Universität für Bodenkultur Wien

hnRNP M and its interactors

A Liquid Yeast Two-Hybrid study

Diplomarbeit

zur Erlangung des akademischen Grades Diplom-Ingenieur der Lebensmittel- und Biotechnologie

eingereicht von

Robin Ristl

Durchgeführt am Institut für angewandte Mikrobiologie der Universität für Bodenkultur Wien

Betreuer: Priv.-Doz. Dipl.-Ing. Dr. Johannes Grillari

Wien, Mai 2007

Vorwort

Diese Arbeit bildet den Abschluss meines Studiums der Lebensmittel- und Biotechnologie. Sie beschäftigt sich mit interessanten Details über spezifische Proteininteraktionen im Zellkern, die sich hoffentlich in ein größeres Bild fügen werden, um unser Wissen über die Abläufe des Lebens zu erweitern. Die zentrale Methode dieser Arbeit, Yeast Two-Hybrid, wurde hier erstmals als Liquid Yeast Two-Hybrid durchgeführt, und ich hoffe, auch damit einen kleinen Beitrag zur Weiterentwicklung unserer Wissenschaft geleistet zu haben.

Eine derartige Forschungsarbeit benötigt natürlich viele Ressourcen und Unterstützung von vielen Seiten.

Dank geht daher zu aller erst an meinen Betreuer Priv.-Doz. Giovanni Grillari. Er hat das Thema vorgeschlagen, mich beraten und bei entscheidenden Fragen den Weg gewiesen. Er hat auch, was ich für besonders wichtig halte, genug Freiraum für eine eigenständige Entwicklung der Arbeit gelassen.
Mein Dank gilt auch Marlies Löscher, die mich in das Geheimnis des Yeast Two-Hybrid eingewiesen hat.
Mein Dank gilt ebenso Prof. Hermann Katinger, dessen Aufbau unseres Instituts für angewandte Mikrobiologie schon legendhaft ist.

Acht Monate Laborarbeit waren für diese Arbeit notwendig, und sie wurden in einer großartigen Gruppe verbracht. Besonders hervorheben möchte ich hier Hanna Dellago, die für mich die beste und auf jeden Fall lustigste Laborpartnerin war. Christine Marizzi soll ebenso genannt werden, sie hat ein von mir begonnenes Teilprojekt in ihrer Diplomarbeit vollendet. Allen Gruppenmitgliedern danke ich für die angenehme und freundschaftliche Arbeitsatmosphäre, die ich erfahren habe.
Noch viele Mitglieder der Labors im vierten Stock unseres Instituts sind mir bei großen und kleinen Problemen mit Wissen und Material zur Seite gestanden. Gordana Wozniak-Knopp hat mir entscheidend beim Klonieren und DNA Sequenzieren geholfen, Michael Wuczkowski hat mit mir die mikroskopischen Aufnahmen der Hefe gemacht. All jenen, die mir geholfen haben, danke ich.

Doch nicht nur ein interessantes Thema, umfassende Betreuung und eine gute Forschungsgruppe sind bedeutsam für das gelingen einer Diplomarbeit, sondern auch Unterstützung aus dem persönlichsten Umfeld.

Ganz besonders danke ich daher meiner Freundin Alexandra. Sie hat mich ermutigt, wenn es notwendig war, und meine Arbeit mit Interesse verfolgt. Durch ihre Kochkunst hat sie bei mehreren unserer wöchentlichen LabMeetings ganz ausgezeichnet für unser leibliches Wohl gesorgt und dadurch das Niveau der Verpflegung bei derartigen Treffen dauerhaft gehoben.
Ich danke auch meinen Eltern, die mir immer aufmerksam zugehört haben. An langen Abenden haben sie einiges über Proteine, RNA und vor allem die wundersame Welt der Hefe erfahren.

Wohin die Zukunft führt, weiß niemand. Doch der erfolgreiche Abschluss meiner Diplomarbeit bestärkt meinen Glauben an eine gute Zukunft.

Table of Contents

1 Summary

Heterogeneous nuclear ribonucleoproteins (hnRNPs) are an abundant group of proteins that are mainly located to the nucleus. They bind to mRNA or pre-mRNA and have functions in all aspects of mRNA processing, including transcription, pre-mRNA splicing, polyadenylation, mRNA transport and translation. hnRNP M in particular is associated with pre-mRNA during splicing. It probably has a regulative function in response to cellular stress, causing pre-mRNA splicing to stop.

Pre-mRNA splicing is the process of removing introns from eukaryotic protein-coding RNA transcripts and joining the remaining exons. For a great majority of transcripts splicing is performed by the spliceosome, a large and dynamic machinery that consists of many protein factors and a few catalytic RNA molecules that catalyse the actual splicing reactions. One protein complex physically and functionally associated with the spliceosome is the CDC5L complex. It is composed of the proteins CDC5L, PLRG1, SNEV and SPF27.

The main aim of this work was to identify the interaction domain of hnRNP M with PLRG1 and CDC5L. Deletion mutant coding sequences of hnRNP M were created according to its domain structure. The quality of interactions of the resulting hnRNP M deletion mutants with CDC5L and PLRG1 was examined using the new method of Liquid Yeast Two-Hybrid.

In the Yeast Two-Hybrid system one protein is expressed fused to a DNA binding domain, the other one is expressed fused to a transcriptional activation domain. If the two proteins interact their fusion moieties are brought into proximity and their combined function activates transcription of reporter genes. Transcription of these reporter genes restores the ability of the otherwise auxotrophic yeast strain to synthesise certain nutrients. The ability of the yeast reporter strain to grow on medium lacking the selecting nutrients therefore indicates a protein-protein interaction.
In a standard Yeast Two-Hybrid approach interaction-dependent growth of colonies is observed, but expression of hnRNP M negatively influenced the reporter strain and so inhibited visible colony growth. Liquid Yeast Two-Hybrid was developed as a more sensitive and accurate but still simple method to detect even very weak growth. In this approach the reporter strain is incubated in small amounts of liquid selective medium. Increase in biomass is then easily monitored by consecutive OD measurement. The resulting Liquid Yeast Two-Hybrid data is independent from eye-judgement and can be applied to statistical evaluation.

Using the method of Liquid Yeast Two-Hybrid, the interaction domain of hnRNP M with both CDC5L and PLRG1 was found to be located between amino acids 345 and 583 of the hnRNP M splicing variant 2. This region includes a methionine and arginine rich repeat motif but no RNA binding domains.

Full length hnRNP M has been found to have a detrimental effect on the reporter strain. This effect has been further examined by studying the cellular phenotype of transformed yeast. Yeast that were co-transformed with hnRNP M and PLRG1 were found to be significantly smaller in cell size than control cells. In additional examinations cells co-transformed with hnRNP M and CDC5L were found not to be reduced in size. This implies that hnRNP M may inhibit increase in yeast cell size, though this effect is depending on additional factors.

