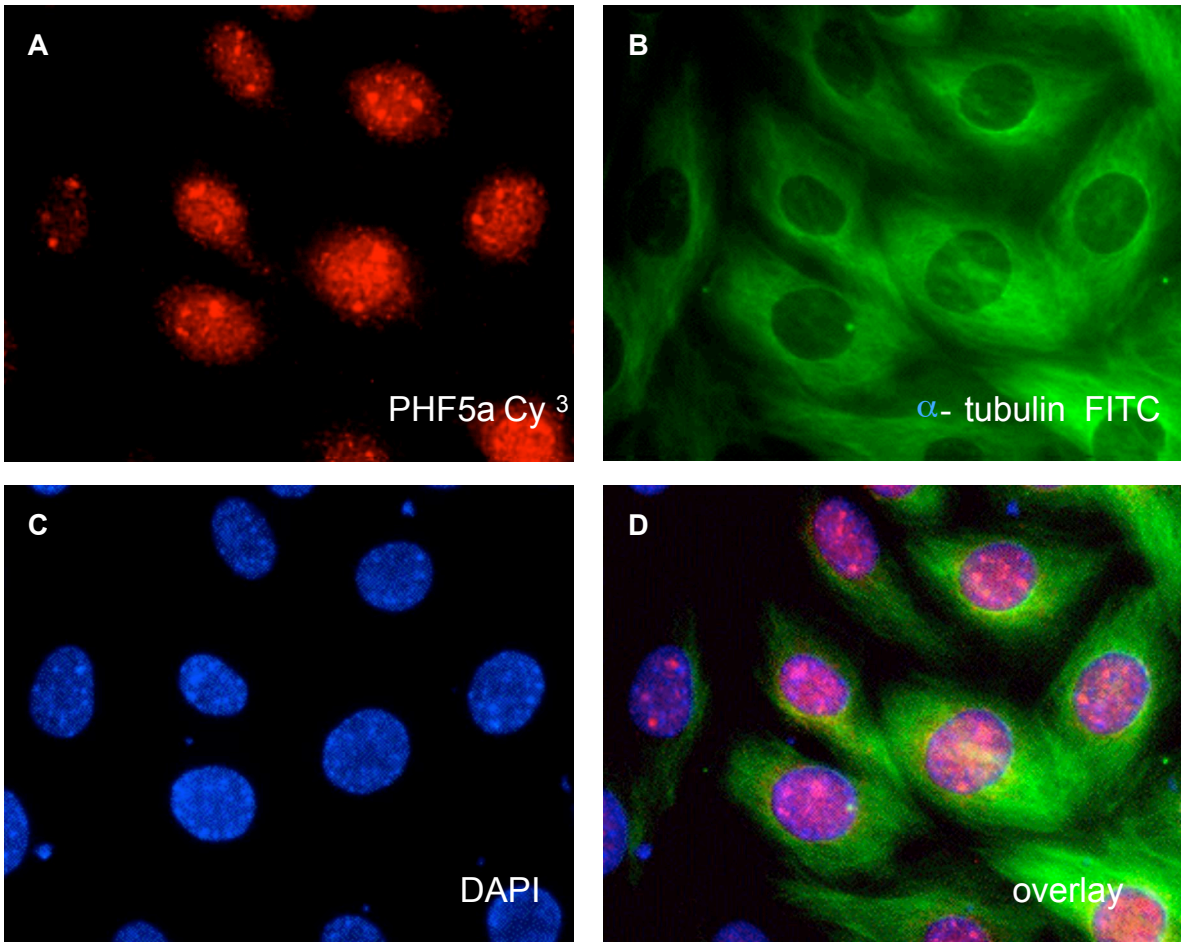


Figure 18. Subcellular localization of the PHF5a protein in NIH 3T3 cells observed under fluorescent microscope. **A**. NIH 3T3 cells stained with polyclonal anti-PHF5a + anti-rabbit Cy^3 staining; **B**. NIH 3T3 cells stained with anti-_ -tubulin antibody + anti-mouse FITC staining; **C.** Nucleus of fibroblast cells was counterstained with DAPI; **D**. Merged picture.

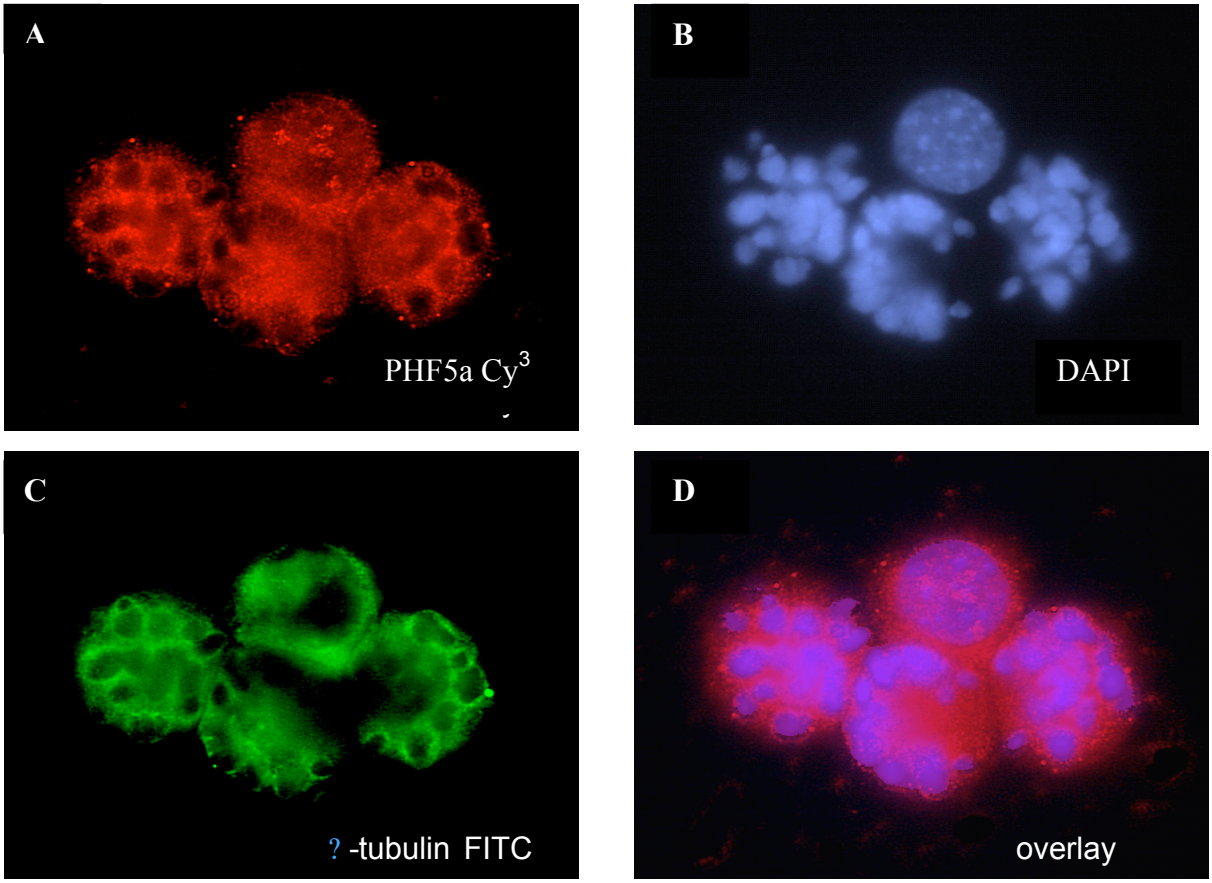


Figure 19. Subcellular localization of PHF5a protein in NIH 3T3 cell arrested in metaphase by using colcemid observed under fluorescent microscope. **A**. NIH3T3 cells stained with polyclonal anti-PHF5a antibody + anti-rabbit Cy^3; **B**. DAPI staining; **C**. anti-_-tubulin + anti mouse FITC staining; **D**. merged picture.

3.5 Colocalisation of PHF5a with splicing proteins SRp40 and $U2AF^{35}$ in NIH3T3 cells

The subcellular distribution of endogenous PHF5a and PHF5a-GFP fusion proteins and possible colocalization with splicing proteins SRp40 and $U2AF^{35}$ in mouse NIH 3T3 cells were analyzed by immunocytochemical studies. Appropriate PHF5a-GFP and SRp40-DsRed fusion protein constructs were prepared (Figure 20).

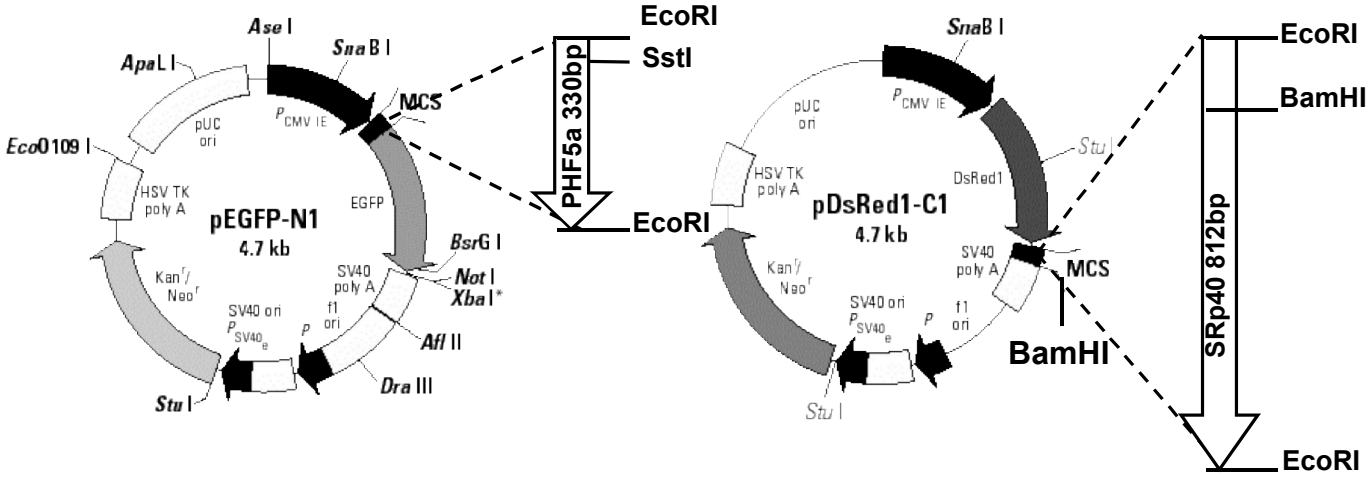


Figure 20. Schematic representation of the constructs used for colocalization studies. The cDNA sequence containing the complete ORF of PHF5a was cloned into the pEGFP-N1 vector (generated and described by Trappe et. al. , 2001). The fragment containing the complete ORF of SRp40 was generated by RT-PCR on RNA from 11.5-days murine embryos using primers SRp40 FP, SRp40 RP. The RT-PCR product of SRp40 was digested by *EcoR*I and cloned into the pDsRed1-C1 vector. The correct orientation of the SRp40 sequence in the pDsRed1-C1vector was determined by digestion with the enzyme *BamH*I.

NIH3T3 cells were transiently transfected with DNA of the vector encoding the PHF5a-GFP fusion protein. Subsequently, transfected cells were stained with anti-U2AF^{35} antibody. Signals from both the green PHF5a fusion protein and for U2AF^{35} protein (red staining) partially colocalize and are distributed throughout the nucleus, excluding the nucleolus, but are enriched in multiple speckles (Figure 21 A-F). The staining pattern for PHF5a and U2AF^{35} in the nucleus is similar to that known for other splicing proteins present in a granular/punctuate pattern probably representing perichromatin fibrilles (PFs) (Monneron and Bernhard, 1969) occurring mainly on the periphery of condensed chromatin regions and interchromatin granule clusters (IGCs) (Thiry, 1995). These speckled regions represent the major nucleoplasmic ribonucleoprotein constituents rich in different splicing components.

In addition, double transfection of NIH3T3 cells with DNA of vectors expressing recombinant GFP/PHF5a and DsRed/Srp40 fusion proteins revealed that PHF5a protein could also be colocalized with the second interaction partner- SRp40 in the nucleus of NIH3T3 cells. The staining pattern is restricted to the speckled regions (Figure 21 G-L).

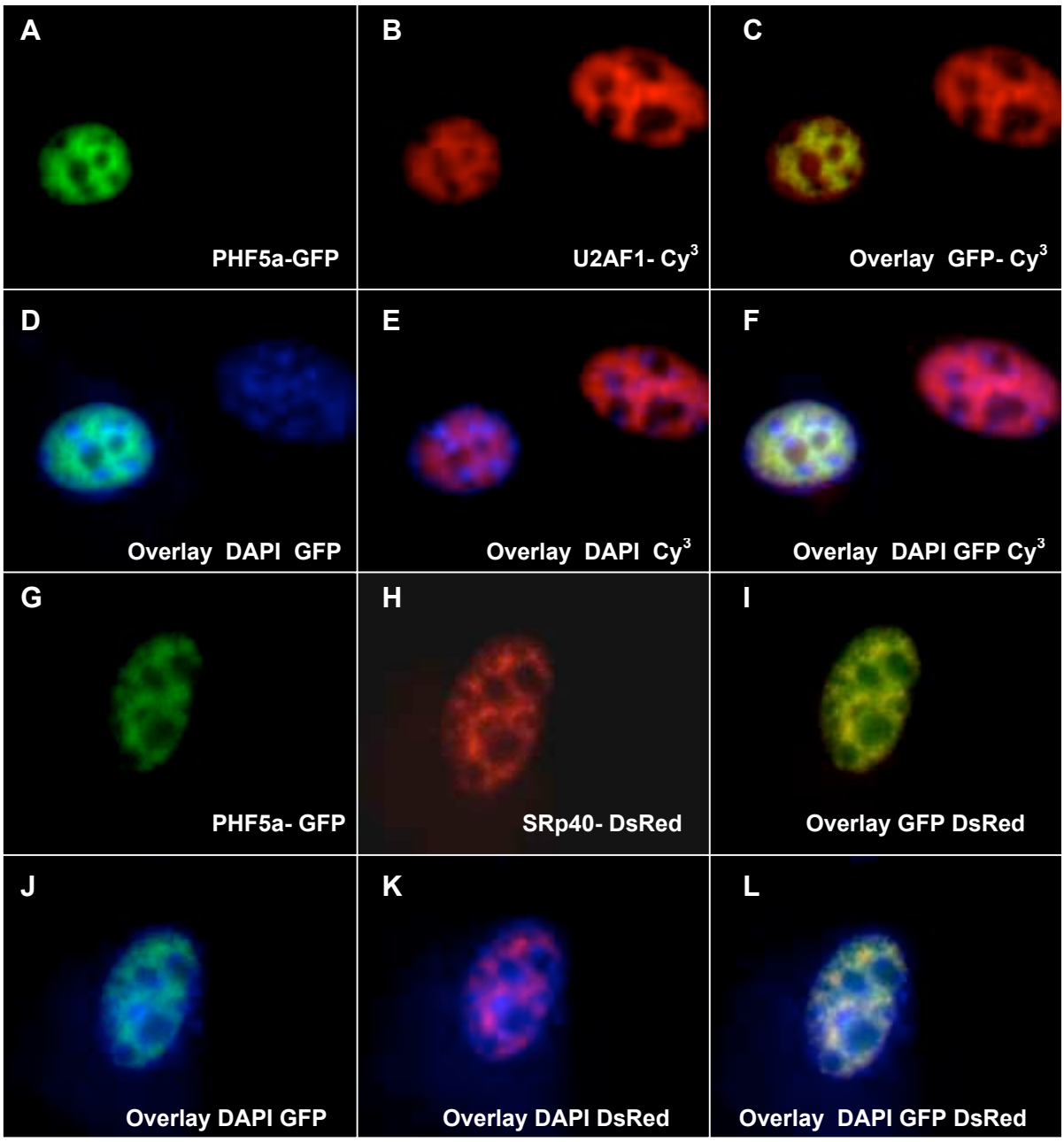


Figure 21. Sub-cellular colocalization of PHF5a with splicing proteins U2AF^{35} and SRp40 in nuclear speckles of NIH3T3 cells observed under a fluorescent microscope: **A-F.** The PHF5a-GFP fusion protein colocalizes with

the signals for the U2AF^{35} protein detected by staining with anti-U2AF^{35} and anti-rabit Cy3 antibodies. **G-L.** PHF5a-GFP fusion protein colocalizes with SRp40-DsRed fusion.

3.6 Characterization of U2AF1 splicing forms

It was previously described that U2AF is a heterodimer consisting of a 35 kDa and a 65 kDa subunit and the U2AF^{65}-interacting domain of the U2AF^{35} protein is localized in the N-terminal part (Wu et al., 1994).

Shaoping et al. (1999) have demonstrated that two polypeptides: 35kDa U2AF and an unknown protein at 33kDa could be effectively crosslinked to the conserved AG dinucleotide at 3' splice site. Using an anti-U2AF^{35} antibody both polypeptides could be detected by Western blot analysis and coimmunoprecipitation on NIH3T3 cell extracts (Figure 22). This result prompted us to focus on possible alternative transcripts of U2AF^{35}. Interestingly only one form of U2AF^{35} could be coimmunoprecipitated with the U2AF^{65} protein (Figure 22, lane 1).

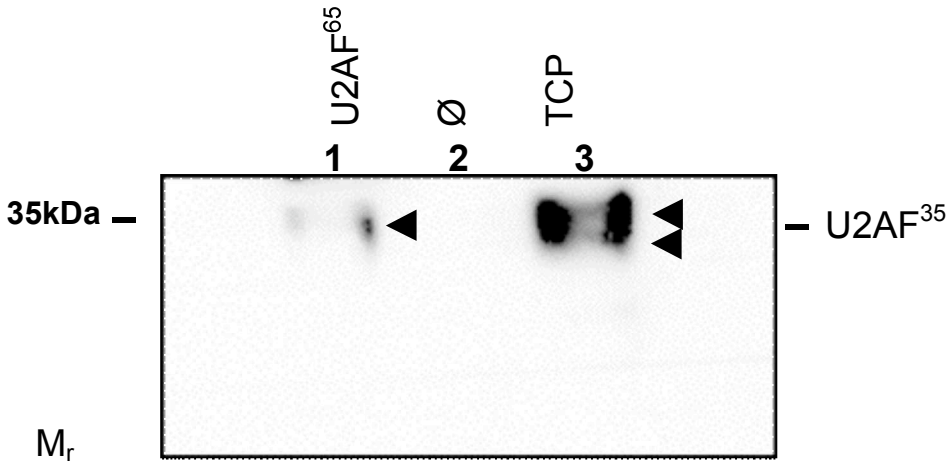


Figure 22. Western blot analysis demonstrating that only one alternative form of U2AF^{35} could be precipitated using an anti-U2AF^{65} antibody and that the anti-U2AF^{35} antibody recognizes two polypeptides isolated from the NIH3T3 cell extract. **Lane 1.** Proteins immunoprecipitated with an anti-U2AF^{65} antibody were resolved on the SDS-PAGE gel, transferred onto a nylon membrane and probed with an anti-U2AF^{35} antibody. **Lane 2**. As a

negative control protein A beads precipitate was probed with an anti-U2AF^{35} antibody. **Lane 3**. Total cell proteins (TCP) without immunoprecipitation were probed with an anti-U2AF^{35} antibody.

The genomic organization of the U2AF1 gene was determined by comparison of the cDNA sequence (Accession No.:NM_024187) with the mouse genomic DNA database, using the Mouse Genome Blast tool (NCBI) www.ncbi.nlm.nih.gov/genome/seq/MmBlast.html (Figure 23). One clone corresponding to a splicing form of U2AF^{35} (U2AF^{35}d) was isolated by the yeast two-hybrid screening. It revealed that the isolated clone contains a cDNA missing fragment corresponding to sequences of exons 2 and 3. In addition, this U2AF^{35} splicing form contains a premature internal stop codon as a result of a frame shift.

To examine the presence of other splicing forms of the U2AF1 gene RT-PCR experiments on RNA extracted from 11.5-days old mouse embryos were performed using U2AF1 specific primers U2AF1FP and U2AF1RP, which bind to both ends of the complete ORF of the U2AF^{35}. cDNA sequence analyses revealed the presence of a third splicing form (U2AF^{35}c) with a missing exon 3 characterized by shortened open reading frame leading to a truncated U2AF^{35} (Figure 22). Furthermore, a fourth splicing form of U2AF^{35} (U2AF^{35}b) was isolated corresponding to the recently described human U2AF^{35} variant (Pacheco et al., 2004) where exon 3 is substituted by an alternative exon Ab (Figure 23). Similarly, as described for the human U2AF^{35} variant this alternative exon Ab is localized in intron 2 of the murine U2AF1 gene (Figure 23). This U2AF splicing form contains a full open reading frame and in the deduced amino acid sequence of this U2AF^{35} variant only 7 amino acids are changed.

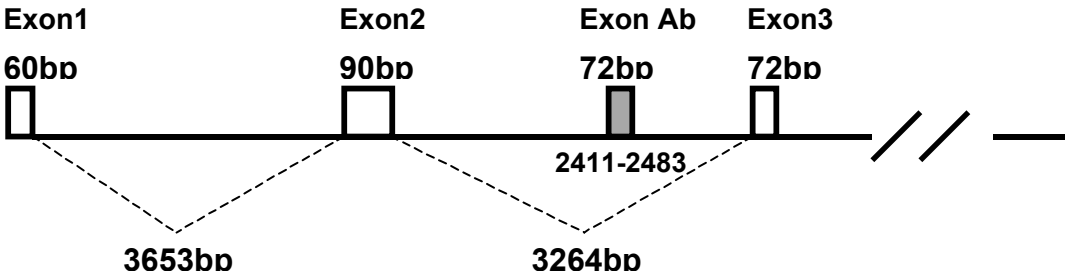


Figure 23. Schematic view of the genomic fragment containing the first 3 exons of the U2AF1 gene. The alternative exon Ab is presented as a grey block. Exon Ab is located in intron 2 of the U2AF1 gene (nucleotide position 2411-2483) and substitutes exon 3 in the alternative splicing form U2AF^{35}b.

Splicing variants were grouped (Figure 24) and assigned as: U2AF^{35}a corresponding to the previously described form (Zhang et al., 1992); U2AF^{35}b identical to human splicing pattern b, where exon Ab substitutes exon 3 (Pacheco et al., 2004); U2AF^{35}c with missing exon 2 and U2AF^{35}d with missing exons 2 and 3.

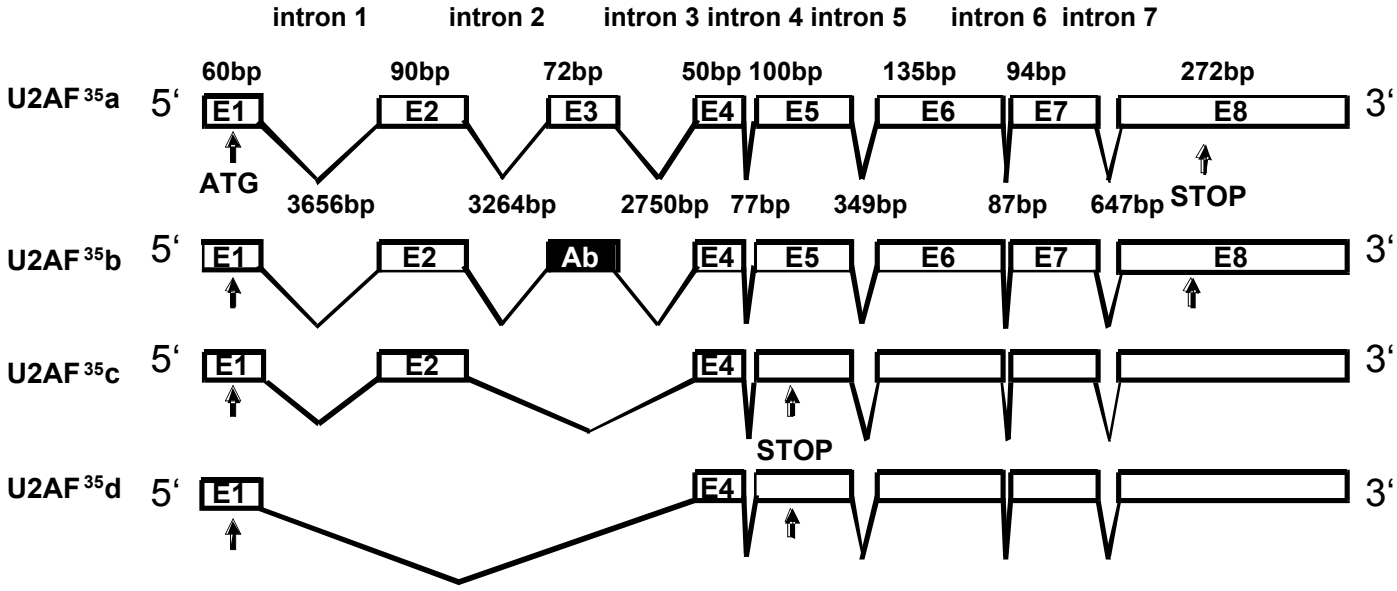


Figure 24. Schematic representation of the genomic organization and splicing forms of the murine U2AF1 gene. Alternative exon Ab in splicing variant U2AF^{35}b is marked as a black segment. Alternative variants U2AF^{35}c and U2AF^{35}d have premature stop codons in exon 5 as a result of a frame shift.

To determine if the alternative variant U2AF^{35}a has the same bindings affinity to the PHF5a protein as compared with variants U2AF^{35}b and the truncated polypeptide U2AF$^{35}\Delta$1-67 immunoprecipitation assays were used. The construct U2AF$^{35}\Delta$1-67 has a deletion in the N-terminal part and contains the sequence of U2AF^{35} downstream of exon 3. The N-terminal region of the U2AF^{35} protein is known to be responsible for the association with U2AF^{65}.

These proteins were expressed as tagged HA and c-Myc fusion proteins and used in the immunoprecipitation assays (Figure 25). All the variants of U2AF^{35} could be efficiently coimmunoprecipitated with PHF5a, indicating, that the N-terminal part of U2AF^{35} with exon 3 or with exon Ab does not modify the binding properties of the U2AF^{35} protein to PHF5a.

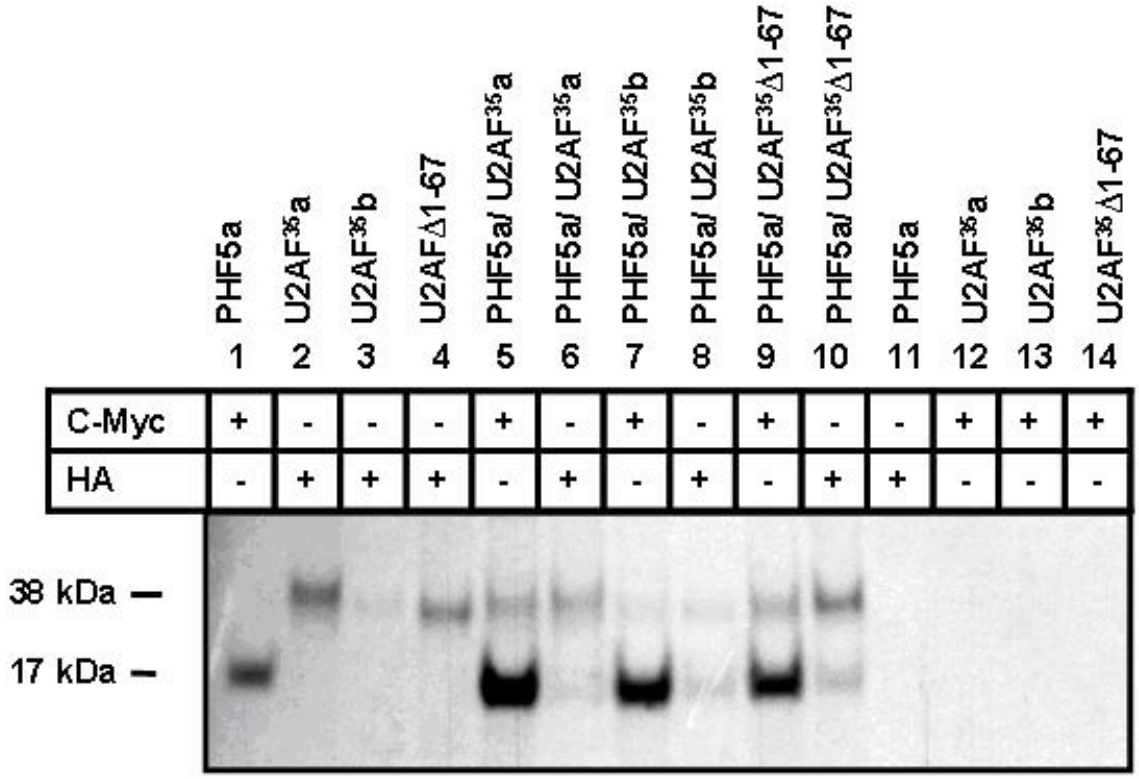


Figure 25. SDS-PAGE analysis showing that murine PHF5a protein coimmunoprecipitates with splicing forms U2AF^{35}a, U2AF^{35}b and the U2AF^{35}Δ1-67 fragment, spanning a sequence from exon 4 to the stop codon. Radioactively labeled ^{35}S proteins were *in vitro* transcribed/ translated from the construct BK/PHF5a (c-Myc epitope), AD/U2AF^{35}a (HA epitope), AD/U2AF^{35}b (HA epitope) and AD/U2AF^{35}Δ1-67 (HA epitope). Generated proteins: PHF5a-c-Myc, U2AF^{35}a-HA, U2AF^{35}b-HA, U2AF^{35}Δ1-67-HA, were used in the immunoprecipitation experiment. 10µl of the immunoprecipitate was loaded onto a 4-12% SDS polyacrylamide gel: **Lane 1**: PHF5a + c-Myc antibody; **Lane 2**: U2AF^{35}a + HA antibody; **Lane 3**: U2AF^{35}b + HA antibody; **Lane 4**: U2AF^{35}Δ1-67 + HA antibody; **Lane 5**: PHF5a + U2AF^{35}a + c-Myc antibody; **Lane 6**: : PHF5a + U2AF^{35}a + HA antibody; **Lane 7**: PHF5a + U2AF^{35}b + c-Myc antibody; **Lane 8**: PHF5a + U2AF^{35}b + HA antibody; **Lane 9**: PHF5a + U2AF^{35}Δ1-67 + c-Myc antibody; **Lane 10**: PHF5a + U2AF^{35}Δ1-67 + HA antibody;

Lane 11: Negative control PHF5a + HA antibody; **Lane 12**: Negative control U2AF^{35}a + HA antibody: **Lane 13**: Negative control U2AF^{35}b + HA antibody; **Lane 14**: U2AF^{35}Δ1-67 + HA antibody.

3.7 Expression of PHF5a and U2AF^{35} in the testis

Previous studies (Trappe et al., 2001) have indicated that the PHF5a gene might be ubiquitously expressed in all tissues and organs. However it was also shown that PHF5a is differentially expressed between the spermatogonia-specific GC-1spg cells and spermatocyte-specific GC-4spc cells (Figure 26).

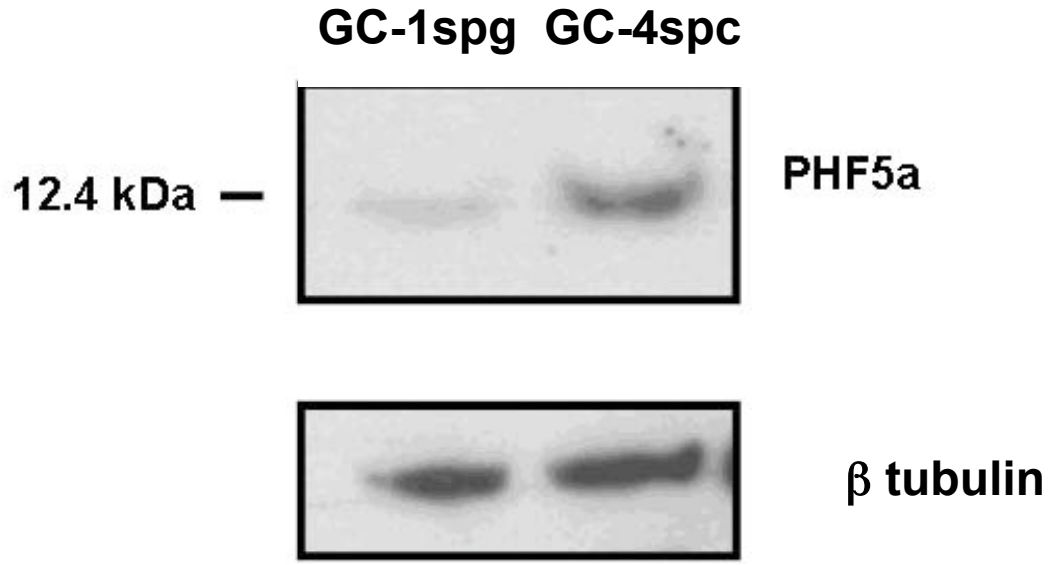


Figure 26. Western blot analysis demonstrating that the PHF5a gene expression is up-regulated in GC-4spc cells as compared to GC-1spg cells. Proteins were separated on a 4-12% SDS polyacrylamide gel, transferred onto nylon membrane and hybridized with anti-PHF5a and tubulin antibodies, respectively.

To obtain a more detailed picture of PHF5a expression within the murine testis immunostaining on testis sections was performed. (Figure 27 A-B, E-F). Immunofluorescence staining of testis sections reveals that PHF5a protein is mainly confined to the nuclei of primary spermatocytes. The strongest labelling was observed in pachytene spermatocytes, but also earlier meiotic cells such as leptoten and zygotene spermatocytes were also positive. Nuclear staining of PHF5a was in a speckled fashion. PHF5a immunoreactivity was almost

absent in interstitial cells, except for a very weak labelling in some scattered cells, which can not represent the entity of Leydig cells. No reaction was visible in spermatids and spermatozoa.

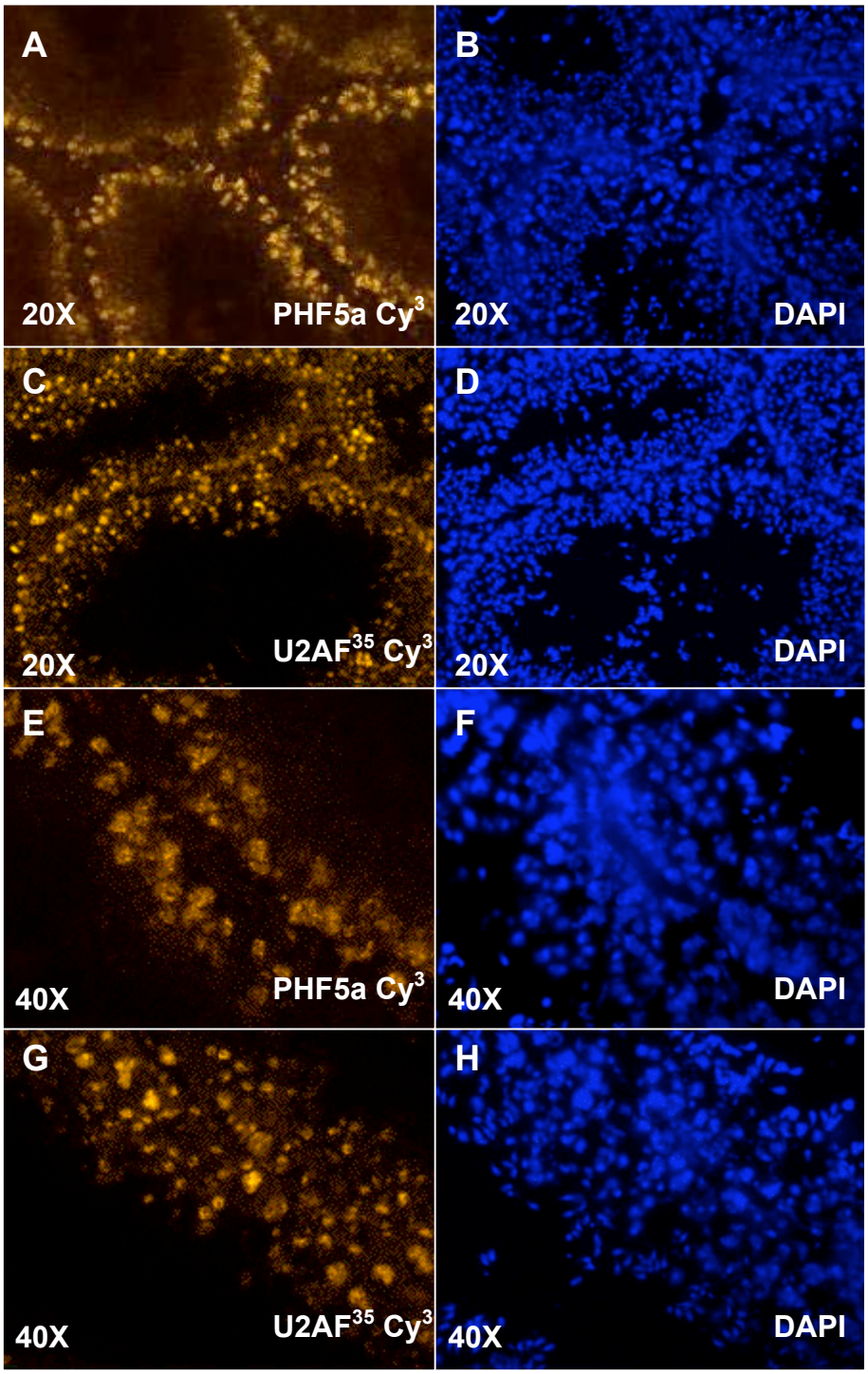


Figure 27. Immunofluorescence localization of PHF5a and U2AF^{35} in mouse testis. **A, E**: PHF5a immunoreactivity is visible in nuclei of primary spermatocytes and few scattered cells in the interstitium.

Spermatids and spermatozoa are negative for PHF5a. **C, G**: U2AF^{35}. Higher magnifications (E, G) show that the immunoreaction for PHF5a and U2AF^{35} appears in multiple punctuate sites in the nuclei of primary spermatocytes, whereas U2AF^{35} labeling in nuclei of spermatids concentrates into few, but larger foci (G). **B, D, F, H**: corresponding DAPI staining.

The expression of U2AF^{35} was also examined as the binding studies suggested at least a partial temporospatial overlap with PHF5a synthesis in testis (Figure 27 C-D, G-H). The U2AF^{35} protein was detectable in all types of spermatocytes as well as in round and some elongating spermatids, but not in spermatozoa. The strongest fluorescence signal for both proteins was clearly evident in pachytene spermatocytes.

Western blotting revealed that the expression of PHF5a increases during postnatal testis development (Figure 28). There is only a weak signal of PHF5a in the protein extract isolated from testis of 5 day-old mice, where spermatocytes are still absent and from testis of W/W^{v} mutant mouse lacking cells from germ line (Figure 29). This would support the observation that PHF5a is apparently also weakly expressed in somatic (interstitial) cells. Expression of PHF5a is noticeably higher in the Tfm mutant mouse, which is characterized by an arrest of spermatogenesis at the early spermatocyte level and massively increases in the Insl3 mutant, which produces no postmeiotic cells such as spermatids and spermatozoa.

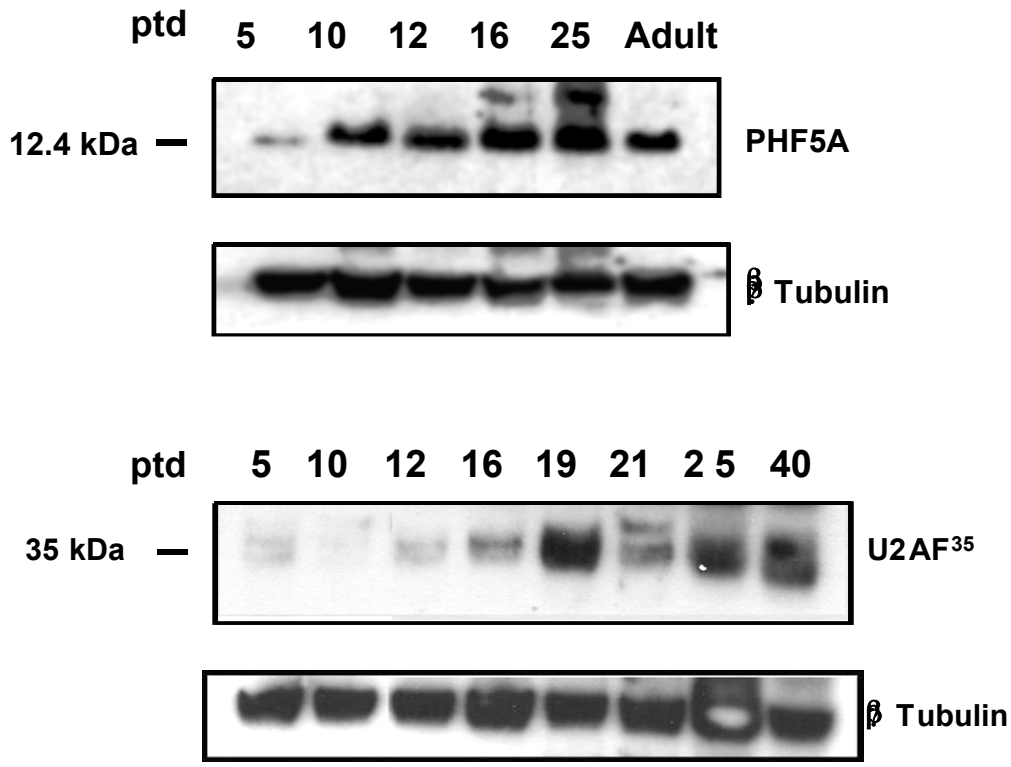


Figure 28. Western blot analysis representing that both PHF5a and U2AF^{35} expression increase during postnatal mouse testis development. Proteins extracted from testis at different days in postnatal development: 5, 10, 12, 16, 25 (ptd), and from adult animals were separated and probed with an anti-PHF5a and anti-_-tubulin antibody, respectively; Proteins from testis of 5, 10, 12, 16, 19, 21, 25, 40 day-old mice in postnatal development, and an adult animal, were separated and probed with an anti-U2AF^{35} and an anti-_-tubulin antibody as described in materials and methods section.