In addition to studying interactions of hnRNP M a search for novel interaction partners of NOPSI was started. Nopsi is one of several genes deleted in Williams-Beuren Syndrome. In this search a human cDNA library was to be screened by Yeast Two-Hybrid. In this work the Yeast Two-Hybrid reporter strain was transformed with NOPSI and the cDNA library in large scale and then grown on selective medium.

2 Introduction

2.1 Pre-mRNA processing with special emphasis on pre-mRNA splicing

RNA originates upon transcription of DNA. While in prokaryotic cells transcribed RNA is translated directly to form proteins, the protein coding transcripts in eukaryotes require further processing and transport to the cytoplasm.

2.1.1 RNA Polymerase II CTD

In eukaryotes mRNA transcripts are synthesised by RNA polymerase II, while genes encoding for direct RNA products such as ribosomal RNA or transfer RNA are transcribed by RNA polymerases I and III. Among these enzymes only RNA polymerase II has a special, highly functional carboxy-terminal domain (CTD), which is hyperphosphorylated during transcription. The CTD carries or at least mediates several factors that are essential for the pre-mRNA capping, assembly of the splicing complex and the final addition of a poly-A tail. These factors are mediated to the nascent RNA when their target position is synthesised. RNA polymerase II therefore not only produces pre-mRNA but also is essential for its further processing (reviewed by Hirose and Manley, 2000).

2.1.2 End modification of pre-mRNA and mRNA export

Capping of the 5' end

As soon as 25 to 30 nucleotides have been transcribed, the 5' end of the nascent pre-mRNA is capped with 7-methylguanosine. The cap is formed by three enzymes. RNA triphosphatase removes the utmost 5' phosphate. Then RNA guanyl transferase adds GMP 5' to 5' to the pre-mRNA. The guanosine is thus bound over a 5'-5' triphosphate bridge. Finally a methyl group is added to the 7 position of the terminal guanine by guanine-7-methyl transferase. In multicellular eukaryotes the base at the original 5' end is methylated as well. (Lewin, p. 130.) It is bound by a cap-binding complex, which is recognised in mRNA export from the nucleus and initiation of mRNA translation (Alberts, p. 316-317).

Addition of a poly-A tail to the 3' end

The end of a gene contains poly-A signals which are recognised by specific factors immediately after transcription. A hexamer of typically AAUAAA 10-30 nucleotides upstream the cleavage site is bound by the Cleavage and polyadenylation specifity factor (CPSF). A GU or U rich region downstream of the cleavage site is bound by the Cleavage stimulation factor F (CstF).
Both factors are multisubunit proteins that travel on the CTD. Additional factors assemble and cleavage is performed just downstream of a conserved CA sequence. Poly-A polymerase then adds a tail of approximately 200 adenosine nucleotides. Poly-A binding proteins assemble onto the tail and stay there. The help in protein synthesis and may also define the length of the poly-A tail. (Alberts, p. 326-327.)

mRNA export from the nucleus

Export from the nucleus is highly selective. Only mRNA that has been successfully spliced and processed is exported to the cytosol through the nuclear pore complex. Complete mRNA is identified by the set of proteins bound to it. Among others cap-binding proteins are required as well as proteins that mark exon-exon junctions. These proteins are placed on the mRNA after splicing and indicate that splicing has been completed. Certain factors such as snRNPs have to be absent for the export to be allowed. (Alberts, p. 327-329.)

2.1.3 Introns and splicing

Eukaryotic genes typically contain translated regions, termed exons, and longer intervening sequences called introns. Introns strongly vary in size from below 100 to over 30,000 nucleotides. Exons have a more uniform length with an approximate average of 150 nucleotides (Alberts et al. p. 322). RNA splicing is the process of removing introns from the nascent pre-mRNA. One splicing event removes one intron. For most mRNA molecules splicing is catalysed by the spliceosome (Alberts et al., p. 317, 319, 322).

2.1.4 Intron structure

Four positions within an intron are required for spliceosomal splicing. At the 5' splice site introns start with the short but highly conserved sequence GU, at the 3'-splice site they end with AG. They include a branch site which is typically much closer to the 3'-splice site than to the 5'-splice site. The branch site is not well conserved, but most positions show a

prevalence of either purine or pyrimidine nucleotides. In animals its consensus sequence is $Py_{80}NPy_{80}Py_{87}Pu_{75}A_{100}Py_{95}$, in yeast it is well conserved to UACUAAC. The only invariant position is the 3' A, which is directly involved in the splicing reaction. A pyrimidine tract lies between the branch site and the 3' splice site. There, the splicing factor U2AF binds in the first stage of spliceosome assembly (Lewin, p. 689, 691, 694).

Since the conserved sequences that define introns are very short it is extremely difficult to predict splice sites from sequence data alone. In vivo, the rearrangements of the spliceosome that are described below and binding of splicing factors ensure correct splicing (Alberts, p. 321-322).

2.1.5 The splicing reaction

The splicing reaction proceeds over two transesterifications. In the first step the 2'-OH of the invariant adenosine of the branch site performs a nucleophilic attack on the phospho-diester bond of the 5'splice site. The G of the 5' intron end is therefore covalently linked to the branch site A in a 5'-2' bond, resulting in a lariat structure. In the second step the newly created 3'-OH of the 5' exon attacks the 3' splice site. It is linked to the 3' exon and the intron is released. The intron is then debranched and degraded, while the two exons have been joined to form one continuous sequence. (Alberts, p. 317-318)

2.1.6 Catalysis of the reaction by the spliceosome

The reaction is catalysed by small nuclear RNA molecules (snRNAs). Their binding to each other and to the pre-mRNA is performed by base pairing. Five snRNAs are involved in the predominant form of splicing. They are termed U1, U2, U4, U5 and U6, due to their richness in uracil. In higher eukaryotes their length is 100-300 nucleotides, in lower eukaryotes they are longer. They exist as ribonucleoprotein particles (snRNPs). Each snRNP consists of one molecule snRNA and eight common proteins, called Sm-proteins. Only U6 snRNP has a set of slightly different Sm-like proteins. In addition each snRNP contains a set of individual proteins. The overall spliceosome is composed of snRNPs and additional proteins that are generally called splicing factors. A well known and often referred to group of splicing factors are SR proteins, named after the presence of an Arginine-Serine rich region. (Lewin, p. 692-694)

The actual reaction is catalysed when U6 snRNP is bound to the 5' splice site, U2 snRNP is bound to the branch site and both are bound to each other. U2 snRNA pairs with the branch site, but its sequence lacks a binding partner for the invariant A. This nucleotide is therefore

forced in a prominent position and brought into proximity to the 5' splice site. Therefore the first transesterification proceeds easily. The free 3'-OH of the first exon is then brought to the 3' splice site by a rearrangement of U5 snRNP which pairs with both exons. In this way the second transesterification is catalysed. (Alberts, p. 321)