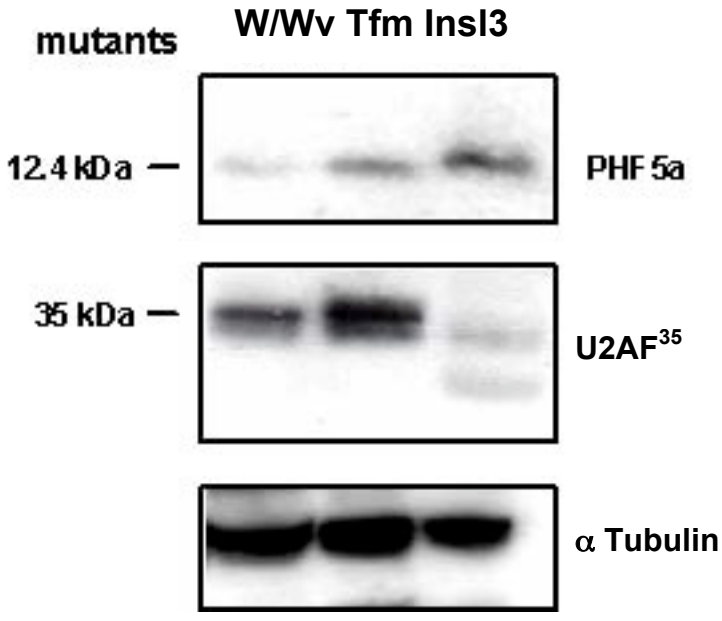


Figure 29. Western blots analysis showing expression of U2AF^{35} and PHF5a in testes of mouse mutants W/Wv, Tfm and Insl3. Proteins were separated on SDS/4-12% polyacrylamide gels, transferred onto nylon membrane and hybridized with an anti-PHF5a, anti-U2AF^{35} and an anti-_-tubulin antibody, respectively.

Expression of U2AF^{35} in the postnatal testis reminds to that previously found for PHF5a. (Figure 28) The highest expression starts between day 16 and 19 in postnatal testis development. Surprisingly, expression of U2AF^{35} in the testis of the Insl3 mutant with targeted disruption of the Insl3 gene is highly decreased (Figure 29).

Interestingly the PHF5a nuclear immunoreactivity pattern is dynamic during progression of mouse spermatogenesis in PHF5a. Multiple punctuate sites in primary spermatocytes concentrate into few, but extended positive foci within the nucleus of elongating spermatids (Figure 30). Spermatogonia and mature spermatocytes were negative for PHF5a protein.

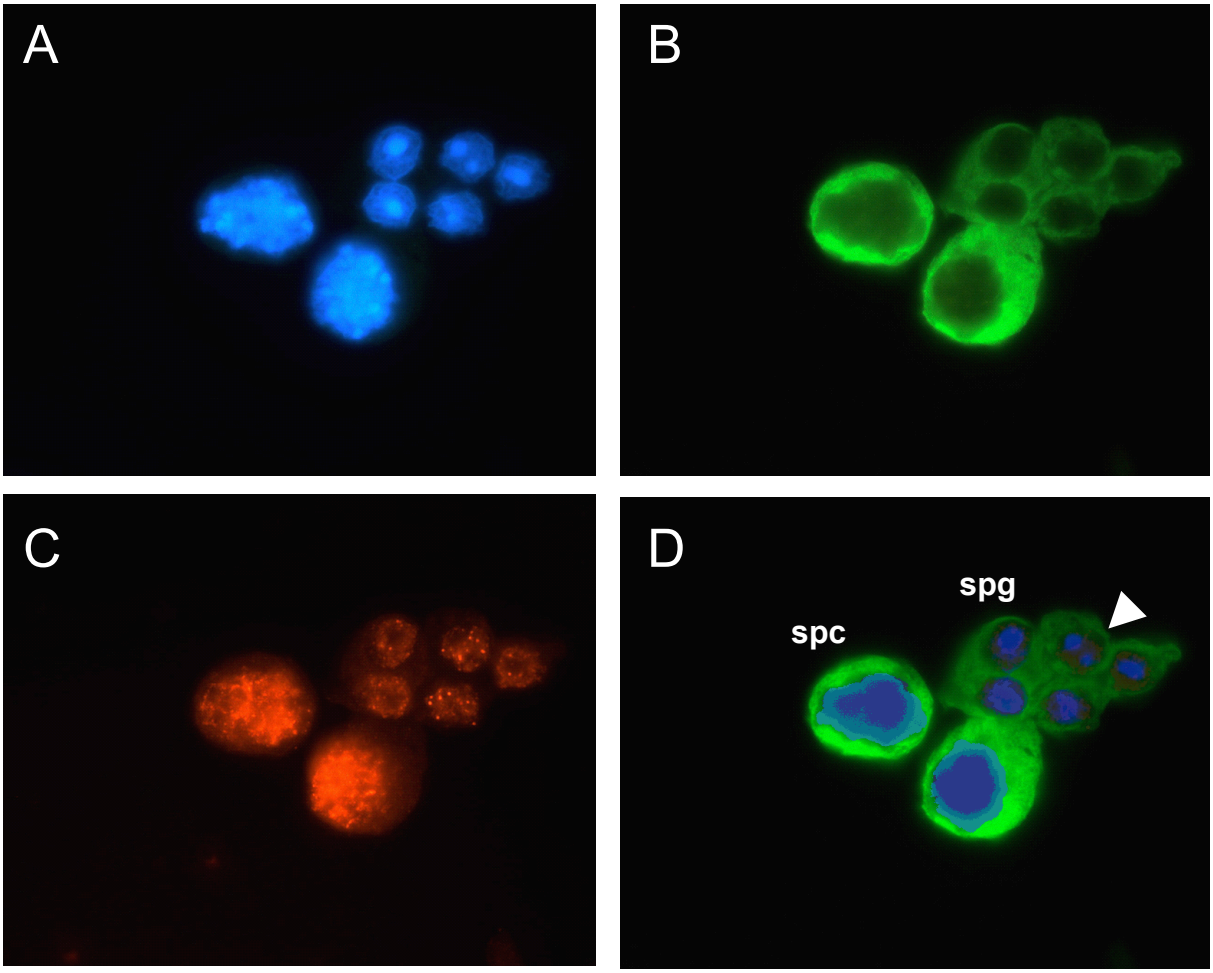


Figure 30. Immunostaining of testicular cell extracts with anti-_-tubulin and anti-PHF5a antibodies observed under a fluorescent microscope. **A.** DAPI staining; **B**. staining with anti-tubuline and anti-mouse FITC conjugated antibody; **C**. staining with anti-PHF5a and anti-rabbit Cy^3 antibody; **D**. overlay DAPI, FITC and Cy^3. Sites of PHF5a accumulation in spermatids are marked with an arrow. (Spc) spermatocytes, (Spg) spermatogonia.

3.8 *In vivo* coimmunoprecipitation of PHF5a and $U2AF^{35}$ in GC- 4spc cells

The results of immunostaining on testis sections revealed that the highest expression of both PHF5a and $U2AF^{35}$ proteins was detectable in pachytene spermatocytes. Based on this result, murine GC-4spc cells derived from murine spermatocytes were selected to verify that the interaction between PHF5a and $U2AF^{35}$ could exist *in vivo*. Using an anti-PHF5a antibody, the $U2AF^{35}$ protein could be efficiently coimmunoprecipitated from GC-4spc cell extracts (Figure 31). Subsequently, both interaction partners were visualized by Western blot analysis using anti-$U2AF^{35}$ and anti-PHF5a antibodies, respectively. The data presented in figure 31 indicate that an interaction between PHF5a and $U2AF^{35}$ exists *in vivo*. As a positive control coimmunoprecipitation of $U2AF^{35}$ was performed with an anti-$U2AF^{65}$ antibody.

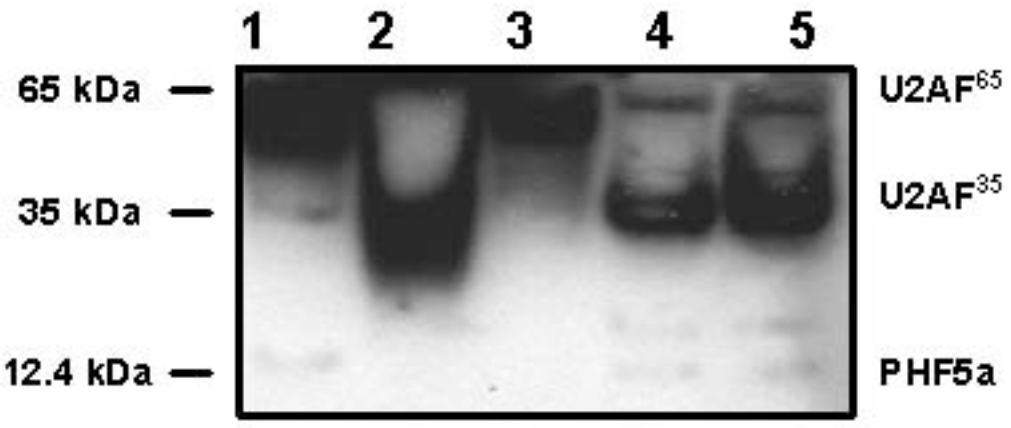


Figure 31. Western blot analysis showing that PHF5a proteins coimmunoprecipitate with U2AF^{35} in the GC-4spc protein cell extract. Cell lysate from GC-4spc cells was incubated with following antibodies: **Lane 1.** Anti-PHF5a; **Lane 2.** Anti-U2AF^{65} (positive control); **Lane 3.** Negative control incubated with protein A beads only. After elution from protein A beads 10µl of the immunoprecipitate was separated on a SDS/4-12% polyacrylamide gel transferred onto a nylon membrane and probed with an anti-PHF5a, U2AF^{35} and U2AF^{65} antibodies, respectively. **Lane 4-5**: Protein extract from GC-4spc cells probed with the same antibody. Note that the unspecific precipitate is also giving a strong background in the negative control- (lane 3).

3.9 *In vivo* splicing assay

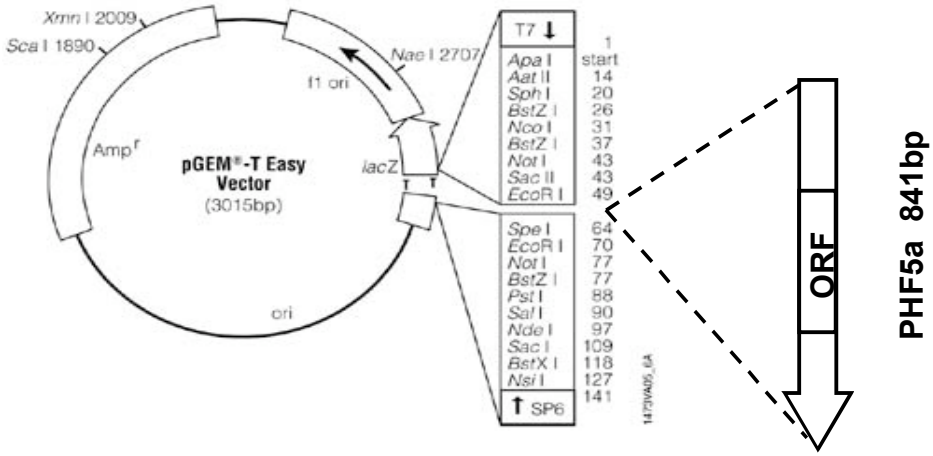
The coimmunoprecipitation experiments indicated that the PHF5a protein is interacting with components of the splicing machinery. Therefore i*n vivo* splicing assay was designed to analyze if RNAi downregulation of PHF5a can affect splicing of the ß-globin mini-gene. ß-globin mini-gene consisted of first two exons separated with the intron and was expressed under control of CMV promoter in NIH3T3 cells. Expression and splicing of the mini-gene was monitored by RT-PCR in cells cotransfected with PHF5a RNAi. 48 hours after transfection when PHF5a expression reached lowest level, splicing of _-globin was assessed. There were no detectable changes in splicing of PHF5a RNAi treated cells as compared to parental and RNAi luciferase control.

TCACATTTGCTTCTGACATAGTTGTGTTGACTCACAACCCCAGAAACAGACATCATGGTGCACCTGACTGATGCT
GAGAAGGCTGCTGTCTCTGGCCTGTGGGGAAAGGTGAACGCCGATGAAGTTGGTGGTGAGGCCCTGGGCAGgttg
gtatccaggttacaaggcagctcacaagaagttgggtgcttggagacagaggtctgcttaccagcaggcactaac
tttgagtgtcccctgtctatgtttccctttttagGCTGCTGGTTGTCTACCCTTGGACCCAGCGGTACTTTGATA
GCTTTGGAGACCTATCCTCTGCCTCTGCTATCATGGGTAATGCCAAAGTGAAGGCCCATGGCAAGAAAGTGATAA
CTGCCTTTAACGATGGCCTGAATCACTTGGACAGCCTCAAGGGCACCTTTGCCAGCCTCAGTGAGCTCCACTGTG
ACAAGCTGCATGTGGATCCTGAGAACTTCAGG

Figure 32. Sequence of the ß-globin mini gene. 5'UTR is marked in red colour. Exons are presented in capital letters and the intron in small blue letters. Primer binding sites used for generation and detection of ß-globin transcription were underlined.

3.10 PHF5a is ubiquitously expressed during embryonic development

C. elegans ortholog of human PHF5a shows a muscle-specific expression domain and is essential for *C. elegans* morphogenetic development (Trappe et al., 2001). To determine if PHF5a is differentially expressed during embryonic mouse development whole mount *in situ* hybridization on 11.5-days old mouse embryos was performed. Antisense RNA probe containing complete cDNA of PHF5a was used according to the protocol described in the Materials and Methods section (Figure 33). Specific signals were detectable in almost every part of the embryo suggesting that PHF5a is ubiquitously expressed during embryonic development. The most intensive signal was observed in the region of the yolk sack, the limb buds and the spinal cord (Figure 34).



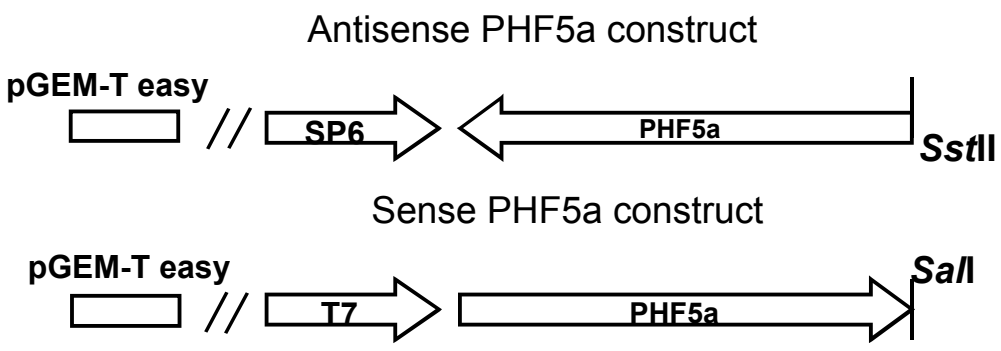


Figure 33. Schematic representation of the constructs used for the generation of sense and antisense whole mount in situ hybridization probes. The pGEM-T easy construct containing the full cDNA sequence of PHF5a was linearized with the enzyme *Sst*II (antisense probe) and with *Sal*I (sense probe)

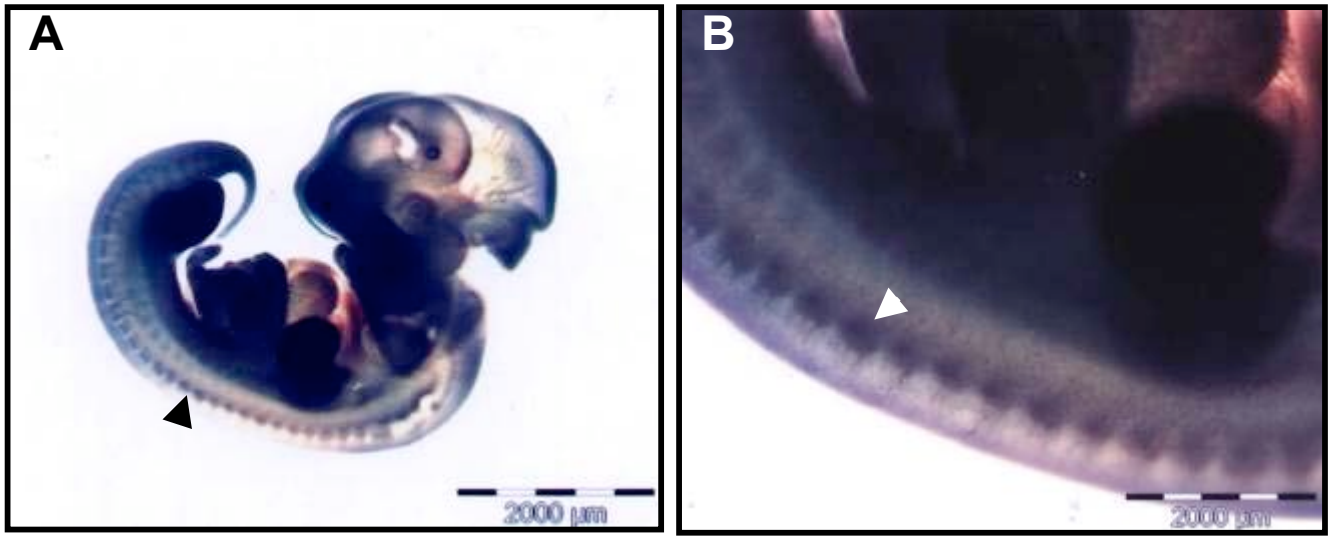


Figure 34. The picture of PHF5a expression in 11.5 day mouse embryo found by the whole mount *in situ* hybridization with RNA antisense probe. **A**. PHF5a is highly expressed in all parts of embryo with particularly strong expression in the yolk sack, spinal cord and limbs regions. **B.** Magnification of spinal cord region

3.11 PHF5a DNA-binding assays

PHF5a shows a weak homology to the homeo-domain (PHD) zinc finger proteins. PHD fingers are described to be engaged in DNA binding and at least a few PHD fingers are present in known transcription factors (Aasland et al., 1995). Therefore, a recombinant PHF5a-GST fusion protein was generated, purified and used in the DNA-binding assay. (Figure 35 and Figure 36).

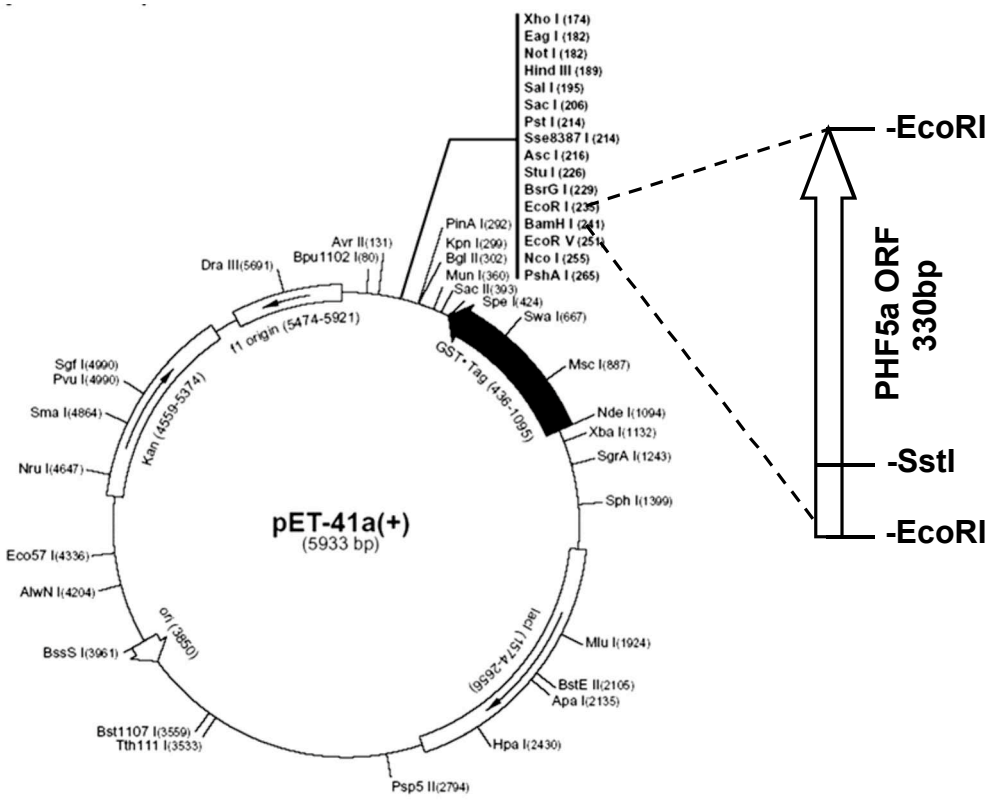


Figure 35. Schematic representation of the construct used for the expression of the recombinant PHF5a-GST fusion protein. The complete ORF of PHF5a was cloned into the *Eco*RI site in the MCS of the pET41c vector. The correct orientation of the PHF5a fragment in the vector was determined by a test digestion with the enzyme *Sst*I.

First, genomic DNA fragments with an average size of 200 bp were generated by sonification and subsequently labelled with ^{32}P cCTP. The recombinant PHF5a-GST fusion protein was incubated with ^{32}P labelled genomic DNA fragments with an average size of 200bp in the buffer containing Zn^{2+} ions (Figure 37 A). DNA fragments that bind to the PHF5a protein were purified and the activity was quantified by using the scintillation counter (Figure 37 B). In addition, radioactive DNA fragments were also visualized by PAGE and subsequently exposed to X-Ray films as it can be seen in figure 37 C. In all experiments a GST protein was used as a negative control. Genomic DNA fragments could be pulled down only in the presence of PHF5a protein.

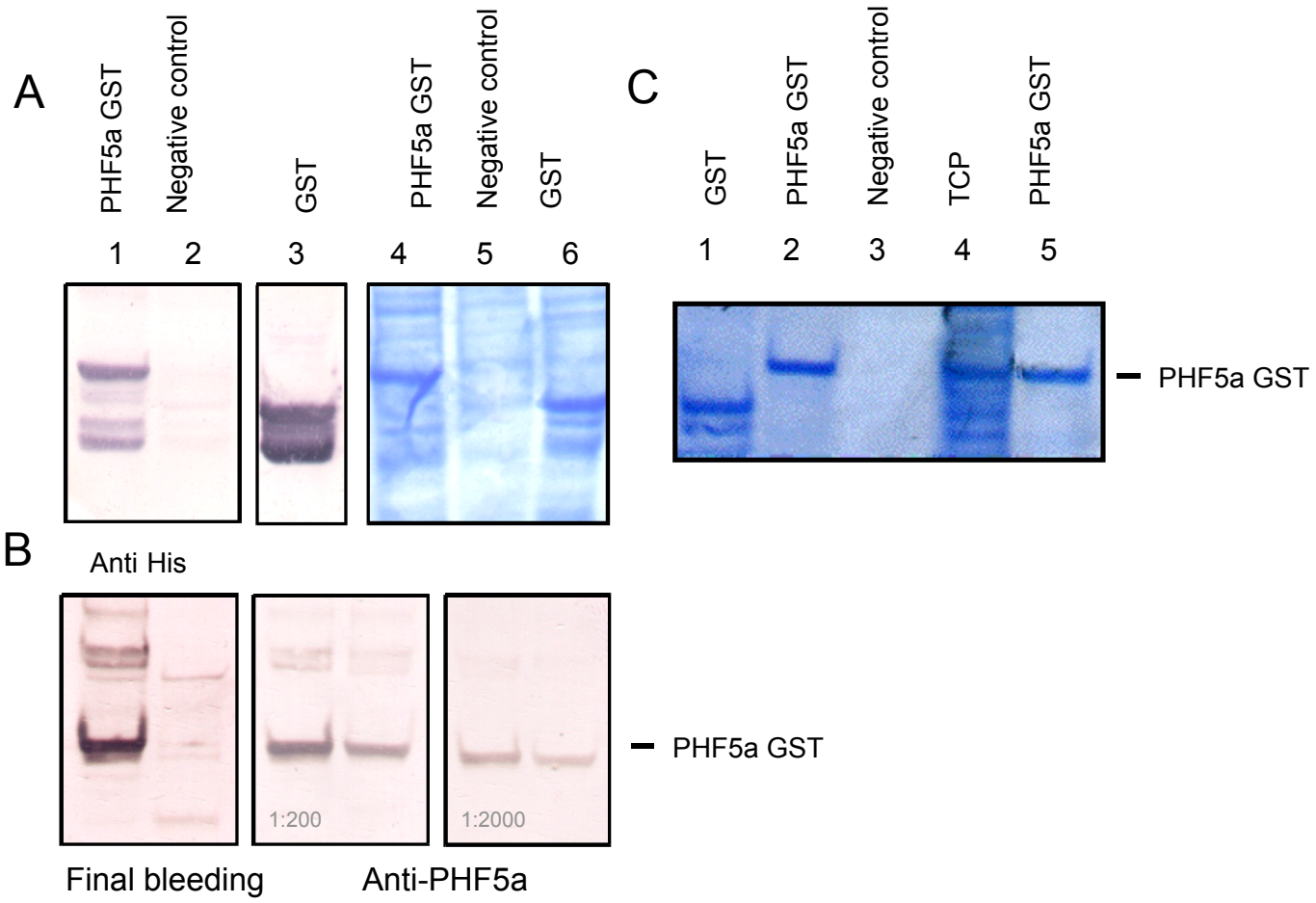


Figure 36. A. SDS- PAGE analysis of the bacterial protein cell extract after induction of PHF5a expression by IPTG. Proteins were transferred onto a nylon membrane and stained with anti-His antibody: **Lane 1**. Bacteria transformed with pET-41c PHF5a construct; **Lane 2**. Noninduced control; **Lane 3.** Bacteria transformed with empty pET-41c vector; **Lane 4-6**. SDS-PAGE gel stained with comassie blue; **B**. SDS- PAGE analysis of bacterial protein cell extract after induction of expression of recombinant proteins with IPTG. Again, proteins were transferred onto a nylon membrane and stained with an anti-PHF5a final bleeding and an anti- PHF5a antibody. **C.** SDS- PAGE gel analysis of proteins purified by the GST resin. **Lane 1**. Purified GST protein; **Lane 2**. Purified PHF5a–GST protein; **Lane 3.** Negative control (non induced sample); **Lane 4**. Total cell proteins from an induced PHF5a-GST sample; **Lane 5.** Purified PHF5a- GST protein.

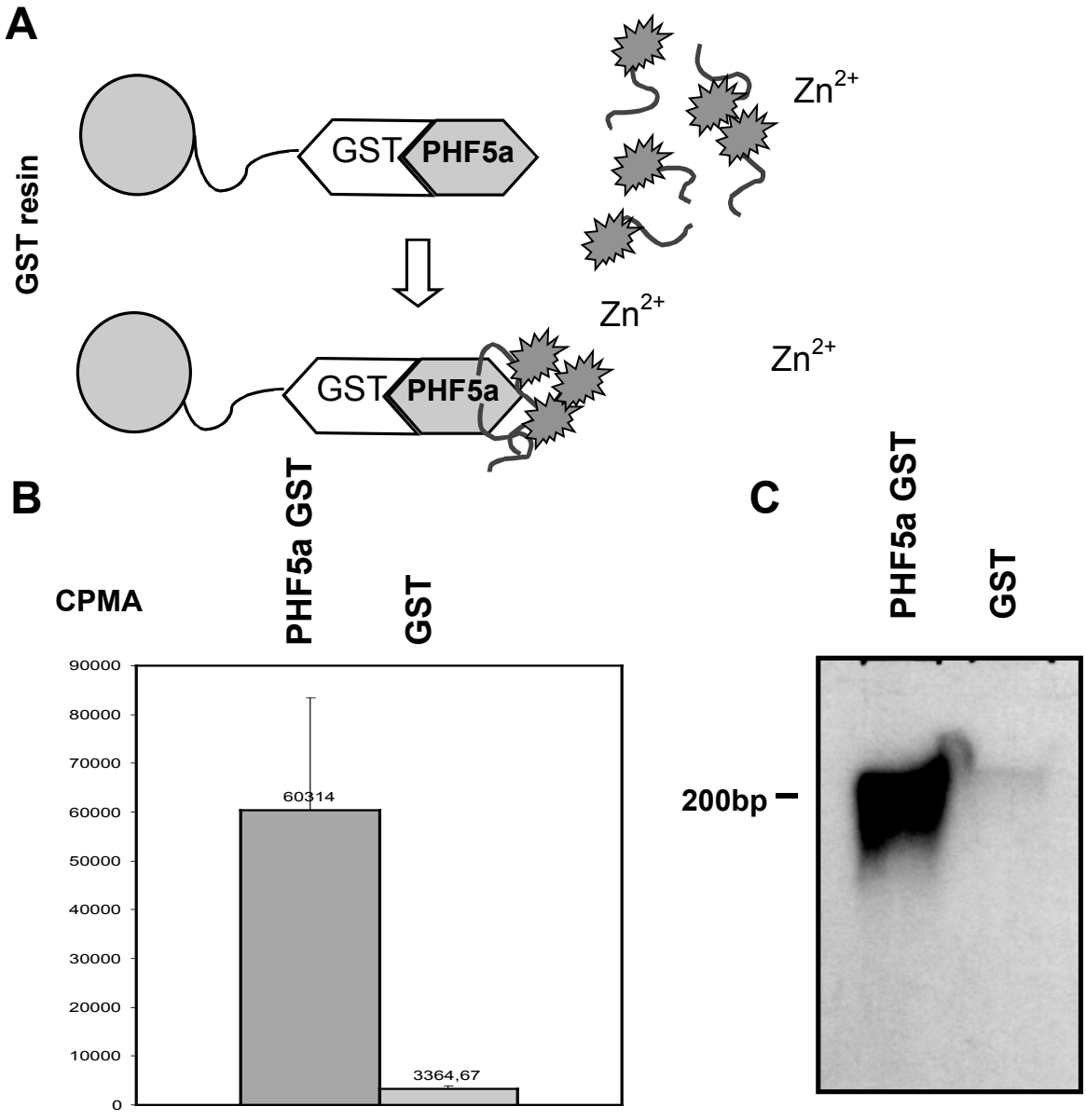


Figure 37. DNA-binding assay. **A.** Schematic representation of the strategy used for the investigation of DNA-binding properties of the PHF5a protein. PHF5a-GST fusion protein was coupled with the GST binding resin and incubated with radioactively labelled genomic DNA fragments. **B.** DNA fragments were pulled down and the radioactivity was measured by using a scintillation counter. **C**. DNA fragments pulled down by recombinant PHF5a-GST and GST proteins were resolved on a PAGE gel and visualized on an X-Ray film.

3.12 PHF5a Whole Genome PCR assay

In order to identify DNA target sequences which are recognized by the PHF5a protein the whole genome PCR assay (WG-PCR) was employed. Human genomic DNA was digested with the enzyme *Mbo*I, thus generating DNA fragments with an average size of 400bp in length. Subsequently, DNA fragments were ligated to the *Mbo*I linker. Target sequences of PHF5a were selected by binding to the PHF5a–GST fusion protein in the presence of GST resin followed by a DNA pull down. Then, DNA was subsequently amplified with the primer WG-PCR I. After 6 cycles of selection and amplification DNA fragments containing PHF5a target sequence were amplified with nested primer WG-PCR II and cloned into the pGEM-T easy vector (Figure 38). Sequence analyses of the cloned inserts revealed that selected DNA fragments share homologous DNA sequences (Figure 37). However, using the Human Genomic Blast tool (NCBI) and Ensembl Genome Browser, www.ensembl.org failed to identify promoter regions and other 5' regulatory elements matching the DNA sequences identified by the WG-PCR assay.

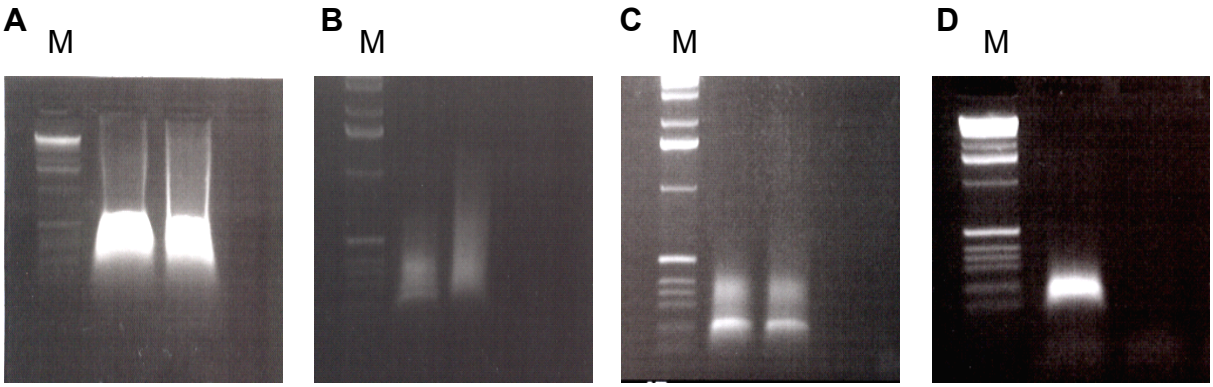


Figure 38. Agarose gel electrophoresis of PCR products obtained after different steps of the WG-PCR assay. **A**. Genomic DNA digested with the enzyme *Mbo*I; **B**. Genomic DNA ligated with *MboI* linker; **C**. Product of PCR amplification with WG-PCR I primer after 6 cycles of selection; **D**. PCR product of the nested PCR with primer WG-PCR II. M- size marker.

Fragment	Sequence
WGP6	ATTCTTATTTTCATAATCTAACAGGTATGTATTTATTAGCCAAATTTTTT**AGAAA**CT GCATTAG**AATTCTGTAT**ATTC**AGAAAT**GGAGATGA**GAATGT**AAACAA ATTATTACATCATTCACCTAACAGCATTTCCCGGAACCGCATAGGCCA
WGP7	CGGTATATCAGCACACACACCCCAAAAAAAGGGAAAAATTATGTTTGCCTTTGAAGA GCTCCAAATATG**AGAAAT**CCAGGAAAGATGTT**AATTCCTGTAT**CACCCTA TTCTTTTATTG**GAATGT**ACAGAGAAAGACACAA A
WGP1	TGGCCTATGCGGTAAGCATCTGCTTTTCGGAGAGCCATCAGAGATTTTTATTGCCTTCC TTAAGCCTCCATGACATCACTAAATAAACCTGACACA**ATTCTTTT**CTGCACTAGTAGT TAGAGTTTTCCAAGCAGAA CGGTTTTTATG
WGP2	TGGCCTATGCGGGAGTGCAGTGGCGCGATCC**AAGAATGT**GCAGCCTAAACTTGG ATTTCAGGACATTACACCTAATGCATAGTTGGCCTTCGAAAAGTCAAGCTCCTTTATAA TCCACTCCCTCTCCTATACTAACTTTGATAT

Figure 39. DNA fragments identified by the WG-PCR assay. Regions of homology are assigned in bold letters. Spacer sequences between homologous sequences in fragment WGP6 and WGP7 are underlined as they contain the same nucleotide size of 18bp.

3.13 SELEX

Another strategy was employed to identify consensus DNA sequences that are recognized by PHF5a protein. Single stranded SELEX library of random oligonucleotides (14 nt oligomers with primer starter) was made double stranded in one cycle of PCR using reverse primer SEL R. The double stranded SELEX library was incubated with the recombinant PHF5a-GST fusion protein and GST-binding beads. After 6 rounds of selection and amplification with primers SEL F and SEL R, PCR products were cloned into the pGEM-T easy vector and were sequenced (Figure 40). Multi alignment revealed a putative consensus DNA binding sequence for PHF5a: $CGAA^{C}/_{G}A^{C}/_{T}CCGCC$

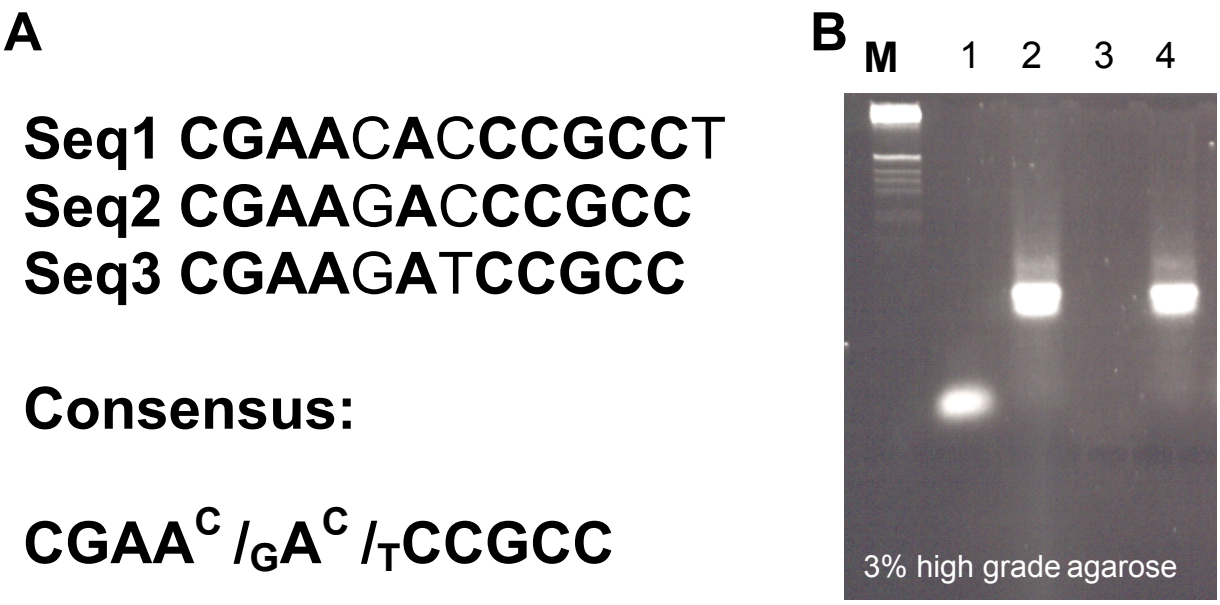


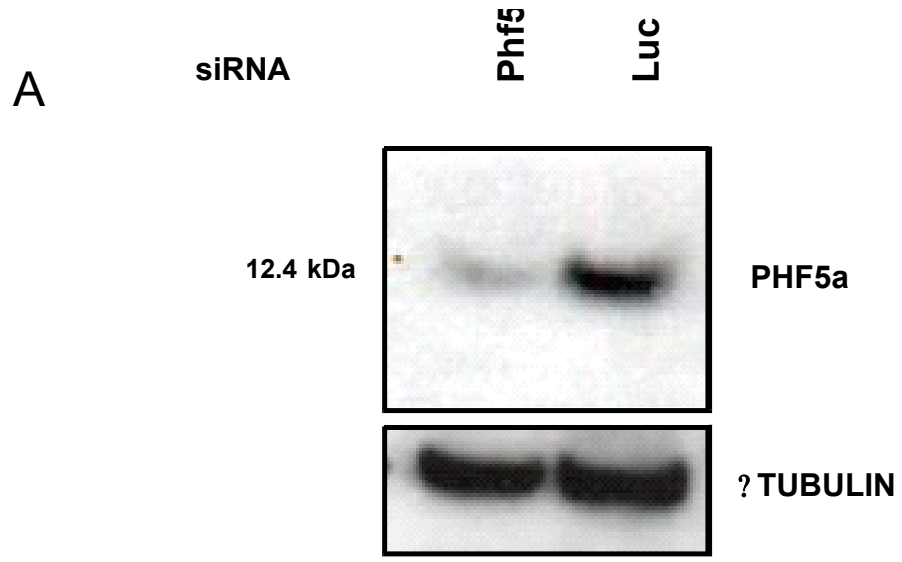
Figure 40. A. Putative consensus binding sequence of PHF5a protein selected by the SELEX assay. **B.** Agarose gel electrophoresis of PCR products after 6 rounds of selection and amplification. DNA fragments were purified, cloned into pGEM-T easy vector and sequenced. M-sizemarker, Lane 1- single stranded SELEX oligonucleotides, Lane 2- double stranded SELEX oligonucleotides after 6 rounds of selection and amplification, Lane 3- negative control.