This catalytic structure is not obtained at once though, but reached in a stepwise assembly. The necessary rearrangements require additional proteins and are ATP consuming. The 5' splice site is first bound by U1 snRNP which is later exchanged for U6 snRNP. At the branch site the first factors to assemble are the branch-point binding protein (BBP) and U2AF, later on U2 sRNP binds instead of BBP. This double recognition of the intron defining sequences ensures correct splicing. Errors, such as exon skipping or splicing to cryptic splice sites are prevented. (Alberts, p. 321-322)
The spliceosome is therefore a dynamic machinery that undergoes several changes. Different complexes, referring to the different stages of splicing, can be identified.
The E complex is formed by binding of U1 snRNP at the 5' splice site and U2AF at the pyrimidine tract. When U2 snRNP has bound to the branch site the pre-spliceosomal A complex is formed. Binding of a U5/U4/U6 trimer creates the B1 complex. U1 is then released and U6 pairs with the 5' splice site to form the B2 complex. U4 snRNP is bound to U6 and inhibits the binding between U6 and U2. When U4 is released from U6 in an ATP dependent step, the C1 complex is formed: U6 and U2 pair and catalyse the first reaction. A conformational change in U5 leads to the C2 complex and catalysis of the second reaction. (Lewin, p. 694-696)
In addition an H complex may be defined as the set of proteins that bind pre-mRNA ATP independently and independently from splicing (Jurica et al., 2002).

Assembly of the splicing machinery occurs co-transcriptionally as it is mediated by the CTD, but the actual splicing processes may happen much later (Alberts, p. 323). Usually the order in which introns are removed seems not to be entirely random, but follow a preferred scheme for a given pre-mRNA sequence (Lewin, p.690).

2.1.7 Alternative splicing

At which position splicing occurs is not defined by the pre-mRNA sequence alone. Many genes are found to be spliced in different ways. Alternative splicing creates different versions of proteins from the same primary transcript or simply decides between functional and non-functional forms of a protein. It is found as response to environmental stimulus or to produce

cell type specific forms of a protein and can change during development (Sanford and Bruzik, 1999).
Alternative splicing can occur by removing optional exons or including optional introns. In these cases the combination of given splice sites is altered. Additionally, alternative splicing can be performed using completely alternative splice sites.
Different splice sites may be thought to be competing for detection by the splicing apparatus. In negative regulation a regulatory molecule prevents access to a certain splice site. Often the splicing machinery then detects a usually overlooked weaker splice site. In positive regulation the machinery is directed to otherwise overlooked splice sites. (Alberts, p. 436-437.) Selection of alternative splice sites is therefore regulated by splicing regulatory proteins due to changes in their concentration, activity or intracellular localisation (reviewed by Stamm, 2002).

The possibility to create various products from one primary sequence greatly increases the coding potential of genomes. Recent research points out that probably 42 % or more of human genes are subject to alternative splicing and that alternative splicing is a highly specific way to control protein function. It seems to especially abundant in the immune system and the nervous system (Modrek et al., 2001).
To understand the regulation of pre-mRNA splicing and alternative splicing is therefore of major interest.

2.2 The CDC5L complex

2.2.1 Composition and functions

The CDC5L complex is an evolutionary well conserved protein complex that is located in the nucleus. It consists of at least four core proteins: CDC5L, PLRG1, SNEV, and SPF27 (Ajuh et al., 2000). The complex is associated with the spliceosome. As will be described in more detail below, a direct CDC5L–PLRG1 interaction and the self interaction of SNEV are essential for pre-mRNA splicing.

The CDC5L complex is also required for DNA damage repair. In this context it is linked to WRN, the DNA helicase deficient in Werner's syndrome, over a direct CDC5L-WRN interaction (Zhang N. et al., 2005).

Throughout this chapter it will become evident that the members of the CDC5L complex are involved in multiple processes within the nucleus. To investigate their functions even closer

may therefore greatly improve our understanding of the interconnections of different nuclear pathways.

2.2.2 Identification of CDC5L as the human orthologue of Schizosaccharomyces pombe Cdc5.

In an early screen for cell division cycle mutants in S. pombe the cdc5+ gene was identified in this organism (Nurse et al., 1976). It was then shown that cdc5+ is essential for transition from G2 phase to mitosis and that its gene product shares homology with the DNA binding domain of the Myb family transcription factors (Ohi et al., 1994). Gene products with significant homology to Ccd5 were found in human, mouse, nematode and budding yeast (Bernstein and Coughlin, 1997). The cell division cycle regulatory role was found again for the human form CDC5L, when overexpression of CDC5L in mammalian cells turned out to shorten G2 phase and reduce cell size, while its inactivation slowed cell cycle progression (Bernstein and Coughlin, 1998). The nuclear localisation of CDC5L was confirmed and its gene locus was mapped to chromosome 6 region p21 (Groenen et al., 1998).

2.2.3 Identification of the CDC5L complex and its association with the spliceosome

Again in S. pombe, Cdc5 was found to belong to a multiprotein complex. By immunoaffinity purification the homologues of PLRG1, SNEV, and SPF27 were identified in this complex among other factors. Additionally, S. pombe Cdc5 was shown to associate with U2, U5 and U6 snRNA and to be essential for pre-mRNA splicing (McDonald at al., 1999).

The human orthologue CDC5L was co-purified with the human spliceosome as well (Neubauer et al., 1998) and in a subsequent study the human CDC5L complex was shown to be essential for the second catalytic step of pre-mRNA splicing, but not for the precedent assembly of the pre-spliceosomal complex (Ajuh et al., 2000). These findings are supported by recent research using mass spectrometry. Here the CDC5L complex was found in the B complex prior to its activation but it was not detected in the pre-spliceosomal A complex (Deckert et al., 2006). The role of the CDC5L complex in splicing is depicted in Figure 2.1.

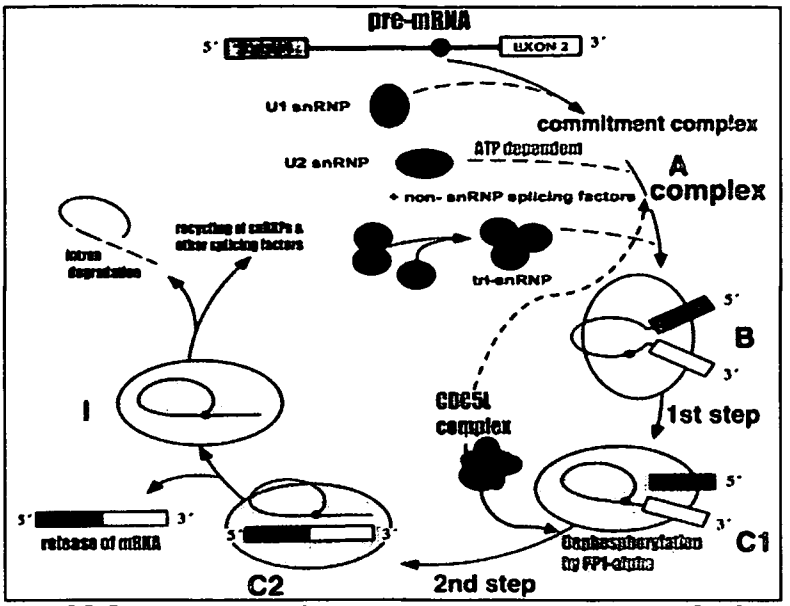


Figure 2.1 The CDC5L complex in pre-mRNA splicing. *The CDC5L complex is associated with the spliceosome prior to spliceosome activation, indicated by a broken lined arrow. It is essential for the second catalytic step of splicing, indicated by a continuous arrow. The figure further illustrates the process of pre-mRNA splicing as described in 2.1.*