3.14 Down-regulation of PHF5a expression in NIH3T3 cells by using the RNAi technique

In order to evaluate the function of PHF5a in mouse NIH3T3 cells, the approach of gene silencing through RNA interference (RNAi) was applied. Transfection of NIH3T3 cells was accomplished with either PHF5a sequence-specific small interfering RNA (siRNA) duplex oligonucleotides or with control duplex siRNA olignucleotides against the firefly luciferase gene (luc) as described in the Material and Methods section. Forty eight hours after transfection NIH3T3 cells were collected and used for the determination of down-regulation of PHF5a gene expression. To investigate the knockdown of PHF5A expression in duplex siRNA-transfected NIH3T3 cells at the protein level, Western blot analysis using a polyclonal

antibody against PHF5a was performed. As shown in figure 41 A, the expression of PHF5a was specifically reduced by the cognate duplex siRNA, but not when control oligonucleotides against the firefly luciferase gene (luc) were used. The expression of a non-targeted housekeeping gene, α-tubulin, was found to be unaffected and the reduction in PHF5a protein was more than 60**%** complete as quantified by Western blotting.

Furthermore, to assess the effect of down-regulated PHF5a expression on cell proliferation, we analyzed PHF5a duplex siRNA-transfected NIH3T3 cells as well as non-transfected parental cells and control-transfected cells with siRNA duplex against the luciferase gene. After the transfection step with siRNA duplex, the EZ4U-based proliferation assay was performed at different time points (0 hours, 24 hours, and 48 hours). The 0 hours incubation time was performed as a control point for equal amount of cell numbers where the optical density at 450 nm was at a similar level for both PHF5a duplex siRNA-transfected NIH3T3 cells and control-transfected NIH3T3 cells (Figure 41 B). As shown in figure 41 B after 24 hours and 48 hours we observed a significant inhibition of cell proliferation in PHF5A duplex siRNA-transfected NIH3T3 cells of 50% and 60%, respectively, as compared to control-transfected and parental NIH3T3 cells.



B

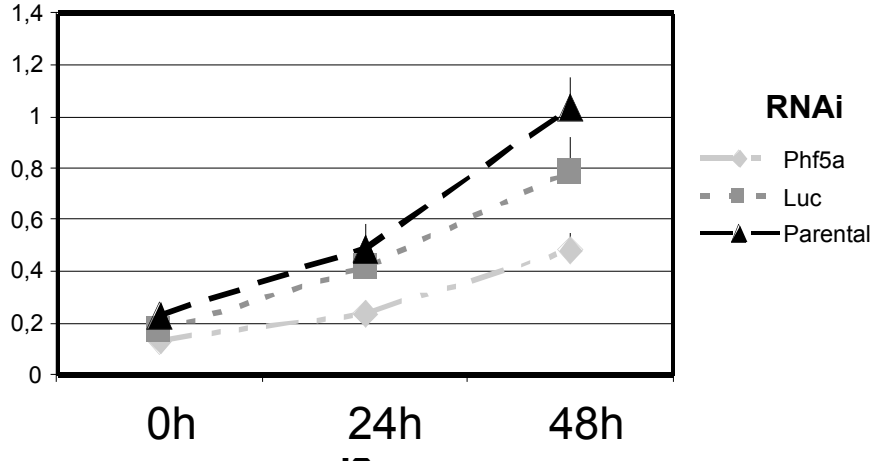


Figure 41. **A.** Down-regulation of PHF5a expression in NIH3T3 cells transfected with sequence-specific duplex siRNA oligonucleotides against PHF5. At twenty-eight hours after transfection NIH3T3 cells were collected and used for protein isolation. As negative control, duplex siRNA olignucleotides against the firefly luciferase gene (luc) were used for transfection experiments. **B.** Western blot analysis of transfected NIH3T3 cells using PHF5a (upper panel) or _-tubulin-specific (lower panel) antibodies. The Western blot was stripped and re-probed with an _-tubulin antibody to check for equal loading (20µg) of total protein.

4 Discussion

4.1 Summary of the results

The present work discuss the characterization of unique protein interactions of the PHF5a protein. In the present study following results were obtained:

- Four proteins namely U2AF^{35}, SRp40, mDomino and Ddx1 were identified as PHF5a interacting partners by using the murine 11.5-days embryo yeast two-hybrid library screening.
- PHF5a interacting domains of SR proteins U2AF^{35} and SRp40 were restricted to C-terminal arginine-serine rich domains (RS).
- Other members of the SR protein family namely SRp20, SRp30c and ASF/SF2 were shown that they can not bind to the PHF5a protein by using a directed yeast two-hybrid assay and the _-galactosidase assay.
- Protein interactions between PHF5a and U2AF^{35}, SRp40, mDomino and Ddx1 were verified by using *in vitro* coimmmunoprecipitation assays.
- Mapping of binding sites by using coimmunoprecipitation experiments and a directed yeast two-hybrid assay demonstrated that both ATP-dependent helicases mDomino and Ddx1 interact with the C-terminal segment of PHF5a. In addition, the N-terminal part of PHF5a was identified as a region responsible for binding to splicing proteins U2AF^{35} and SRp40.
- By using the yeast-three hybrid assay it was demonstrated that PHF5a is able to play a role as a bridge protein which can bind splicing proteins U2AF^{35}, SRp40 and ATP-dependent helicases Ddx1, mDomino simultaneously.

- Immunostaining on NIH3T3 cells demonstrated that the PHF5a protein is localized predominantly in the nucleus of stained cells.
- Signals from both the PHF5a-GFP fusion protein and for $U2AF^{35}$ or SRp 40 partially colocalize and are distributed in multiple speckles throughout the nucleus.
- Four splicing forms of $U2AF^{35}$ were isolated by RT-PCR and the yeast two-hybrid experiment and characterized.
- Expression of PHF5a and $U2AF^{35}$ proteins in the testis were analyzed. It was demonstrated that PHF5a shows a time and spatial coexpression with the $U2AF^{35}$ protein predominantly in the nuclei of pachytene spermatocytes. In addition, the PHF5a protein was found to be up-regulated in spermatocyte-specific GC-4spc cells as compared to spermatogonia- specific GC-1spg cells.
- PHF5a and $U2AF^{35}$ interact *in vivo* in spermatocyte-derived GC-4spc cells which was demonstrated by using *in vivo* coimmunoprecipitation assays.
- By using the *in vivo* splicing assay we could demonstrate that the PHF5a protein does not directly affect splicing of the ß-globin mini gene.
- Whole mount *in situ* hybridization on 11.5-days mouse embryos suggests that PHF5a is expressed ubiquitously during embryonic development.
- A PHF5a-GST fusion protein was able to pull down genomic DNA. Furthermore, by using the whole genome PCR and SELEX putative target DNA sequences of the PHF5a protein were identified.
- Down-regulation of PHF5a expression in NIH3T3 cells by using the RNAi technique inhibit cell proliferation.

4.2 Detection and characterization of PHF5a protein interactions

Eukaryotic messenger RNA precursors (pre-mRNAs) are typically processed before they are exported to the cytoplasm. Pre-mRNA processing steps include the addition of a cap-structure to the 5'-end, the removal of internal non-coding sequences by splicing and the generation of a new 3'-end by endonucleolytic cleavage and polyadenylation. Transcription, capping, splicing and 3'-end processing of pre-mRNAs in cells are coupled processes, but they can be analyzed independently *in vitro* (reviewed in Barabino and Keller, 1999; Bentley, 1999; Zhao et al., 1999; Hirose and Manley, 2000; Proudfoot, 2000). In RNA splicing noncoding sequences are removed which typically requires a large ribonucleoprotein complex called the spliceosome and multiple auxillary factors. Relations between different components of the RNA processing machinery are a key to understand both constitutive and alternative splicing and its coordination together with gene expression (Proudfoot, 2000).

4.2.1 PHF5a interacts with splicing proteins U2AF^{35} and SRp40

Intron removal from mRNA precursors is an essential step of gene expression in eukaryotes. The precise recognition of the intron boundaries, the 5' and 3' splice sites, is achieved by small nuclear RNPs (snRNPs) and non-snRNP proteins. The 5' splice site is initially recognized by U1 snRNP, and the 3' splice site region is recognized by U2 snRNP. Subsequent addition of the U4/U6/U5 tri-snRNP forms the spliceosome, the macromolecular complex within which splicing catalysis takes place (reviewed Burge et al., 1999; Krämer, 1996). Several sequence elements help to define the 3' splice site region in higher eukaryotes (reviewed by Reed, 2000): the branchpoint (BP) sequence, usually followed by a pyrimidine-rich sequence (the polypyrimidine tract or Py tract), and a conserved AG dinucleotide at the 3' end of the intron.

Spliceosomes are very dynamic and complex structures. Among all of the proteins involved in splicing, proteins containing arginine-serine rich domain (RS) are one of the most

characteristic and prominent. RS domains are typically localized in the C- terminal part of SR proteins and other related proteins, e.g. the U2AF- heterodimer, composed of two subunits with a size of 35 kDa (U2AF1) and 65 kDa (U2AF2), respectively.

PHF5a is implicated in the processing of pre-mRNAs (Wang et al., 2003). It was shown that homologous protein Rds3p in yeast is a critical pre-mRNA splicing factor and an integral component of the yeast spliceosome. In addition, the human homologous protein SF3b 14b was identified as a novel protein associated with the human 17S U2 snRNP and one of its stable subunits, SF3b. In the present study we could also demonstrate that the PHF5a protein binds to four proteins, namely $U2AF^{35}$, SRp40, Ddx1 and mDomino.

Using the yeast two-hybrid library screening for PHF5a interacting partners we isolated multiple overlapping cDNA clones corresponding to splicing proteins $U2AF^{35}$ and SRp40. Both proteins bear strong similarities in domain structure and contain C-terminal arginine-serine-rich (RS) domains and RNA recognition motifs (RRM) (Zhang et al., 1992). First, $U2AF^{35}$ is part of the heterodimeric complex consisting of 65-kDa ($U2AF^{65}$) and 35-kDa ($U2AF^{35}$) subunits (Zamore and Green, 1989). $U2AF^{65}$ binds specifically to the poly-pyrimidine tract via its RNA recognition motifs (RRMs) (Zamore et al., 1992) and contacts the branch point site via its RS domain, whereas $U2AF^{35}$ has been shown to function by directly interacting with the conserved AG dinucleotide at the 3' splice site of pre-mRNAs (Figure 42) (Merendino et al., 1999; Wu et al., 1999; Zorio and Blumenthal, 1999). $U2AF^{35}$ is also responsible for the recruitment of $U2AF^{65}$ by serine-arginine-rich (SR) proteins in enhancer-dependent splicing (Graveley et al., 2001). Exonic splicing enhancers (ESE) are purine-rich sequences bound by SR proteins, which stimulate splicing of pre-mRNAs containing weak 3' splice sites (Cooper, 1999; Tacke and Manley, 1999). Therefore, SR proteins can modulate alternative splice-site selection.

Second, SRp40 is a serine-arginine-rich (SR) protein—a member of a conserved family of essential splicing factors (Graveley, 2001). SR proteins share a distinctive domain structure, which consists of one or two copies of an RNA-recognition motif (RRM), followed by a characteristic C-terminal arginine/serine-rich (RS) domain (Birney et al., 1993). It has been proposed that exonic splicing enhancer-bound SR proteins recruit U2AF which involves protein interactions with the RS domain of $U2AF^{35}$ (Zu and Maniatis, 1996).

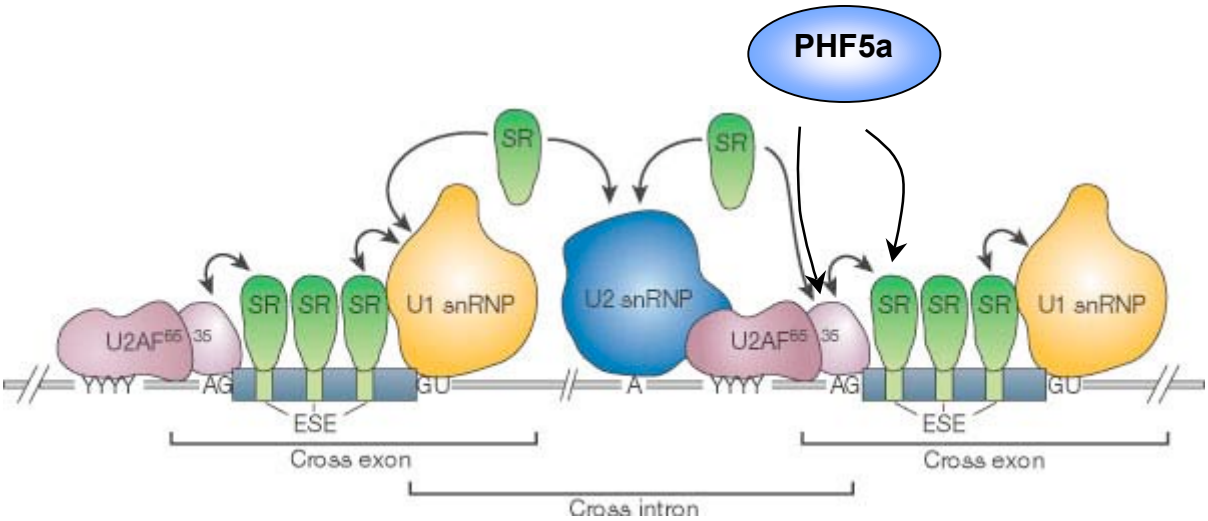


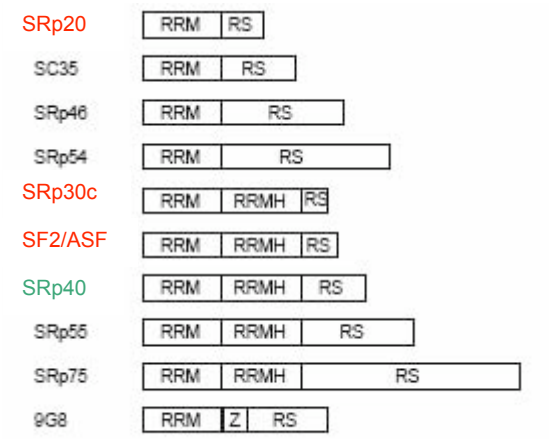
Figure 42. Schematic depiction of exon recognition and PHF5a protein-protein interactions with splicing proteins. The correct 5' (GU) and 3' (AG) splice sites are recognized by the splicing machinery on the basis of their proximity to exons. The exons contain exonic splicing enhancers (ESEs) that are binding sites for SR proteins. When bound to an ESE, the SR proteins recruit U1 snRNP to the downstream 5' splice site, and the splicing factor U2AF (65 and 35 kDa subunits) to the pyrimidine tract (YYYY) and the AG dinucleotide of the upstream 3' splice site, respectively. In turn, U2AF recruits U2 snRNP to the branchpoint sequence (A). Thus, the bound SR proteins recruit splicing factors to form a 'cross-exon' recognition complex. SR proteins also function in 'cross-intron' recognition by facilitating the interactions between U1 snRNP bound to the upstream 5' splice site and U2 snRNP bound to the branchpoint sequence. (modified and adapted from Maniatis and Tasic, 2002)

A minimal PHF5a-interacting domain of the SRp40 protein was restricted to the C-terminal RS domain. Subsequently, by using a directed yeast two-hybrid we also identified the C-terminal part of $U2AF^{35}$ containing an RS domain as a sufficient fragment maintaining

the protein interaction with PHF5a. The RS domains are required for protein–protein interactions of SR proteins with each other and with other components of the splicing machinery (Wu and Maniatis, 1993; Kohtz et al., 1994). RS domain-containing proteins are summarized in the Figure 43. RS domains also mediate subcellular localization of SR proteins (Cáceres et al., 1997). Reversible phosphorylation of SR proteins at the serine residues within the RS domains can modify protein-RNA (Tacke et al., 1997) and protein–protein interactions (Xiao and Manley 1997, 1998), as well as localization of SR proteins and recruitment to sites of active transcription.

In addition, we could show by using coimmunoprecipitation that the specific interaction between PHF5a and the RS domain of $U2AF^{35}$ is salt independent suggesting a rather hydrophobic than an ionic type of interaction between these proteins and strong association between these proteins.

A. Human SR Proteins



B. Human SR Related Proteins

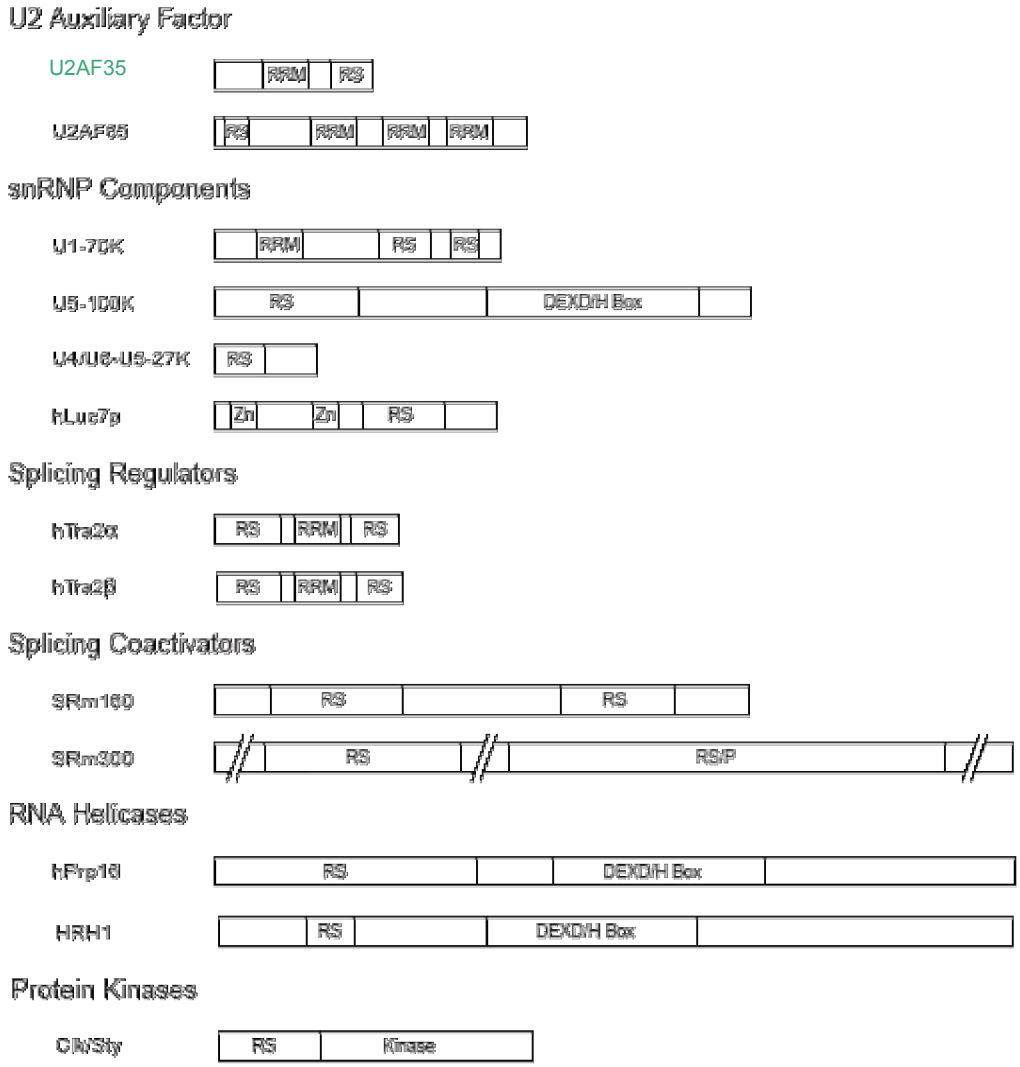


Figure 43. Schematic diagram representing human SR proteins and SR related proteins **A**: The domain structures of the known members of the human SR protein family. (RRM) RNA recognition motif; (RRMH) RRM homology; (Z) zinc knuckle, (RS): arginine/ serine-rich domain **B**. The domain structures of the human SRrps that participate in pre-mRNA splicing. All proteins, with the exception of SRm300, are drawn to scale. PHF5a interacting partners U2AF35 and SRp40 are marked in green; SRp30c, Srp20 and ASF/SF2 could not bind to PHF5a and are marked in red. (RRM) RNA recognition motif; (RS) arginine/serine-rich domain; (Zn) zinc finger; (DEXD/H Box) motif characteristic of RNA helicases (adapted from Graveley, 2000)

4.2.2 Specificity of PHF5a binding to RS domains

RS domains are redundant and can functionally substitute for each other. Furthermore, RS domains are *in vitro* functionally interchangeable and splicing specificity is determined by the RRMs in SR proteins (Chandler et al., 1997). For example, it was found that the RS domains from other SR proteins can substitute for the SF2/ASF RS domain in chicken cells lacking endogenous SF2/ASF (Wang et al., 1998).

This prompted us to test if the PHF5a protein can also interact with others RS domains or if PHF5a protein-protein interactions are restricted only to RS domains of SRp40 and U2AF^{35}. In the present study, we could demonstrate that RS domains of SRp30c, Srp20 and ASF/SF2 could not bind to PHF5a by using a directed yeast-two hybrid and a quantitative _-galactosidase assay. These results suggest that individual RS domains are not functionally equivalent *in vivo* and different protein interactions could have distinct RS domain sequence requirements. In addition it can be hypothesized that specific protein interactions between PHF5a and different RS domains can determine the regulation, or localization of individual members of the SR family of splicing factors.

Recent studies have identified PHF5a as a part of the SF3b subunit of the human 17S U2snRNP (Will et al., 2002) and several proteins (but not SR proteins) were copurified with the PHF5a yeast homologous protein Rds3p. Four of them are known components of the yeast

U2snRNP complex, namely Cus1p, Hsh49p, Hsh155p, and Rse1p (Wang et al., 2003). However, by using the yeast-two hybrid screening with PHF5a as a bait we could not identify any of proteins homologous to these reported by Wang and co-workers. This result may suggest that a different nature of interaction may exist. Moreover, complete sequencing of the yeast genome has failed to reveal proteins with the explicit characteristics of the SR family described in numerous metazoan species. This might explain why Wang et al. failed to detect interactions with $U2AF^{35}$ and SRp40. In addition, although basic mechanisms of splicing are conserved, the lack of SR proteins in yeast can be explained by the stringent consensus sequence of the splice site and by a small number of genes containing introns (Siebel et al., 1999).

4.2.3 PHF5a colocalization with $U2AF^{35}$ and SRp40 in nuclear speckles

We could show that signals from both PHF5a-GFP fusion protein and from splicing proteins $U2AF^{35}$ and SRp40 partially colocalize and are distributed throughout the nucleus, excluding the nucleolus, but are enriched in multiple speckles. From recent studies it is known that splicing factors are concentrated in a few tens of nuclear domains known as speckles. They are also more diffusely spread throughout euchromatin (Pombo and Cook, 1996). The nature of speckles is still being debated; nonetheless interchromatin granule speckles are genuine nuclear structures containing many proteins and ribonucleoprotein (RNP) complexes that play a direct role in pre-mRNA processing. Nuclear speckles also contain RNAs produced by RNA polymerase II (Pol II)-dependent transcription (Bregman et al., 1994) and nascent pre-mRNA transcripts are apparently synthesized in the nucleoplasm adjacent to the nuclear speckles (reviewed by Spector, 1993).

4.2.4 PHF5a interacts with ATP-dependent helicases Ddx1 and mDomino

The superfamily of ATP-dependent helicases (reviewed by Tannerand and Linder, 2001; Luking et al., 1998) is another rapidly expanding group of proteins involved in the RNA metabolism. Whereas some of these proteins are exclusively involved in RNA metabolisms, others have been shown to be more universal and are capable of utilising RNA-DNA or DNA substrates as well. The DEAD box which is characterized by the conserved motif Asp-Glu-Ala-Asp is a common component of RNA helicases unwinding RNA duplexes in an ATP- dependent manner. This vast family of proteins is involved in each step of RNA metabolism and processing including transcription, pre-mRNA splicing, spliceosome assembly, transport to the cytoplasm, translation, RNA decay and organellar gene expression (Luking et al., 1998).

By using the yeast two-hybrid library screening for PHF5a interacting partners we could also isolate multiple overlapping cDNA clones corresponding to ATP-dependent helicases Ddx1 and mDomino. Ddx1 is a member of the DEAD box family of putative RNA helicases (Abdelhaleem et al., 2003). They are characterized by eight conserved amino acid motifs, including the core DEAD (Asp-Glu-Ala-Asp) motif involved in ATP hydrolysis and coupling of ATPase and RNA helicase activity (Pause and Sonenberg., 1992). Ddx1 may play a role in the processing of the 3' end of pre-mRNAs (Bléoo et al., 2001) and has been found to colocalize with cleavage bodies, nuclear structures enriched in proteins involved in 3' processing of mRNAs. *In vitro* binding and co-immunoprecipitation studies confirmed the protein-protein interaction between hnRNP K and Ddx1 (Chen et al., 2002)**.** The hnRNP K protein has been identified as a component of the heterogeneous nuclear ribonucleoprotein complexes. These hnRNP proteins bind pre-mRNAs directly and appear to facilitate various stages of mRNA biogenesis (Dreyfuss et al., 1993; Weighardt et al., 1996; Krecic et al., 1999). In addition, there are observations that *in vivo* 3'-end formation often occurs before the removal of introns (Bauren et al., 1998, Nevins and Darnell, 1978). In conclusion, interactions

of PHF5a with both splicing factors U2AF^{35}, SRp40 and the RNA helicase Ddx1 may support the hypothesis that PHF5a is implicated in different aspects of pre-mRNA processing.

In contrast, the second PHF5a-interacting ATP-depended helicase mDomino contains a SWI2/SNF2-type ATPase/helicase domain, a putative DNA binding domain (a SANT domain), and a glutamine-rich (Q-rich) domain and is thought to regulate general transcriptional activity (Ogawa et al., 2003). The mDomino protein seems to be the part of the ATP-dependent chromatin-remodelling complex and belongs to the subfamily of SWI2/SNF2-type helicases. Interestingly, the C-terminal Q-rich domain of mDomino which has been found to mediate the interaction with the PHF5a protein physically associates with the transcription activation domain of C2H2 zinc finger transcription factor MZF-2A in mammalian cells as well as in yeast. Moreover, expression of the mDomino Q-rich domain, together with MZF-2A in myeloid LGM-1 cells, enhanced the MZF-2A-mediated activation of a reporter gene (Ogawa et al., 2003). These results strongly suggest that an ATP-dependent chromatin-remodelling complex containing mDomino interacts with MZF-2A to regulate gene expression in myeloid cells. In conclusion, these results support the theory that specific protein interactions between mDomino and the second zinc finger of PHF5a are responsible for the reported function of the PHF5a protein as a transcriptional activator (Oltra et al., 2003).

4.3 The PHF5a and its role in coordination of splicing and chromatin activation

Several lines of evidence indicate that chromatin remodelling and transcription occur in coordination with pre-mRNA splicing. Although i*n vitro* studies have clearly demonstrated that the reactions of transcription, capping, splicing, and mRNA 3'-end formation can take place independently, recent work indicates that *in vivo* the cell couples all of these processes

on each RNA polymerase II transcript (Neugebauer and Roth., 1997). Therefore, in this context it is not surprising that the PHF5a protein may associate with both helicases, i.e. with the pre-mRNA processing protein Ddx1 and the chromatin-remodelling protein mDomino.

Furthermore, Ddx1 represents the first DEAD box protein associated with a 3'-cleavage and polyadenylation of pre-mRNA (Bléoo et al., 2001) and contains a region with homology to heterogeneous nuclear ribonucleoprotein U (hnRNP U) (Godbout et al., 1994). The domain common to both hnRNP U and Ddx1 mediates binding of TFIIH to the RNA polymerase II holoenzyme (Kim and Nikodem, 1999). The C-terminal domain (CTD) of the largest subunit of RNA polymerase II appears to provide a platform on which to link physically transcription to splicing and other gene expression steps, such as capping, polyadenylation and termination of transcription (reviewed in Hirose and Manley, 2000; Bentley, 2002; Proudfoot et al., 2002). The mammalian RNA polymerase II CTD is composed of 52 tandem repeats of the YSPTSPS consensus peptide (Corden et al., 1985). Phosphorylation of Ser2 and/or Ser5 of this heptapeptide is a key event during the transcription cycle, with the hypophosphorylated form of RNA polymerase II being found in pre-initiation complexes and the hyperphosphorylated form in elongation complexes (ECs).

Truncation of the C-terminal domain (CTD) of the RNA polymerase II inhibits splicing, polyadenylation, and transcription termination (McCracken et al., 1997). Interestingly, the CTD binds directly to protein factors essential for RNA processing such as capping enzymes, snRNPs, and serine-arginine (SR)-like proteins (Cramer et al., 1999) Moreover, a critical role for the CTD of the RNA polymerase II was also reported for the intranuclear targeting of splicing factors to transcription sites *in vivo* (Mistreli and Spector, 1999). Subsequent functional studies showed that the identity of the RNA Polymerase II promoter affects splicing patterns of the nascent transcript (Cramer et al, 1997, 1999), further connecting transcription and splicing.

4.4 Mapping of PHF5a protein-protein interaction domains and models of PHF5a protein folding

The PHF5a protein sequence is rich in metal ligand-binding cysteines and histidines (Figure 44).

H-x(6)-C-x(12)-C-x(3)-C-x(4)-C-x(3)-C-x(13)-C-x(3)-C-x(9)-C-x(3)-C-x(11)-C-x(3)-C

Figure 44. Schematic representation of metal ligand-binding cysteines and histidines resesidues in the PHF5a sequence and the schematic diagram representing the PHF5a folding model proposed by Trappe et al. (2003) with a central PHD finger-like domain flanked by two basic regions.

These residues were proposed to form three distinct zinc-finger motifs (Klug et al., 1987). The two most carboxy-terminal zinc-fingers can fold into a plant homeodomain (PHD) finger-like structure (Aasland et al., 1995, Trappe et al., 2001). The plant homeodomain (PHD) finger is a C4HC3 zinc-finger-like motif found in nuclear proteins thought to be involved in chromatin-mediated transcriptional regulation. The PHD finger motif is reminiscent of, but distinct from the C3HC4 type RING finger (according to the PFAM database http://pfam.wustl.edu).

Alternatively, the two amino-terminal zinc fingers almost match the consensus sequence for a RING-finger motif (Kosarev et al., 2002) as proposed by Oltra et al. (2003) (Figure 45). The RING-finger is a specialized type of Zn-finger of 40 to 60 residues that binds

two atoms of zinc, and is probably involved in mediating protein-protein interactions. There are two different variants, the C3HC4-type and a C3H2C3-type, which are clearly related despite the different cysteine/histidine pattern. The latter type is sometimes referred to as 'RING-H2 finger' (according to the PFAM database http://pfam.wustl.edu)

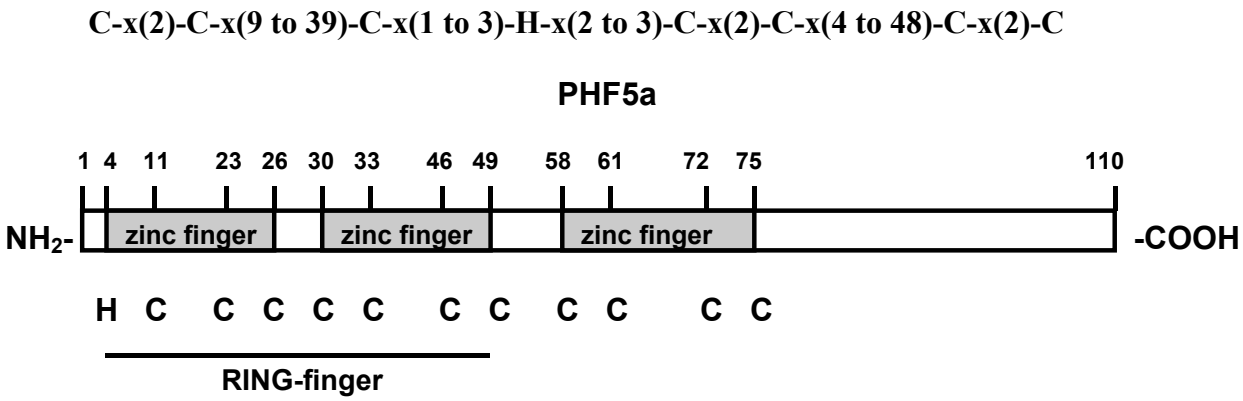


Figure 45. Schematic representation of the consensus sequence for a RING-finger motif and the schematic diagram representing the PHF5a folding model proposed by Oltra et al. (2003) with two N-terminal zinc fingers folding into the RING-finger domain.

Mapping experiments demonstrated that a minimal fragment of the PHF5a protein is able to interact with RS domains of U2AF^{35} and SRp40 and this region could be restricted to the N-terminal part of PHF5a (Figure13). A PHF5a deletion mutant bearing all three putative zinc fingers was sufficient to maintain an interaction with RS domains of U2AF^{35} and SRp40 in the yeast two-hybrid assay, whereas a PHF5a deletion mutant with two N-terminal zinc fingers (putative RING-finger) shows only weak affinity to RS domains. Interestingly, all three N-terminal zinc fingers of the PHF5a protein have been previously shown to be required for the activation of the estrogen receptor _ (ER_) (Oltra et al., 2003).

Both Ddx1 and mDomino associated with the C-terminal segment of PHF5a which could be more precisely restricted to the C-terminal zinc finger. Previously, this domain has

been described as a fragment capable of binding to the promoter of connexin 43 (Oltra et al., 1999, 2003).

In addition, Trappe et al. (2001) proposed that a central PHD-finger-like motif is flanked by two highly basic domains. Possible protein interactions with PHF5a by ionic attraction of basic domains were excluded by using mutagenesis studies and subsequently, by using the coimmunoprecipitation assay.

Taken together, PHF5a seems to posses at least two distinct functional protein-binding domains. The PHF5a protein contains one domain for binding to RS domain-containing proteins, i.e. U2AF^{35}, SRp40 and a second one, for binding to ATP-dependent helicases mDomino and Ddx1. Moreover, our mapping studies and the results of Oltra et al. (2003) argue against the folding model proposed by Trappe et al. (2003) with a central PHD finger-like domain. However, the final answer to this question of whether PHF5a folds into either a RING or a PHD finger-like structure will only be solved by determining the three-dimensional structure of the PHF5a protein.

4.5 PHF5a simultaneously binds splicing proteins U2AF^{35}, SRp40 and ATP-dependent helicases Ddx1 and mDomino

By using a yeast-three hybrid assay we could demonstrate that PHF5a is able to bind simultaneously RS domain-containing proteins U2AF^{35}, SRp40 and ATP-dependent helicases Ddx1, mDomino. This result indicates that PHF5a plays a role as a bridge protein.

The large group of ATP-dependent helicases contain a RS domain (Ono et al., 1994, Teigelkamp et al., 1997, Ortlepp et al., 1998), which could allow a direct interaction with SR proteins. Both RNA helicases Ddx1 and mDomino seem to interact with splicing proteins U2AF^{35} and SRp40 only indirectly via the PHF5a protein as it has been proven by the yeast three-hybrid experiment. In this context a functional link between chromatin activation

mediated by mDomino, 3'-pre-mRNA processing mediated by the Ddx1 helicase and 3' splicing site recognition mediated by U2AF^{35} and SRp40 would exist via the PHF5a protein (Figure 46).

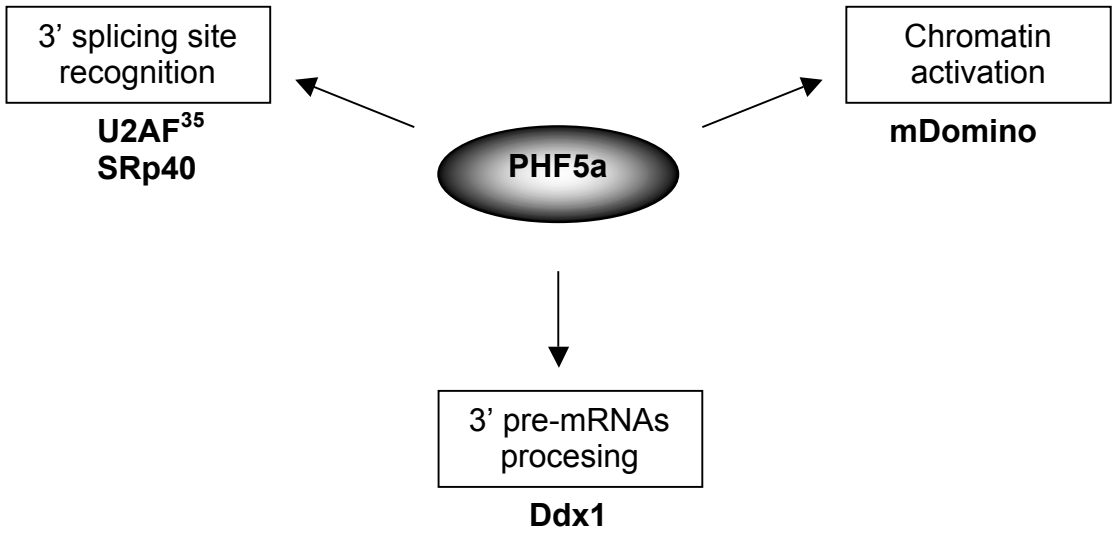


Figure 46. Schematic diagram presenting that the PHF5a protein may act as a linker between chromatin activation mediated by mDomino, 3'-pre-mRNA processing mediated by the Ddx1 helicase and 3' splicing site recognition mediated by U2AF^{35} and SRp40.

4.6 The expression of U2AF1 gene results in four alternative mRNA isoforms by alternative splicing

We also provide evidence for the existence of alternative murine U2Af1 gene transcripts encoding different U2AF^{35} isoforms. Four mRNA isoforms (termed U2AF^{35}a-d) are produced by alternative splicing of the murine U2AF1 gene. In the splicing variants U2AF^{35}c and U2AF^{35}d exon 2 and exon 2-3 are skipped, respectively. Both variants of U2AF^{35} (c, d) contain premature stop codons as a result of frame shifts. *In vivo* mRNAs containing premature stop codons are selectively degraded by a surveillance mechanism termed nonsense-mediated decay (NMD) (Losson. and Lacroute, 1979, Neu-Yilik et al., 2001) to prevent the synthesis of potentially harmful truncated proteins. In mammalian cells, a

termination codon is usually recognized as a premature stop codon if it is located more than 50-54 nucleotides 5' to the final exon-exon junction (Maquat et al., 2004). Recently, it has been presented by Pacheco et al. (2004), that in human alternative splicing forms of U2AF^{35} with a premature stop codon are targeted for the nonsense mediated RNA decay (NMD).

The presence of at least two murine splicing variants with premature stop codons suggests a possible mechanism of U2AF^{35} regulation by the NMD. Relative concentration of splicing factors can affect recognition of splicing sites (Caceres and Krainer, 1997). Moreover, it has been shown that overexpression of the SR protein-SRp20 regulates alternative splicing of its own pre-mRNA by promoting the inclusion of exon 4 (Jumaa and Nielsen, 1997). Therefore, it is possible that overexpression of U2AF^{35} will promote alternative splicing of itself and will result in the accumulation of U2AF^{35} mRNAs containing premature stop codons. This would lead to a feedback regulation, a common mechanism controlling the level of splicing regulators such as tra-2 (Mattox and Baker, 1991), SWAP (Zachar et al., 1987), Sxl (Bell et al., 1991) and the SR-like protein HTRA2-BETA1 (Stoilov et al., 2004).

In conclusion, it has to be proven by using additional experiments which conditions lead to the skipping of exon 2 in the splicing form U2AF^{35}c and exons 2 and 3 in the splicing form U2AF^{35}d.

We also identified alternative U2AF^{35} splicing variants with exon 3 replaced by exon Ab producing U2AF^{35} isoform b (U2AF^{35}b). Recently, this splicing variant has been described in human (Pacheco et al., 2004) and encodes a protein which is 97% identical to the known U2AF^{35} protein. Amino acid changes in the isoform U2AF^{35}b are localized in the RMM domain mediating dimerization with U2AF^{65}. However Pacheco et al. (2004) could present evidences that isoform U2AF^{35}b can replace U2AF^{35} in all experiments and is able to effectively bind to U2AF^{65}. Our results indicate that both U2AF^{35}b and the truncated U2AF^{35}

form containing sequences downstream of exon 3 have the same binding affinity to PHF5a as compared to the previously known U2AF^{35} isoform (U2AF^{35}a).

In addition, we could not detect the recently described alternative human U2AF^{35} splicing variant with a simultaneous inclusion of exon Ab and exon 3 resulting in an in frame premature stop codon (referred by Pacheco et al. as U2AF^{35}c).