(Source: Ajuh et al., 2000)

Independently, the yeast protein Prp19, to which SNEV is the human orthologue, had been previously shown to belong to a spliceosome associated protein complex and here too the complex was found to bind to the spliceosome just before U4 snRNP is released. (Woan-Yuh et al., 1994.). The CDC5L complex is therefore sometimes termed Prp19 complex as well as Prp19/Pso4 complex, Pso4 being a different name for Prp19 that refers to its role in repair of psoralen induced DNA damage (De Andrade et al., 1989) that will be described below.

2.2.4 The CDC5L complex in DNA repair

In a study that aimed to identify factors that are required for DNA interstrand cross link repair all four members of the CDC5L complex and WRN were found to be essential for this DNA repair process (Zhang N. et al., 2005). WRN is a member of the family of RecQ DNA helicases and its deficiency causes Werner's syndrome, a premature aging syndrome. It was further shown that WRN is linked to the CDC5L complex over a specific interaction with CDC5L (Zhang N. et al., 2005). The CDC5L complex is therefore directly involved not only in pre-mRNA splicing but also in DNA repair.

In a very recent study the role of SNEV in response to DNA damage and in association with the CDC5L complex was further examined. SNEV was found to be ubiquitylated and associated with chromatin in HeLA cells after treatment with DNA damaging agents. Additionally, ubiquitylated SNEV had no interaction capability with PLRG1 and CDC5L, thus altering the CDC5L complex in response to DNA damage (Lu and Legerski, 2007). The interaction of ubiquitylated SNEV with SPF27 was not examined. It remains therefore unknown if ubiquitylated SNEV leaves the CDC5L complex or possibly directs it to sites of DNA damage.

2.2.5 PLRG1 and its interaction with CDC5L

PLRG1 has been shown to interact directly with CDC5L. If their interaction is blocked, pre-mRNA splicing is inhibited in in vitro assays. Their direct interaction is therefore essential for pre-mRNA splicing (Ajuh et al., 2001 and 2003).
The interaction domain of PLRG1 to CDC5L is a seven copy WD repeat motif (Ajuh et al., 2001), a structure that is found in various proteins in all eukaryotes and seems to be involved in different functions (reviewed by Temple et al., 1999). The according interaction domain of CDC5L is located at its carboxyl terminus (Ajuh et al., 2001).
The function of PLRG1 in splicing was also found in S. pombe, when mutation of the Plrg1 homologue *Prp5* led to accumulation of unspliced pre-mRNA (Potashkin, 1998). In the same study the mutation also resulted in cell cycle arrest in the G2 phase. If this effect is due to the interaction of Prp5 with Cdc5 has not been examined.

In a frequently cited study concerning PLRG1, its Arabidopsis thaliana orthologue PLR1 has been found to be involved in pleitropic control of glucose and hormone response in this plant and to possibly have a role in regulation of nuclear import (Nemeth, 1998). Although findings made in a plant system might not be simply transferred to heterologous organisms like mammals, PLRG1 turns out to be evolutionary well conserved.

2.2.6 SNEV

SNEV is the human orthologue of the yeast splicing factor Prp19. Its homooligomerisation was found to be essential for spliceosome assembly (Grillari et al., 2005). Like PLRG1 it contains seven copies of the WD repeat motif, which might mediate protein-protein interactions in this case as well.
SNEV also contains a U box domain, a domain that is constitutive for a family of E3 ubiquitin-protein ligases (Hatakeyama et al. 2001). Its E3 ligase activity has not been proven in vitro

so far, but it has been shown to associate directly with the 20S proteasome and might therefore serve as link between the spliceosome and the ubiquitin-proteasome (Löscher et al., 2005).
A most interesting feature of SNEV is its ability to increase the life span of human endothelial cells upon overexpression. SNEV therefore stands for "Senescence Evasion Factor". This effect was found to result from increased stress resistance or improved ability to repair DNA damage (Voglauer et al., 2006).

2.2.7 Interactions within the CDC5L complex

In a related Diploma thesis by Christine Marizzi the interactions of the CDC5L complex proteins with each other were examined using the Yeast Two-Hybrid system. In this work each protein of the CDC5L complex was found to interact with each other (Marizzi, 2007).

2.2.8 CDC5L as transcription factor

CDC5L has been suggested to function as transcription factor. It contains a Myb-like DNA binding domain at its amino terminus (Ohi et al., 1994) and a 12 bp DNA sequence has been identified that is bound by this domain with high specifity (Lei et al., 2000). CDC5L was able to activate transcription of a luciferase reporter gene downstream of this 12 bp sequence (Lei et al., 2000).
This function of CDC5L is yet unlinked to its role in splicing or DNA damage repair, but it is another example for the complex role of the CDC5L complex proteins.

2.3 hnRNP M

2.3.1 Heterogeneous nuclear ribonucleoproteins (hnRNPs)

As was laid out in chapter 2.1 the transcripts of RNA polymerase II are accompanied, altered and directed by protein factors throughout their complete existence. The specific protein composition on an individual mRNA seems to be decisive for its further fate. These proteins that bind RNA and are not part of snRNPs or similar complexes are generally termed heterogeneous nuclear ribonucleoproteins (hnRNPs). hnRNPs are found primarily in the nucleus, but some have the ability to shuttle between nucleus and cytoplasm. Some belong to the most abundant nuclear proteins. At least 20 major hnRNPs have been identified. They typically contain one or more RNA binding domains, domains that mediate protein-protein interactions and domains to determine their localisation. (Reviewed by Dreyfuss et al., 2002.)

The most common RNA binding domains are RNA recognition motifs (RRM), hnRNP K homology domains (KH) and Arginine-Glycine-Glycine (RGG) boxes (Burd and Dreyfuss, 1994). The RNA binding ability of hnRNPs is sequence specific, but other sequences than those of highest affinity might be bound if an hnRNP is present in molar excess over its preferred binding sequence.

hnRNPs participate in all steps of mRNA processing, including transcription, pre-mRNA splicing, polyadanylation, mRNA localisation, export, stability and translation. For example, the concentration of the hnRNP A1 in relation to the concentration of the essential splicing factor SF2 was found to determine selection of alternative 5' splice sites (Mayeda and Krainer, 1992). As another example hnRNP K was shown to accumulate in the cytoplasm and inhibit mRNA translation upon phosphorylation by mitogen-activated protein kinase/extracellular-signal-regulated kinase (MAPK/ERK) (Habelhah et al., 2001).