4.7 PHF5a can act as a DNA binding protein

Of additional interest is the putative role of PHF5a as a transcription regulator. The results presented by Oltra et al. (2003) have shown that the rat PHF5a ortholog Ini actively participates in the estrogen response by binding to the proximal region of the connexin 43 promoter. Therefore, Ini may belong to a new class of small transcription factors or cofactors that are highly conserved throughout evolution. PHF5a expression has been shown to be essential in yeast (Oltra et al., 1999, Giaever et al., 2002) and more recently in *Caenorhabditis elegans* (Trappe et al., 2002). In addition, PHF5a shows a weak homology to the homeo-domain (PHD) zinc finger proteins. PHD fingers are described to be engaged in DNA binding and at least a few PHD fingers are present in known transcription factors (Aasland et al., 1995).

In fact, by using recombinant PHF5a-GST fusion protein we could precipitate genomic DNA fragments which is consistent with the results presented by Oltra et al.(2003) suggesting that PHF5a may bind to the promoter region of the rat gene connexin 43. In order to identify DNA target sequences which are recognized by the PHF5a protein the whole genome PCR assay (WG-PCR) was employed. By using the Human Genomic Blast tool (NCBI) and Ensembl Genome Browser we failed to identify promoter regions and other 5' regulatory elements matching the DNA sequences identified by the WG-PCR assay.

Furthermore, DNA consensus binding sequences recognized by PHF5a and found by using SELEX could not be identified within the fragments selected by the whole genome PCR.

At this point it should be pointed out that in the PHF5a amino acid sequence at positions 77, 84, 91, and 98 leucine, isoleucine or glycine residues are located which could fold into a leucine zipper structure found in many transcription factors (Vinson et al., 1989; Williams et al., 1991; Osada et al., 1997) (Figure 47).

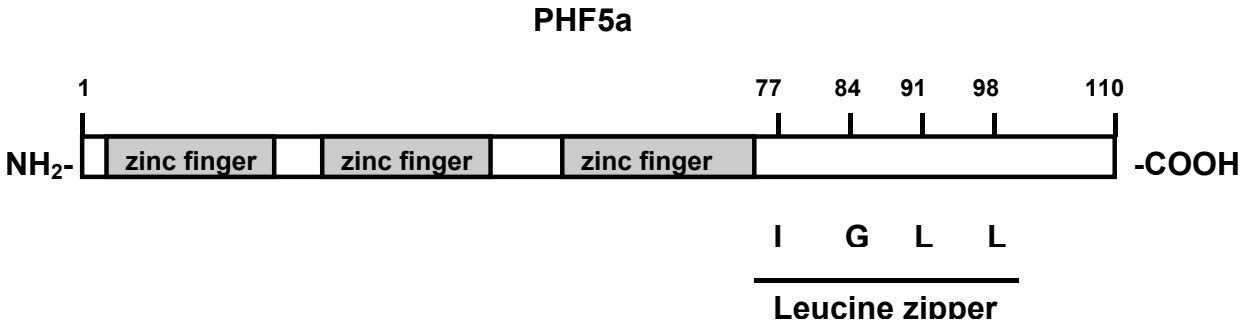


Figure 47. The schematic diagram representing the PHF5a protein with marked amino acid residues (isoleucine^{77}, glycine^{84}, leucine^{91}, leucine^{98}) which can fold into a leucine zipper structure typical for many transcription factors.

Both, a 100% level of conservation of PHF5a protein sequence among vertebrates together with an ubiquitous pattern of expression also in tissues that are not responsive to estrogen suggest that PHF5a could function as a more general transcription factor. Finally, the putative implication of the PHF5a protein in transcriptional activation and specific interactions with splicing proteins raises the possibility that PHF5a is involved in both the coordination of splicing and active transcription.

4.8 PHF5a as a candidate gene for meiotic differentiation

There is very little known about mechanisms that regulate pre-mRNA processing during development and maturation of germ cells in mammals (Venables et al., 1999). Transition from diploid spermatogonias to functional haploid sperm cells is accompanied by

substantial changes in strategies for regulating gene expression. Increase in post-transcriptional regulation by alternative splicing events is one of the possible mechanisms. PHF5a is a putative candidate gene for a factor taking part in this unique process.

Our observation that PHF5a is strongly expressed in the spermatocyte-specific cell line GC-4spc as compared to the spermatogonia specific cell line GC-1spg is consistent with results of the immunolocalization studies. We could demonstrate that PHF5a is co-localized with $U2AF^{35}$ predominantly in the nuclei of pachytene spermatocytes. This raises the possibility that PHF5a and $U2AF^{35}$ take part in the meiotic differentiation of male germ cells. Rearrangements in the PHF5a protein localisation in the later germ cell stages are similar to those found for RMB germ-cell-specific RNA binding protein and other splicing components (Elliot et al., 1998; Richler et al., 1994; Moussa et al., 1994). Transition from spermatocytes to sperm is accompanied with downregulation of chromatin activation and these changes may correlate with splicing and finally result in few punctuates foci being the storage place for proteins involved in mRNA processing. Such global rearrangements in the localisation of snRNPs were also described by Antoniou et al. (1993) for murine erythroleukaemia cells after induction of differentiation. Widespread nuclear snRNPs in later stages of differentiation disassociate from coil bodies and aggregate in foci corresponding to the interchromatin granules. In conclusion, we provide evidence that indicates that the nuclear organization of PHF5a and $U2AF^{35}$ undergoes dynamic modifications in cell types undergoing division and differentiation *in vivo*, and so is more complex than in the more homogeneous populations of cells grown in tissue cultures.

4.9 The role of PHF5a in cell cycle progression

The deletion of the PHF5a homologous gene ini1 from *S. pombe* was lethal, and loss of Ini1 resulted in a block to cell cycle progression (Oltra et al., 2003). Yeast cells lacking

Ini1 are incapable of completing a successful mitosis but a direct mechanism is currently unknown. One attractive hypothesisis is that ini1 gene product activates promoters, which control expression of genes required for cell cycle progression.

S. pombe Ini1 protein is 74% identical to both human and murine PHF5a sequences at the amino acid level. The putative function of PHF5a in mouse NIH3T3 cells was studied by using the gene silencing technique through RNA interference (RNAi). A significant inhibition of cell proliferation in PHF5A duplex siRNA-transfected NIH3T3 cells was observed indicating that PHF5a can also regulate cell cycle progression in mammalian cells. However a direct mechanism how PHF5a may affect cell proliferation is unknown. It can be speculated that the PHF5a protein is required for general splicing of pre-mRNAs of at least some genes which encode proteins required for normal cell cycle progression.

4.10 PHF5a is involved in vertebrate development

Trappe et al. (2001) has reported that in *C. elegans* spatial PHF5a expression is muscle-specific with an expression in the developing pharynx, in body wall muscular structures, and in the anal muscles. Furthermore, it was demonstrated that PHF5a is essential in the morphogenetic phase of *C. elegans* embryonic development as well as in young larvae. In contrast, down-regulated expression of the PHF5a gene by using RNAi did not show an evident phenotype to adult worms. In the insertional mutagenesis screen in the zebra fish PHF5a was identified as a gene essential for early vertebrate development (Golling et al., 2002). The zebra fish PHF5a mutant was characterized by a curved body, a constricted yolk sack extension and necrosis of the central nervous system. Interestingly, in the present study the strongest PHF5a expression in murine embryos was observed in regions of the yolk sack and the spinal cord by using whole mount in situ hybridization. Moreover, the zebra fish insertional mutant of the U2AF1 gene (a PHF5a interacting partner) shows also necrosis of the

central nervous system (Golling et al., 2002). In conclusion, the similarities of both the PHF5a insertional mutant and the U2AF1 mutant in zebra fish together with the expression pattern of PHF5a in the spinal cord indicate that interaction of PHF5a and U2AF1 are evolutionary conserved.

.

4.11 Further perspectives

In view of the observations presented in this studies regarding the protein interaction of PHF5a further experiments are needed to obtain conclusive evidence that the PHF5a protein is indeed involved in processing of pre-mRNAs. The identification of additional proteins that interact with the PHF5a protein will certainly support or negate the potential implications discussed here and could reveal a connection between splicing and gene expression. The yeast two-hybrid screening used in the present study has several advantages, but has also certain intrinsic limitations. The yeast two-hybrid assay is highly sensitive and allows for detecting interactions that are not detected by other methods. Another advantage is that the interactions are detected within the native environment of the cell. In addition, the use of genetic-based organisms such as yeast cells allows for the screening of a large number of variants to detecting both, those that either interact strongly or those which interact less strongly. The main limitation of the yeast two-hybrid system is that interactions are restricted to proteins that can be localized to the nucleus, thus preventing its use with certain extracellular proteins. Furthermore, proteins must be able to fold and exist stably in yeast cells and to retain activity as fusion proteins. In addition, all interactions depending on a posttranslational modification that does not occur in yeast cells will not be detected. Keeping in mind that all limitations which are listed here, alternative methods of detecting PHF5a protein-interactions should be considered. These may include protein affinity chromatography where a protein is covalently coupled to a matrix such as Sepharose under controlled

conditions and used to select ligand proteins that bind and are retained from an appropriate extract. Main advantages over the yeast two-hybrid system include, first, incredible sensitivity and second, this technique tests all proteins in an extract equally.Thus, extracted proteins that are detected have successfully competed with the rest of the population of proteins. Third, interactions that depend on a multisubunit tethered protein can be detected which is crucial for investigating large spliceosome complexes. Moreover, this technique is particularly attractive, because it was shown that the recombinant PHF5a protein can be stably expressed in different systems.

Our polyclonal anti-PHF5a antibodies open new possibilities in detecting novel protein-protein connections as they can be used in coimmunoprecipitation assays. The basic experiment is simple, i.e. cell lysates are generated, the antibody is added, the antigen is precipitated and washed, and bound proteins are eluted and analyzed. This method can bring the most valuable and authentic evidences for an interaction that exists *in vivo,* because the antigen and the interacting proteins are present in the same relative concentrations and complexes are already in their natural state (posttranslational modification.

Of additional interest are interaction studies of PHF5a with both ATP-dependent helicases and RS-containing proteins. The mapping studies presented in this work could not give conclusive evidences which domains of the PHF5a protein are responsible for maintaining specific protein-protein interactions. The isolation of protein interaction modules that bind with high specificity to a given target protein, i.e. RS proteins and ATP-dependent helicases could be performed by using mutagenesis studies spanning crucial cysteines and histidines in the centre of putative PHF5a zinc fingers. Subsequently, interaction studies of the mutated PHF5a protein together with the determination of its three-dimensional structure will reveal which amino acid motifs are responsible for protein-protein interactions.

There is also a clear need for further studies to confirm the reported properties of PHF5a as a general activator of gene expression. Although the DNA-binding properties of the PHF5a protein could be demonstrated in the present work, further studies are needed to select regulatory DNA elements bound by PHF5a. The whole genome PCR and SELEX have been proven to be powerful tools in determining protein binding sites on DNA, however, our experiments failed to identify specific regulatory elements in the 5' region of putative target genes. Therefore, alternative approaches should be performed, e.g. chromatin immunoprecipitations (ChIP). By using ChIP DNA sequences bound by the protein of interest are precipitated and the DNA will be labeled and hybridized to a recently designed genome wide DNA microarray (Ng et al., 2002, Ren et al., 2000, Weinmann et al., 2002). The resulting microarray data then identify DNA sequences which were bound by the protein of interest.

Finally, there is a need to determine splicing requirements for the PHF5a protein. Our anti-PHF5a antibodies enable us to perform *in vitro* splicing assays in immunodepleted nuclear extracts. Splicing reactions are typically carried out using S100 nuclear extracts complemented with SR proteins, or partially purified fractions derived from the crude extracts. Extracts derived from HeLa cells are used most commonly. The pre-mRNA substrates are usually prepared by *in vitro* runoff transcription with a bacteriophage polymerase. The intermediates and products of splicing are most conveniently visualized by urea/polyacrylamide gel electrophoresis (urea-PAGE) and autoradiography, which requires the use of labeled pre-mRNA substrate (Mayeda and Krainer, 1999).

5 References

1. **Aasland, R., Gibson, T. J. and Stewart, A. F.** (1995). The PHD finger: implications for chromatin-mediated transcriptional regulation. *Trends Biochem. Sci.* **20,** 56-59.
2. **Abdelhaleem, M., Maltais, L. and Wain, H**. (2003). The human DDX and DHX gene families of putative RNA helicases. *Genomics.* **81**, 618-622.
3. **Antoniou, M., Carmo-Fonseca, M., Ferreira, J. and Lamond, A. I.** (1993). Nuclear Organization of Splicing snRNPs during Differentiation of Murine Erythroleukemia Cells In Vitro. *J. Cell Biol.* **123,** 1055-1068.
4. **Barabino, S. M. and Keller W.** Last but not least: regulated poly (A) tail formation. *Cell.* **99**, 9-11.
5. **Bauren, G., Belikov, S. and Wieslander, L.** (1998). Transcriptional termination in the Balbiani ring 1 gene is closely coupled to 3'-end formation and excision of the 3'-terminal intron. *Genes Dev.* **12**, 2759-2769.
6. **Bell, L. R., Horabin, J. I., Schedl, P. and Cline, T. W.** (1991). Positive autoregulation of sex-lethal by alternative splicing maintains the female determined state in Drosophila. *Cell.* **65,** 229-239.
7. **Bentley, D.** Coupling RNA polymerase II transcription with pre-mRNA processing. *Curr. Opin. Cell Biol.* **11**, 347-351.
8. **Birney, E., Kumar, S. and Krainer, A. R.** (1993). Analysis of the RNA-recognition motif and RS and RGG domains: conservation in metazoan pre-mRNA splicing factors. *Nucleic Acids Res.* **21**, 5803-5816.

9. **Bléoo, S., Sun, X., Hendzel, M. J., Rowe, J. M., Packer, M. and Godbout, R.** (2001). Association of Human DEAD Box Protein DDX1 with a Cleavage Stimulation Factor Involved in 3'-End Processing of Pre-mRNA. *Mol. Biol. Cell*. **12**, 3046–3059.

10. **Bregman, D. B., Du, L., Li, Y., Ribisi, S. and Warren, S. L.** (1994). Cytostellin distributes to nuclear regions enriched with splicing factors. *J. Cell Sci.* **107**, 387-396.

11. **Burge, C. B., Tuschl, T. and Sharp, P. A.** (1999) Splicing of precursors to mRNAs by the spliceosomes. In Gesteland,R.F., Chech,T.R. and Atkins,J.E. (eds), *The RNA World*, 2nd edn. Cold Spring Harbor Press, Cold Spring Harbor, NY.

12. **Caceres, J. F., Misteli, T., Screaton, G. R., Spector, D. L. and Krainer, A. R.** (1997). Role of the modular domains of SR proteins in subnuclear localization and alternative splicing specificity. *J. Cell Biol*. **138**, 225-238.

13. **Chandler, S. D., Mayeda, A., Yeakley, J. M., Krainer, A. R. and Fu, X. D.** (1997). RNA splicing specificity determined by the coordinated action of RNA recognition motifs in SR proteins. *Proc. Natl Acad. Sci. U S A*. **94**, 3596-3601.

14. **Chen, H. C., Lin, W. C., Tsay, Y. G., Lee, S. C. and Chang, C.J.** (2002).An RNA helicase, DDX1, interacting with poly(A) RNA and heterogeneous nuclear ribonucleoprotein K. *J. Biol Chem.* **277,** 40403-40409.

15. **Chen, Z.Q., Lefebvre, D., Bai, X. H., Reaume, A., Rossant J. and Lye, S. J.** (1995). Identification of two regulatory elements within the promoter region of the mouse connexin 43 gene. *J. Biol Chem.* **270**, 3863-3868.

16. **Cooper, T. A.** (1999). Defining pre-mRNA cis elements that regulate cell-specific splicing. *Methods Mol. Biol.* **118,** 391-403.

17. **Corden, J. L., Cadena, D. L., Ahearn, J. M. Jr and Dahmus, M. E.** A unique structure at the carboxyl terminus of the largest subunit of eukaryotic RNA polymerase II. *Proc. Natl Acad. Sci. U S A*. **82,** 7934-7938.

18. **Cramer, P., Caceres, J. F., Cazalla, D., Kadener, S., Muro, A. F., Baralle, F. E. and Kornblihtt, A. R.** (1999). Coupling of transcription with alternative splicing: RNA pol II promoters modulate SF2/ASF and 9G8 effects on an exonic splicing enhancer. *Mol Cell.* **4**, 251-258.

19. **Cramer, P., Pesc,e C. G., Baralle, F. E. and Kornblihtt, A. R. (**1997). Functional association between promoter structure and transcript alternative splicing. *Proc. Natl Acad. Sci. U S A.* **94**, 11456-11460.

20. **Dreyfuss, G., Matunis, M. J., Pinol-Roma, S. and Burd, C. G.** (1993). hnRNP proteins and the biogenesis of mRNA. *Annu. Rev. Biochem*. **62**, 289-321.

21. **Elliott, D. J., Oghene, K., Makarov, G., Makarova, O., Hargreave, T. B.,Chandley, A. C., Eperon, I. C. and Cooke, H. J.** (1998). Dynamic changes in the subnuclear organisation of pre-mRNA splicing proteins and RBM during human germ cell development. *J. Cell Sci.* **111**, 1255-1265.

22. **Feinberg, A. P., Vogelstein, B.** (1989). A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. *Anal Biochem.* **123,** 6-13.

23. **Giaever, G., Chu, A. M., Ni, L., Connelly, C., Riles, L., Veronneau, S., Dow, S., Lucau-Danila, A., Anderson, K., Andre, B., Arkin, A. P., Astromoff, A., El-Bakkoury, M., Bangham, R., Benito, R., Brachat, S., Campanaro, S., Curtiss, M., Davis, K., Deutschbauer. A,, Entian, K. D., Flaherty, P., Foury, F., Garfinkel, D. J., Gerstein, M., Gotte, D., Guldener, U., Hegemann, J. H., Hempel, S., Herman, Z., Jaramillo, D. F., Kelly, D. E., Kelly, S. L., Kotter, P., LaBonte, D., Lamb, D. C., Lan, N., Liang, H., Liao, H., Liu, L., Luo, C., Lussier, M., Mao, R., Menard, P., Ooi, S. L., Revuelta, J. L., Roberts, C. J., Rose, M., Ross-Macdonald, P. , Scherens, B., Schimmack, G., Shafer, B., Shoemaker, D. D., Sookhai-Mahadeo, S., Storms, R. K., Strathern, J. N., Valle, G., Voet, M., Volckaert, G., Wang, C.**

Y., Ward, T. R., Wilhelmy, J., Winzeler, E. A., Yang, Y., Yen, G., Youngman, E., Yu, K., Bussey, H., Boeke, J. D., Snyder, M., Philippsen, P., Davis, R. W. and Johnston, M. (2002). Functional profiling of the Saccharomyces cerevisiae genome. *Nature*. **418**, 387-391.

24. Gietz, D., St. Jean, A., Woods, R. A. and Schiestl, R. H. (1992). Improved method for high efficiency transformation of intact yeast cells. *Nucleic Acids Res*. **20,**1425.

25. Godbout, R., Hale, M. and Bisgrove, D. (1994). A human DEAD box protein with partial homology to heterogeneous nuclear ribonucleoprotein U. *Gene.* **138**, 243-235.

26. Golling, G., Amsterdam, A., Sun, Z., Antonelli, M., Maldonado, E., Chen, W., Burgess, S., Haldi, M., Artzt, K., Farrington, S., Lin, S. Y., Nissen, R. M. and Hopkins, N. (2002). Insertional mutagenesis in zebrafish rapidly identifies genes essential for early vertebrate development. *Nat. Genet.* **31**, 135-140.

27. Graveley, B. R. (2000). Sorting out the complexity of SR protein functions. *RNA*. **6**, 1197–1211.

28. Graveley, B. R., Hertel, K. J. and Maniatis, T. (2001). The role of U2AF35 and U2AF65 in enhancer-dependent splicing. *RNA*. **7,** 806-818.

29. Guth, S., Martinez, C., Gaur, R. K. and Valcarcel, J. (1999). Evidence for substrate-specific requirement of the splicing factor U2AF(35) and for its function after polypyrimidine tract recognition by U2AF(65). *Mol. Cell. Biol*. **19**, 8263-8271.

30. Guth, S., Tange, T. O., Kellenberger, E. and Valcarcel, J. (2001). Dual function for U2AF(35) in AG-dependent pre-mRNA splicing. *Mol. Cell. Biol*. **21,** 7673-7681.