Many hnRNPs have multiple functions and some are also involved in DNA related events like transcriptional control (e.g. Michelotti et al., 1996), telomere maintenance (e.g. Fiset and Chabot, 2001) and immunoglobulin recombination (Dempsey et al., 1999).

Quite recently an important role for hnRNPs was found in building an exon-exon junction complex as a direct result to pre-mRNA splicing. This marking of exon-exon junctions is especially important for mRNA export and degradation of mRNA that contains premature stop codons using the nonsense-mediated decay pathway (reviewed by Dreyfuss et al., 2002).

2.3.2 Knowledge about hnRNP M

hnRNP M is an authentic component of hnRNP complexes, it is predominantly located in the nucleus and it is an RNA binding protein. It contains three RNA recognition motifs (RRM) and a methionine and arginine rich repeat motif between RRM 2 and RRM 3 (Datar et al., 1993). The domain structure of hnRNP M is depicted in Figure 2.2.

hnRNP M is involved in pre-mRNA splicing (Kafasla et al., 2000). It was detected by antibodies in the pre-spliceosomal A complex and to a lesser extent in the B complex, but not in the prior complexes H and E. Its stable association with pre-mRNA was therefore thought to be confined to the pre-spliceosomal A complex, although its first contact with pre-mRNA was suggested to occur at a very early stage in spliceosome assembly (Kafasla et al., 2002). In a recent work that examined the C complex by mass-spectrometry hnRNP M was found in the H complex and in a splicing intermediates containing C complex (Jurica et al., 2002).

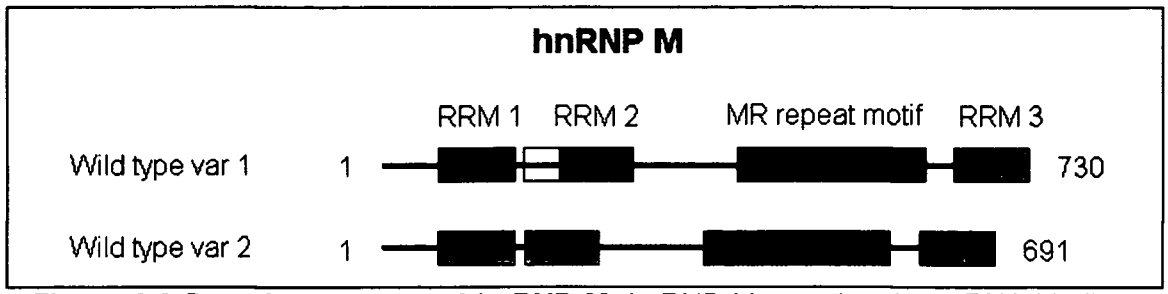


Figure 2.2 Domain structure of hnRNP M. *hnRNP M contains three RNA binding domains of the RRM type (red boxes). A methionine and arginine rich repeat motif (blue box) built from 27 repeats is located between RRM 2 and 3. The sequences of two splicing variants of hnRNP M are known. They result in a 730 and a 691 amino acid polypeptide respectively. The differing region of 39 amino acids (empty square) is located between RRM 1 and 2 and is absent in splicing variant 2.*

An interesting study suggests a role for hnRNP M in splicing regulation in response to heat shock. In HeLa cells hnRNP M was found to temporary leave the hnRNP population and strongly bind to the nuclear matrix after a brief heat shock of 10 minutes at 45 °C. This redistribution of hnRNP M correlated with inhibition of pre-mRNA splicing at the stage of spliceosome assembly (Gattoni et al., 1996). In an additional work hnRNP M was found to accumulate in large distinct structures consisting of merged perichromatin granules upon heat or cadmium induced stress (Chiodi et al., 2000).

Recently hnRNP M was found to be a genuine target for the small ubiquitin like modifier SUMO-2 (Alfred et al., 2004). Sumoylation may affect a proteins function, stability and localisation (Hay, 2001) but the effect of SUMO-2 on hnRNP M is not known.

The *hnRNP M* gene was mapped to band 19p13.3 of the human genome by in situ hybridization (Gattoni et al., 1996).
In different studies hnRNP M proteins of different molecular masses, ranging from 64 to 74 kDa, were found. They are nonetheless highly similar and most probably arise from alternative splicing and posttranslational modifications (Datar et al., 1993, Gattoni et al., 1996, Kafasla et al., 2000). So far the sequences of only two splicing variants are known (Datar et al., 1993, Gattoni et al., 1996). This work was performed with the coding sequence of the shorter splicing variant 2, that lacks 117 bases between RRM 1 and RRM 2 (nucleotide 475 to 591) when compared to the variant 1 coding sequence.

An interaction between hnRNP M and PLRG1 was found by Marco Denegri at the University of Dundee (unpublished results). It is the basis for this work in which the interaction of hnRNP M with both PLRG1 and CDC5L will be further examined.

2.4 NOPSI (WBSCR20A)

Nopsi (WBSCR20A) is one of several genes that are deleted from a specific region on the chromosome band 7q11.23 in Williams-Beuren syndrome (WBS) (Doll and Grzeschik, 2001, Merla et al., 2002). This deletion causes delayed and impaired growth, cardiovascular diseases and psychological limitations (Morris et al., 1988). While Williams-Beuren syndrome patients tend to be overly social and have comparatively good language skills (Bellugi et al., 1999), duplication of the WBS deletion region was found in patients with language delay and even autism (Depienne et al., 2007).
The gene products of this region therefore play an important role in certain functions of the human brain and understanding of their role may provide most interesting knowledge.

At our working group NOPSI was found as interaction partner of the spliceosome associated protein SKIP in a Yeast Two-Hybrid screening (Schörgenhumer, 2005). In the same study a weak Yeast Two-Hybrid interaction of SKIP with SNEV was observed, and NOPSI might therefore have a connection to the CDC5L complex as well.
In this work preparations to find novel interactors of NOPSI will be described.

2.5 The Yeast Two-Hybrid System

2.5.1 Main function of Yeast Two-Hybrid

Yeast Two-Hybrid is a genetic system that was developed to identify protein-protein interactions (Fields and Song, 1989). It makes use of the modular structure of the yeast transcriptional activator Gal4. Gal4 contains a DNA binding domain and an activation domain. The DNA binding domain binds to upstream activation sequences (UAS) and the activation domain is required to activate transcription of the following gene. These two domains have to be in physical proximity to perform their combined function, but they do not have to be on the same protein.
In a Yeast Two-Hybrid assay testing for the interaction of two proteins, one is expressed as fusion protein to the Gal4 DNA binding domain, the other as fused to the Gal4 activation domain. If the two proteins interact the Gal4 domains are brought together and transcription of reporter genes under control of the Gal4 UAS is enabled (Figure 2.3).

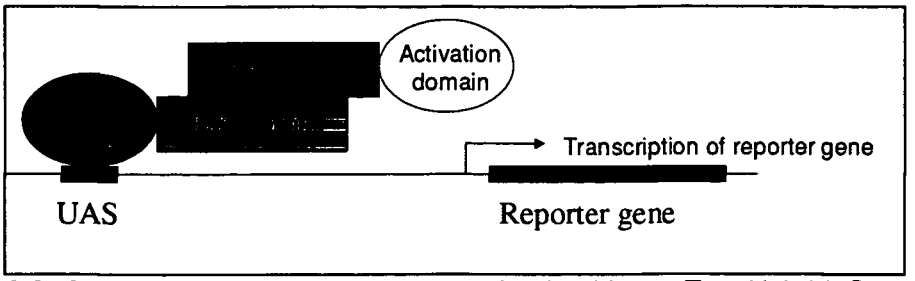


Figure 2.3 Activation of reporter genes. *In the Yeast Two-Hybrid System the examined proteins are expressed as fusion proteins either to the DNA binding domain or to the activation domain of the transcriptional activator Gal4. The protein containing the DNA binding domain is called "bait", while the one fused to the activation domain is the "prey" (also see 2.5.5). Due to its fusion moiety the bait binds to an activation sequence (UAS) upstream of a reporter gene. When a prey protein is able to bind to the bait the activation domain is brought into sufficient proximity to the UAS to activate the transcription of the reporter gene.*

Alternative to the Gal4 domains the DNA binding domain of the E. coli protein LexA and the activation domain of the viral activator VP16 are used in the so called LexA based Yeast Two-Hybrid system (reviewed by Van Criekinge and Rudi Beyaert, 1999).

2.5.2 Yeast Two-Hybrid plasmids

The coding sequences of the proteins of interest are cloned into Yeast Two-Hybrid plasmids. These plasmids contain the sequence for a sufficient part of either the Gal4 DNA binding domain or its activation domain upstream of the multicloning site to generate the fusion proteins. In the same way a nuclear localisation signal is added to the protein. The fusion proteins are expressed from a constitutive promoter in yeast.
The plasmids can be propagated in yeast and in E. coli. In yeast they are selected by nutritional markers, in E. coli by conferring resistance to antibiotics.

2.5.3 Galactose metabolism

Usually Gal4 activates transcription of the six genes the yeast needs to take up galactose and convert it to glucose-6-phosphate (recent findings in Majumdar et al., 2004). In absence of galactose, Gal4 is bound and inhibited by Gal80. When galactose is present an additional factor, Gal3, binds galactose and thus changes its conformation. It then binds Gal80 and removes it from Gal4, which can now activate transcription of the required genes (for recent research see Diep et al., 2006).

2.5.4 Genetic properties of the Yeast Two-Hybrid reporter strain

Yeast Two-Hybrid reporter strains are genetically remodelled. *GAL4* and *GAL80* are non-functional, because the Gal4 function shall be reconstituted from the fusion proteins under examination and this function must not be inhibited by Gal80.

Reporter genes must be set under control of the UAS recognised by Gal4, so they will be transcribed only if the fusion proteins interact.

The plasmids on which the fusion proteins are encoded are selected by nutritional markers. They contain genes that are required for synthesis of certain amino acids and therefore these genes have to be deficient in the reporter strain.

AH109 Genotype:
MATa, trp1-901, leu2-3, 112, ura3-52, his3-200, gal4, gal80, LYS2::GAL1$_{UAS}$-GAL1$_{TATA}$-HIS3, GAL2$_{UAS}$-GAL2$_{TATA}$-ADE2, URA3::MEL1$_{UAS}$-MEL1$_{TATA}$-lacZ, MEL1

Figure 2.4 Genotype of the Yeast Two-Hybrid reporter strain Saccharomyces cerevisiae AH109. *The genes TRP1 and LEU2 are deficient in the reporter strain but are encoded on the Yeast Two-Hybrid plasmids and thus serve as selective markers. The reporter genes HIS3 and ADE2 are required to synthesise histidine and adenine. Their transcription is under control of upstream activation sequences detected by Gal4.*

Figure 2.4 shows the genotype of the reporter strain S. cerevisiae AH109 that was used in this work. The plain reporter strain is auxotroph for four essential nutrients. The genes *TRP1* and *LEU2* that are required to synthesise tryptophane and leucine are mutated and non-functional. These genes are encoded on the Yeast Two-Hybrid plasmids and in this way guarantee that the plasmids are maintained by the reporter strain. The genes *HIS3* and *ADE2* are under control of the UAS that usually belong to GAL1 and GAL2. The reporter strain can therefore synthesise histidine and adenine only if transcription of these genes is activated by interacting fusion proteins. Transcription of these reporter genes is made visible by the ability of the reporter strain to grow on medium lacking both histidine and adenine or at least one of them.

The E. coli *lacZ* gene, coding for β-galactosidase, has been inserted as additional reporter gene. Its activation can be seen by the well known ability of β-galactosidase to hydrolyse X-Gal (5-bromo-4-chloro-3-indolyl-beta-D-galactopyranoside) which results in a blue indigoid dye.

2.5.5 Library screening

The Yeast Two-Hybrid system is not only used to detect interactions between known proteins but also to screen genetic libraries for interaction partners of one known protein. The known protein is fused to the DNA binding domain and is usually referred to as "bait". The library proteins are fused to the activation domain, they are called "prey". The reporter strain is transformed with both plasmids, resulting in yeast cells that contain the bait and at least one prey plasmid. Those cells that contain a prey protein that is able to interact with the bait are able to grow under selective conditions. The colonies they form are then used to obtain the prey plasmid and determine the library sequence it encodes.
In this way a complete library is screened in just one assay.

2.5.6 Advantages and disadvantages of Yeast Two-Hybrid

Overall Yeast Two-Hybrid is a fast an easy to use method to detect and characterise protein-protein interactions but false negative and false positive results occur quite frequently. Most advantages and disadvantages arise from its nature as an in vivo system (reviewed by Sobhanifar, 2003, reviewed by Van Criekinge and Rudi Beyaert, 1999).

Advantages

Yeast Two-Hybrid is an in vivo method. Expression and interaction of the fusion proteins is performed in the yeast host. Therefore the protein-protein interaction is examined in a surrounding that resembles the native environment of most proteins, especially from eukaryotes.
While biochemical methods require proteins and antibodies in sufficient amount and purity, preparation of a Yeast Two-Hybrid assay is performed solely on the DNA level. DNA is stable and easy to multiply and to purify. Truncated proteins can easily be created and studied by using the according DNA sequences.
Novel protein-protein interactions are quickly detected by screening genetic libraries as described above. The interactions in a whole genome can be mapped by producing two sets of genomic libraries, one expressed as bait proteins, the other as prey (Alberts, p. 521).
The strength of an interaction may be semi-quantified by the extent of reporter strain growth at a very rough scale (Estojak et al., 1995).