31. Hanahan, D. (1983). Studies on transformation of E. coli with plasmids. *J. Mol. Biol.* **166**, 557-580.

32. Hill, J., Donald, K. A. and Griffiths, D. E. (1991). DMSO-enhanced whole cell yeast transformation. *Nucleic Acids Res*. **19,** 5791.

33. Hirose, Y. and Manley, J. L. (2000). RNA polymerase II and the integration of nuclear events. *Genes Dev.* **14,** 1415-1429.

34. Hirose, Y., Tacke, R. and Manley, J. L. (1999). *Genes Dev.* **13**, 1234-1239.

35. Hofmann, M. C., Narisawa, S., Hess, R. A., Millan, J.L. (1992). Immortalization of germ cells and somatic testicular cells using the SV40 large T antigen. *Exp. Cell Res*. **201,** 417-435.

36. Ito, H., Fukada, Y., Murata, K. and Kimura, A. (1983). Transformation of intact yeast cells treated with alkali cations. *J. Bacteriol.* **153,**163–168.

37. Jankowsky, E. and Jankowsky, A. (2000). The DExH/D Protein family database. *Nucleic Acid Res.* **28**, 333-334.

38. Jumaa, H. and Nielsen, P. J. (1997). The splicing factor SRp20 modifies splicing of its own mRNA and ASF/SF2 antagonizes this regulation. *EMBO J*. **16**, 5077-5085.

39. Kameoka, S., Duque, P. and Konarska, M. M. p54(nrb) associates with the 5' splice site within large transcription/splicing complexes. *EMBO J.* **231**,782-791.

40. Kim, M. K. and Nikodem, V. M. (1999). hnRNP U inhibits carboxy-terminal domain phosphorylation by TFIIH and represses RNA polymerase II elongation. *Mol. Cell Biol.* **19,** 6833-6844.

41. Kinzler, K. **and Vogelstein B.** (1989). Whole genome PCR: application to the identification of sequences bound by gene regulatory proteins. *Nucleic Acids Res.* **17**, 3645-3653.

42. Klug, A. and Rhodes, D. (1987). Zinc fingers: a novel protein fold for nucleic acid recognition. *Cold Spring Harb. Symp. Quant. Biol.* **52,** 473-482.

43. Kohtz, J. D., Jamison, S. F., Will, C. L., Zuo, P., Luhrmann, R., Garcia-Blanco, M. A. and Manley, J. L. (1994). Protein-protein interactions and 5'-splice-site recognition in mammalian mRNA precursors. *Nature.* **368,** 119-124.

44. Kosarev, P., Mayer, K. F. and Hardtke, C. S. (2002). Evaluation and classification of RING-finger domains encoded by the Arabidopsis genome. *Genome Biol.* **3,** RESEARCH0016.

45. Krainer, A. R. (1995). Identification and characterization of three members of the human SR family of pre-mRNA splicing factors. *EMBO J.* **14,** 4336-4349.

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

47. Krecic, A. M. and Swanson, M. S. (1999). hnRNP complexes: composition, structure, and function. *Curr. Opin. Cell Biol.* **11,** 363-371.

48. Losson, R. and Lacroute, F. (1979). Interference of nonsense mutations with eukaryotic messenger RNA stability. *Proc. Natl Acad. Sci. U S A.* **76**, 5134-5137.

49. Luking, A., Stahl, U. and Schmidt, U. (1998). The protein family of RNA helicases. *Crit. Rev. Biochem. Mol. Biol.* **33,** 259-296.

50. Lyngso, C., Bouteiller, G., Damgaard, C. K., Ryom, D., Sanchez-Munoz, S., Norby, P. L., Bonven, B. J. and Jorgensen, P. (2000). Interaction between the transcription factor SPBP and the positive cofactor RNF4. An interplay between protein binding zinc fingers. *J. Biol Chem.* **275,** 26144-26149.

51. Mattox, W. and Baker, B. S. (1991). Autoregulation of the splicing of transcripts from the transformer-2 gene of Drosophila. *Genes Dev.* **5,** 786-796.

52. Mayeda A., Screaton G. R., Chandler S. D., Fu X. D. and Krainer A. R. (1999). Substrate Specificities of SR Proteins in Constitutive Splicing Are Determined by Their RNA Recognition Motifs and Composite Pre-mRNA Exonic Elements. *Mol. Cell. Biol*. **19**, 1853–1863.

53. McCracken, S., Fong, N., Rosonina, E., Yankulov, K., Brothers, G., Siderovski, D., Hessel, A., Foster, S., Shuman, S. and Bentley, D. L. (1997). 5'-Capping

enzymes are targeted to pre-mRNA by binding to the phosphorylated carboxy-terminal domain of RNA polymerase II. *Genes Dev.* **11**, 3306-3318.

54. Merendino, L., Guth, S., Bilbao, D., Martinez, C. and Valcarcel J. (1999). Inhibition of msl-2 splicing by Sex-lethal reveals interaction between U2AF35 and the 3' splice site AG. *Nature.* **402,** 838-841.

55. Mighell, A. J., Smith, N. R., Robinson, P. A. and Markham, A. F. (2000). Vertebrate pseudogenes. *FEBS Lett*. **468**, 109-114.

56. Misteli, T. and Spector, D. L. (1999). RNA polymerase II targets pre-mRNA splicing factors to transcription sites in vivo. *Mol. Cell.* **3**, 697-705.

57. Monneron, A., and Bernhard, W. (1969). Fine structural organization of the interphase nucleus in some mammalian cells, *J Ultrastruct Res.* **27**, 266-88.

58. Moosmann, P., Georgiev, O., Le Douarin, B., Bourquin, J.P. and Schaffner, W. (1996). Transcriptional repression by RING finger protein TIF1 beta that interacts with the KRAB repressor domain of KOX1. *Nucleic Acids Res*. **24**, 4859-4867.

59. Moussa, F., Oko, R. and Hermo, L. (1994). The immunolocalization of small nuclear ribonucleoprotein particles in testicular cells during the cycle of the seminiferous epithelium in the adult rat. *Cell Tissue Res*. **278,** 363-378.

60. Neugebauer, K. M. and Roth, M. B. (1997). Distribution of pre-mRNA splicing factors at sites of RNA polymerase II transcription. *Genes Dev*. **11,** 1148-1159.

61. Neish, A. S., Anderson, S. F., Schlegel, B. P., Wie, W. and Parvin, J. D. (1998). Factors associated with the mammalian RNA polymerase II holoenzyme. *Nucleic Acids Res.* **26**, 847-853.

62. Neu-Yilik, G., Gehring, N. H., Thermann, R., Frede, U., Hentze, M. W. and Kulozik, A. E. (2001). Splicing and 3' end formation in the definition of nonsense-mediated decay-competent human beta-globin mRNPs. *EMBO J.* **20**, 532-540.

63. Nevins, J.R. and Darnell, J. E. Jr. (1978). Steps in the processing of Ad2 mRNA: poly(A)+ nuclear sequences are conserved and poly(A) addition precedes splicing. *Cell.* **15,** 1477-1493.

64. Ng, H. H., Robert, F., Young, R. A. and Struhl, K. (2002). Genome-wide location and regulated recruitment of the RSC nucleosome-remodeling complex. *Genes Dev.* **16,** 806-819.

65. Ogawa, H., Ueda, T., Aoyama, T., Aronheim, A., Nagata, S. and Fukunaga, R. (2003). A SWI2/SNF2-type ATPase/helicase protein, mDomino, interacts with myeloid zinc finger protein 2A (MZF-2A) to regulate its transcriptional activity. *Genes Cells*. **8**, 325-339.

66. Oltra, E., Pfeifer, I. and Werner, I. (2003). Ini, a Small Nuclear Protein that Enhances the Response of the Connexin43 Gene to Estrogen. *Endocrinology*. **144**, 3148–3158.

67. Oltra, E., Verde, F., Werner, R. and D'Urso, G. (2004). A novel RING-finger-like protein Ini1 is essential for cell cycle progression in fission yeast. *J. Cell Sci*. **117**, 967-974.

68. Oltra, E. and Werner, R. (1998). Cloning of a potential regulatory protein involved in Connexin43 gene expression. Gap junctions. *Amsterdam IOS Press*, 321–325.

69. Ono, Y., Ohno, M. and Shimura, Y. (1994). Identification of a putative RNA helicase (HRH1), a human homolog of yeast Prp22. *Mol. Cell. Biol.* **14**, 7611-7620.

70. Ortlepp, D., Laggerbauer, B., Mullner, S., Achsel, T., Kirschbaum, B. and Luhrmann, R. (1998). The mammalian homologue of Prp16p is overexpressed in a cell line tolerant to Leflunomide, a new immunoregulatory drug effective against rheumatoid arthritis. *RNA.* **4,** 1007-1018.

71. **Osada, S., Yamamoto, H., Nishihara, T. and Imagwa, M.** (1997). DNA binding specificity of the CCAAT/enhancer-binding protein transcription factor family. *J. Biol. Chem.* **271** , 3891–3896

72. **Pause, A. and Sonenberg, N.** (1992). Mutational analysis of a DEAD box RNA helicase: the mammalian translation initiation factor eIF-4A. *EMBO J*. **11**, 2643-2654.

73. **Pombo, A. and Cook, P. R.** (1996). The localization of sites containing nascent RNA and splicing factors. *Exp Cell Res.* **229**, 201-203.

74. **Proudfoot, N.** Connecting transcription to messenger RNA processing. (2000). *Trends Biochem Sci.* **25**, 290-293.

75. **Reed, R.** (2000). Mechanisms of fidelity in pre-mRNA splicing. *Curr. Opin. Cell Biol.* **12,** 340-345.

76. **Ren, B., Robert, F., Wyrick, J. J., Aparicio, O., Jennings, E. G., Simon, I., Zeitlinger, J., Schreiber, J., Hannett, N., Kanin, E., Volkert, T. L., Wilson, C. J., Bell, S. P. and Young, R. A.** (2000). Genome-wide location and function of DNA binding proteins. *Science.* **290**, 2306-2309.

77. **Rogers, J. H.** (1985). The origin and evolution of retroposons. *Int. Rev. Cytol.* **93**, 187-279.

78. **Saha, R., Banik, U., Bandopadhyay, S., Mandal, N. C., Bhattacharyya, B. and Roy, S.** (1992). An operator-induced conformational change in the C-terminal domain of the lambda repressor. *J. Biol Chem.* **267,** 5862-5867.

79. **Sambrook, J., Fritsch, E. F. & Maniatis, T.** (1989). Molecular Cloning: a Laboratory Manual, 2nd edn. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory

80. **Schiestl, R. H. and Gietz, R. D.** (1989). High efficiency transformation of intact cells using single stranded nucleic acids as a carrier. *Curr. Genet*. **16,** 339–346.

81. Southern, E. M. (1975). Detection of specific sequences among DNA fragments separated by gel electrophoresis, *J Mol Biol.* **98**, 503-17

82. Siebel, C. W., Feng, L., Guthrie, C. and Fu, X. D. (1999). Conservation in budding yeast of a kinase specific for SR splicing factors. *Proc. Natl Acad. Sci. U S A.* **96,** 5440-5445.

83. Smith, C. W., Valcarcel, J. and Carmo-Fonseca, M. (2004). Diversity of vertebrate splicing factor U2AF35: Identification of alternatively spliced U2AF1 mRNAs. *J. Biol. Chem*. (in press).

84. Spector, D. L. (1993). Nuclear organization of pre-mRNA processing. *Curr. Opin. Cell Biol.* **5**, 442-447.

85. Stanchi, F., Bertocco, E., Toppo, S., Dioguardi, R., Simionati, B., Cannata, N., Zimbello, R., Lanfranchi, G. and Valle, G. (2001). Characterization of 16 novel human genes showing high similarity to yeast sequences. *Yeast.* **18,** 69-80.

86. Stoilov, P., Daoud, R., Nayler, O. and Stamm, S. (2004). Human tra2-beta1 autoregulates its protein concentration by influencing alternative splicing of its pre-mRNA. *Hum. Mol. Genet.* **13**, 509-524.

87. Tacke, R. and Manley, J. L. (1999). Determinants of SR protein specificity. *Curr. Opin. Cell Biol.* **11**, 358-362.

88. Tascou, S., Nayernia, K., Samani, A., Schmidtke, J., Vogel, T., Engel, W. and Burfeind, P. (2000). Immortalization of murine male germ cells at a discrete stage of differentiation by a novel directed promoter-based selection strategy. *Biol. Reprod.* **63**, 1555-1561.

89. Tascou, S., Nayernia, K., Uedelhoven, J., Bohm, D., Jalal, R., Ahmed, M., Engel, W. and Burfeind, P. (2001). Isolation and characterization of differentially expressed

genes in invasive and non-invasive immortalized murine male germ cells in vitro. *Int. J. Oncol*. **18,** 567-574.

90. Teigelkamp, S., Mundt, C., Achsel, T., Will, C. L. and Luhrmann, R. (1997). The human U5 snRNP-specific 100-kD protein is an RS domain-containing, putative RNA helicase with significant homology to the yeast splicing factor Prp28p. *RNA*. **3,**1313-1326.

91. Thiry, M. (1995). Behavior of interchromatin granules during the cell cycle. *Eur J Cell Biol.* **68,** 14-24.

92. Trappe, R., Ahmed, M., Glaser, B., Vogel, C., Tascou, S., Burfeind, P. and Engel, W. (2002). Identification and characterization of a novel murine multigene family containing a PHD-finger-like motif. *Biochem. Biophys. Res. Commun*. **293,** 816-826.

93. Trappe, R., Schulze, E., Rzymski, T., Frode, S. and Engel, W. (2002). The Caenorhabditis elegans ortholog of human PHF5a shows a muscle-specific expression domain and is essential for C. elegans morphogenetic development. *Biochem. Biophys. Res. Commun*. **297**, 1049-1057.

94. Tuschl, T., Zamore, P. D., Lehmann, R., Bartel, D. P. and Sharp, P. A. (1999). Targeted mRNA degradation by double-stranded RNA in vitro. *Genes Dev.* **13,** 3191-3197.

95. Wang, Q. and Rymond, B. C. (2003). Rds3p Is Required for Stable U2 snRNP Recruitment to the Splicing Apparatus. *Mol. Cell. Biol*. **23,** 7339–7349.

96. Wang , J., Xiao, S. H. and Manley, J. L. (1998). Genetic analysis of the SR protein ASF/SF2: interchangeability of RS domains and negative control of splicing. *Genes Dev.* **12**, 2222-2233.

97. Weighardt, F., Biamonti, G. and Riva, S. (1996). The roles of heterogeneous nuclear ribonucleoproteins (hnRNP) in RNA metabolism. *Bioessays.* **18,** 747-756.

98. Weinmann, A. S., Yan, P. S., Oberley, M. J., Huang, T. H. and Farnham, P. J. (2002). Isolating human transcription factor targets by coupling chromatin immunoprecipitation and CpG island microarray analysis. *Genes Dev.* **16**, 235-244.

99. Wilkinson, D. G. (1992). In Situ Hybridization: A Practical Approach. Oxford, IRL Press

100. Will, C. L, Urlaub, H., Achsel, T., Gentzel, M., Wilm, M. and Luhrmann, R. (2003). Characterization of novel SF3b and 17S U2 snRNP proteins, including a human Prp5p homologue and an SF3b DEAD-box protein. *EMBO J.* **21**, 4978-4988.

101. Williams, S. C., Cantwell, C.A. and Johnson, P.F. (1991). A family of C/EBP related proteins capable of forming covalently linked leucine zipper dimers *in vitro*. *Genes Dev.* **5**, 1553–1567.

102. Wu, S., Romfo, C. M., Nilsen, T. W. and Green, M. R. Functional recognition of the 3' splice site AG by the splicing factor U2AF35. *Nature.* **402**, 832-835.

103. Vanin, E. F. (1985). Processed pseudogenes - characteristics and evolution. *Ann. Rev. Genet.* **19**, 253-272.

104. Venables, J. P. and Eperon, I. (1999). The roles of RNA-binding proteins in spermatogenesis and male infertility. *Curr. Opin. Genet. Dev*. **9**, 346-354.

105. Vinson, C.R., Sigler, P.B. and McKnight, S.L. (1989). Scissors-grip model for DNA recognition by a family of leucine zipper proteins. *Science.* **246**, 911–916.

106. Xiao, S. H. and Manley, J. L. (1997). Phosphorylation of the ASF/SF2 RS domain affects both protein-protein and protein-RNA interactions and is necessary for splicing. *Genes Dev.* **11**, 334-344.

107. Xiao, S. H. and Manley, J. L. (1998). Phosphorylation-dephosphorylation differentially affects activities of splicing factor ASF/SF2. *EMBO J.* **17**, 6359-6367.

108. Zachar, Z., Chou, T. B. and Bingham, P. M. (1987). Evidence that a regulatory gene autoregulates splicing of its transcript. *EMBO J.* **6,** 4105-4111.

109. Zamore, P. D., Patton, J. G. and Green, M. R. (1992). Cloning and domain structure of the mammalian splicing factor U2AF. *Nature*. **355,** 609-614.

110. Zhang, M., Zamore, P. D., Carmo-Fonseca, M., Lamond, A. I. and Green, M. R. (1992). Cloning and intracellular localization of the U2 small nuclear ribonucleoprotein auxiliary factor small subunit. *Proc. Natl Acad. Sci. U S A*. **89,** 8769-8773.

111. Zhao, J., Hyman, L. and Moore C. (1999). Formation of mRNA 3' ends in eukaryotes: mechanism, regulation, and interrelationships with other steps in mRNA synthesis. *Microbiol. Mol. Biol. Rev.* **63**, 405-445.

112. Zimmermann, S., Steding, G., Emmen, J. M. A., Brinkmann, A. O., Nayernia, K., Holstein, A. F., Engel, W. and Adham, I. M. (1999). Targeted Disruption of the *Insl3* Gene Causes Bilateral Cryptorchidism. *Mol. Endo*. **13,** 681 – 691.

113. Zolotukhin, A. S., Michalowski, D., Bear, J., Smulevitch, S. V., Traish, A. M., Peng, R., Patton, J., Shatsky, I. N. and Felber, B. K. (2003). PSF acts through the human immunodeficiency virus type 1 mRNA instability elements to regulate virus expression. *Mol. Cell. Biol.* **23**, 6618-6630.

114. Zorio, D. A. and Blumenthal, T. (1999). Both subunits of U2AF recognize the 3' splice site in Caenorhabditis elegans. *Nature.* **402**, 835-838.

115. Zuo, P. and Maniatis, T. (1996). The splicing factor U2AF35 mediates critical protein-protein interactions in constitutive and enhancer-dependent splicing. *Genes. Dev*. **10,** 1356-1368.

Publications

Bolcun, E., Rzymski, T., Nayernia, K. and Engel, W. (2003). ADAM family genes testase 2alpha and 2beta are chromosomally linked and simultaneously expressed in male germ cells. *Mol Reprod Dev*.**65**, 19-22.

Trappe, R., Schulze, E., Rzymski, T., Frode, S. and Engel W. (2002). The Caenorhabditis elegans ortholog of human PHF5a shows a muscle-specific expression domain and is essential for C. elegans morphogenetic development. *Biochem Biophys Res Commun*. **297**, 1049-57.

Curriculum vitae

Educational background:

- **2004:** Dissertation for Ph.D Thesis: "Protein interactions, expression studies and functional analysis of PHF5a"
- **2001 to 2004**: Ph.D studies at the Institute of Human Genetics, University of Goettingen, Germany; Project: Functional analyses and protein interactions of PHF5a. Supervisor: Peter Burfeind, Ph.D.
- **2001:** Master degree in Biology, Nicholas Copernicus University. Faculty of Biology and Earth Sciences
- **2000 to 2001**: Experimental studies (Master thesis) at the Institute of Human Genetics, University of Goettingen, Germany; Project: Functional analyses of two testis expressed genes: Adam 25 and Adam27. Supervisor: Wolfgang Engel, Prof. Dr. med.
- **1996 to 2000**: Studies in the Institute of General and Molecular Biology, Nicholas Copernicus University in Torun, Poland; 1998 to 2000 laboratory work in the Laboratory of Genetics: Project: Identification of genes responsible for flowering of Pharbitis nil. Supervisor: Anna Goc, Ph.D
- **1996:** High School Diploma
- **1983-1996**: Primary and Secondary School, Elblag, Poland.

.

Personal Informations:

Name:	Tomasz Marek Rzymski
Date and place of birth:	14.10.1977 Torun, Poland
Nationality:	Polish
Education:	PhD in Developmental and Reproductive Biology University of Goettingen Institut of Human Genetics, November 2004
	Master degree in Biology, Nicholas Copernicus University Faculty of Biology and Earth Sciences, 2001