False negative results

The fusion proteins are expressed constitutively and targeted to the nucleus. In an ideal case they will not do anything more than interact with each other, but it is well possible that they interfere with systems of the yeast metabolism. Even if the fusion proteins interact, the host might be unable to grow due to this interference. Detrimental effects of the fusion proteins on the yeast host are therefore a main reason for false negative Yeast Two-Hybrid results.
False negative results may also occur if a Gal4 domain is in some way shielded by the rest of the protein and so cannot participate in transcriptional activation. In both cases the use of truncated proteins may avert negative effects.
Certainly, proteins form other organisms might be folded or modified incorrectly in the yeast environment, which can disable their interaction ability and lead to a false negative result.

False positive results

Some proteins have the ability to activate transcription of the reporter genes without the missing Gal4 domain, resulting in a false positive signal. The fusion proteins should therefore be tested for this auto-activation ability.
Also, proteins might bind to other proteins in an unspecific way, although they are no interaction partners in their native environment. This interaction then yields a false positive result in the Yeast Two-Hybrid assay.
False positive results are quite frequently obtained using Yeast Two-Hybrid. They are a major disadvantage of this system and so caution should be taken to avoid or identify them. The use of more than one reporter gene reduces the frequency of false positive results (Van Criekinge and Rudi Beyaert, 1999).

Indirect interaction

The Yeast Two-Hybrid system cannot detect if the fusion proteins interact directly or are bridged by an additional factor. If an interaction is direct has to be verified by independent methods.

2.5.7 Methods derived from Yeast Two-Hybrid

The Yeast Two-Hybrid system has become a widely used molecular biology technique and so it has been modified in different ways to complete new tasks using the basic Yeast Two-Hybrid idea.

Reverse Two-Hybrid System

Often it is not only of interest to know which protein binds to which other ones, but also what may cause this interaction to be disturbed. The reverse Two-Hybrid System is used to detect cases in which binding of usual interaction partners is inhibited, for instance by mutations in one protein or by additional compounds.
The usual binding partners are expressed as fusion proteins like in the standard Yeast Two-Hybrid approach, but the reporter gene that they activate upon interacting is lethal to the cell. Only if the interaction is inhibited and the reporter gene is not transcribed, the reporter strain is able to grow. A fusion protein expressed by this strain must therefore contain a mutation or a factor that negatively effects the interaction.
For example, in a recent study the interaction domain of the mitogen-activate protein kinase JNK to an interacting regulatory protein was identified by reverse two-hybrid. A prey library contained randomly mutated sequences of the kinase was screened for those mutations that inhibited binding to the bait protein. The mutated residues had thus replaced those necessary for interaction (Barr et al., 2004).

Yeast Three-Hybrid

In a Three-Hybrid System three interacting proteins are expressed, two of which are fusion proteins as in Yeast Two-Hybrid and one is a linking factor. This system can be applied to screen a library for a protein that mediates interaction between to known proteins (Zhang J. and Lautar, 1996).

Yeast One-Hybrid

Aim of the One-Hybrid variation is to find proteins that bind to a target DNA sequence. The proteins are fused to an activation domain and so activate transcription of a reporter gene downstream their target sequence.

RNA interactions in Yeast Two-Hybrid related systems

Interactions that include RNA can be examined in systems derived from Yeast Two-Hybrid.
In a variation of the Three-Hybrid System a fusion protein is expressed that contains the Gal4 DNA binding domain and a fusion moiety to bind a recombinant RNA over an added specific sequence. The RNA is then bound by an RNA binding protein fused to the Gal4

activation domain. This system may be used to detect RNA binding proteins as well as target RNAs transcribed from a library. (Putz et al., 1996.)

RNA-RNA interactions are detected by the Yeast RNA Hybrid System. One RNA is attached to a fusion protein that contains a DNA binding domain. The other RNA is fused to an RNA-based transcriptional activator. (Piganeau and Schroeder, 2006.)

3 Aims of this work

3.1 hnRNP M interaction domain mapping using Yeast Two-Hybrid

3.1.1 Interaction domain mapping

hnRNP M was shown to be an RNA binding protein with a probably regulative role in pre-mRNA splicing (see 2.3). To further elucidate its function in this essential nuclear process is of great interest. The CDC5L complex, consisting of four main proteins, associates with the spliceosome and is required for splicing (see 2.2).

The major aim of this wok is to identify the interaction domains of hnRNP M to the CDC5L complex proteins CDC5L and PLRG1. Fulfilling this major aim shall contribute to understanding mechanisms and regulation of pre-mRNA processing.

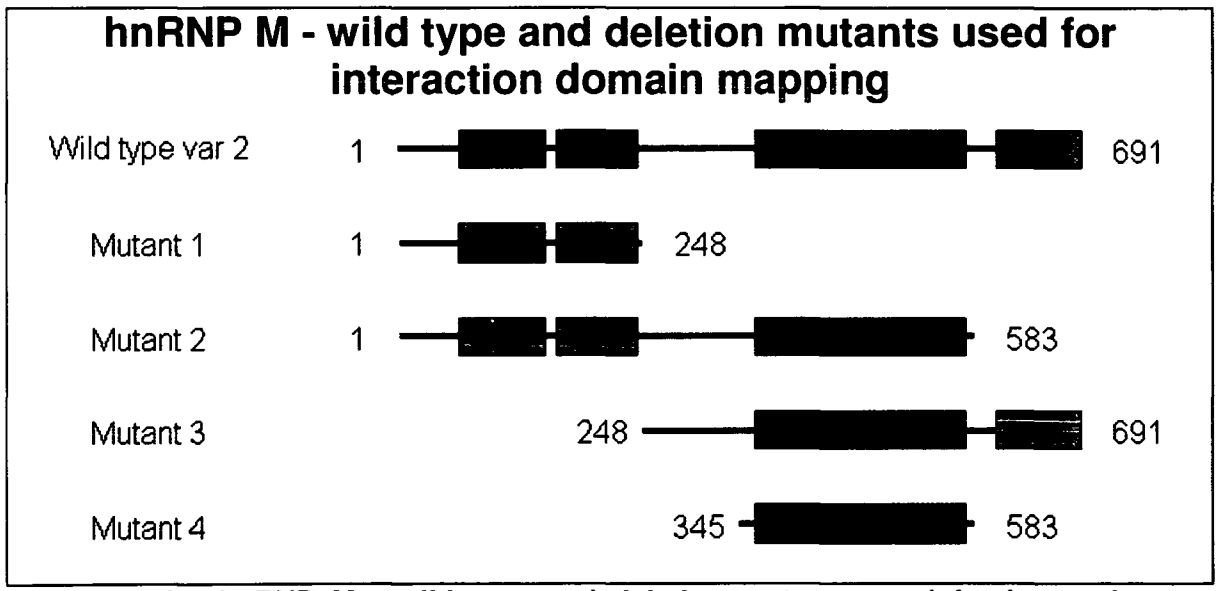


Figure 3.1 hnRNP M - wild type and deletion mutants used for interaction domain mapping. *Aim of this work was to identify the interaction domain of hnRNP M with CDC5L and PLRG1. To do so deletion mutant sequences of hnRNP M had to be created. Using the Yeast Two-Hybrid System it is possible to find out which deletion mutant interacts with CDC5L or PLRG1 and thus contains the sought for interaction domain. Red boxes indicate RNA recognition motifs (RRM). The blue box indicates the methione arginine rich repeat motif (compare 2.3.2).*

To identify the hnRNP M region that contains an interaction domain towards CDC5L or PLRG1, four deletion mutants of hnRNP M were defined as shown in Figure 3.1. By testing which of these mutants interact with CDC5L or PLRG1 a limited hnRNP M region that contains the according interaction domains can be identified.

The method chosen to examine the interaction of the hnRNP M deletion mutants with CDC5L and PLRG1 was the Yeast Two-Hybrid System (described in 2.5). The coding sequences of hnRNP M, its deletion mutants, CDC5L and PLRG1 were to be cloned into Yeast Two-Hybrid vectors.

3.1.2 Modification of the Yeast Two-Hybrid approach

The examined proteins had detrimental effects on the yeast host though, and so it was not possible to show any interaction using a standard Yeast Two-Hybrid protocol.
An additional aim was therefore to modify the Yeast Two-Hybrid protocol to enable detection of interactions despite weak growth of the reporter strain. This modification should be able to detect very weak yeast growth and make quantification of the growth independent from human judgement. Nonetheless the modified method should still be easy to use.

3.2 Phenotype examination of transformed yeast

The yeast growth inhibiting effect of the examined proteins had to be characterised further. Transformed yeast clones used in the interaction domain mapping assay were observed under a light microscope. Aim of this examination was to identify phenotypic traits that occurred due to expression of the fusion proteins and that may be connected to inhibited growth.

3.3 Screening for Nopsi interaction partners

A second aim of this work was to perform the first stage of a Yeast Two-Hybrid screening of a human cDNA library for interaction partners of Nopsi (see 2.4). This included cloning of the Nopsi coding sequence into a Yeast Two-Hybrid vector, co-transformation of the reporter strain on a library scale and plating the transformed yeast on selective plates.

4 Materials and Methods

Buffer and media recipes are listed in Appendix F.

4.1 Cloning of cDNA and insertion into plasmid vectors

4.1.1 DNA Polymerase Chain Reaction – PCR

PCR is a standard molecular biology technique to amplify DNA exponentially in vitro by performing a temperature controlled cycle of DNA denaturing, primer annealing and enzymatic DNA replication. A PCR mixture consists of template DNA, DNA primers, a mixture of all four deoxynucleotidetriphosphates (dNTPs) needed to build DNA, a heat stable DNA Polymerase and a buffer that provides appropriate working conditions for the DNA Polymerase.

PCR primers are synthetic oligonucleotides of usually 18 to 25 nucleotides. Their nucleotide sequence is complementary to the intended starting points of DNA elongation, where they bind and enable the DNA Polymerase to initiate DNA replication. Two different primers are needed in order to allow replication of both DNA strands.

Though the 3' end of a primer has to match the template sequence exactly, the 5' end can contain several non-matching nucleotides. In this way the sequence recognisable by specific restriction enzymes can be added to the 5' end of the original sequence. Restriction cannot be properly performed at the very end of a DNA molecule, therefore one to six nucleotides are added at the 5' end, depending on the actual restriction enzyme in use.
If necessary, nucleotides must be added between the restriction site and the start codon to correctly keep the coding sequence in the plasmid reading frame.

Sequences that allow primers to hybridise with each other must be avoided, for this will prevent successful primer-template annealing. The melting temperature is the temperature at which 50 % of primer molecules form a double helix and 50 % are single stranded. The melting temperature of PCR primers should approximately be between 55 °C and 65 °C. It can be roughly calculated as the sum of 4 °C for each resulting G-C bond and 2 °C for each resulting A-T bond.

A standard PCR mixture had a total volume of 50 µl. Three mixtures were assayed for each amplicon in separate PCR tubes.

Standard PCR mixture

1 µl template
1 µl sense primer (c = 10 pmol/µl)
1 µl anti-sense primer (c = 10 pmol/µl)
1 µl dNTPs (c = 10 nmol/µl for each dNTP)
10 µl 5x Green Go Taq Reaction Buffer (Promega)
1 µl Go Taq DNA Polymerase (Promega) (5 units/µl)
35 µl H_2O

Standard PCR temperature cycle program

1x	95°C 2 min
25x	94°C 30 s
	55°C 30 s
	72°C 1 min/bp
1x	72°C 5 min

Table 4.1 shows which restriction sites were added to which amplicons.

Table 4.1 Restriction sites added to amplicons during PCR

Coding Sequence	5' restriction site	3' restriction site	target vector
hnRNP M (all mutants)	BamH I	Pst I	pGBKT7
PLRG1	BamH I	Xho I	pGADT7
CDC5L	Nde I	Cla I	pGADT7

Plasmids containing the coding sequence of hnRNP M and PLRG1 were kindly provided by Marco Denegri from Dundee.
A Plasmid containing CDC5L was available at our laboratory, prepared by Marlies Löscher.

4.1.2 Gel Electrophoresis

A mixture of DNA fragments of different length or conformation can be separated in a gel made of agarose by applying an electric current that causes the negatively charged DNA to travel towards the anode. The DNA fragments move at a speed inversely related to their size, thus separating into bands.
The DNA is made visible by ethidium bromide staining. Ethidium bromide fluoresces upon excitation with UV light. The size of resulting DNA bands can be determined by comparison

to the bands of a DNA standard that is composed of DNA fragments of known size and concentration. The amount of DNA within a given band can be assessed by comparing the strength of fluorescence of this band to the signal strength of a DNA standard band of similar size and known amount. DNA bands of desired size can be cut out of the agarose gel to retrieve the DNA fragment.

Gel electrophoresis was performed in TAE ethidium bromide buffer and using 1% Agarose Gel slides. A constant voltage of 80 V was applied for approximately 45 min. The band of interest of each sample was cut out and put into a pre-weighed microcentrifuge tube. The weight of the gel slices was determined subsequently.

When used only for determination of the size of DNA fragments a voltage of 115 V was used. The appropriate amount of 6x Orange dye (Fermentas) was added to samples that did not already contain a similar dye.

4.1.3 DNA Purification

DNA is bound to a silica membrane in the presence of the chaotropic salt guanidine isothyocyanate, then washed with 80 % Ethanol Buffer that cannot solve DNA but removes contaminations. The DNA is finally eluted with water.

Purification of DNA amplicons was performed using the Promega Wizard® SV Gel and PCR Clean-Up System according to the manufacturer's instructions, following the centrifugation protocol.
Identical amplicons were unified in one Wizard® SV minicolumn by performing the DNA binding step three times consecutively.
Each sample of purified DNA was eluted in 80 µl nuclease-free water.

4.1.4 DNA restriction digest

Restriction enzymes are endonucleases. They recognise specific DNA sequences and cleave them, unless the DNA is methylated in a host and site specific pattern. Linear DNA is immediately destroyed by exunucleases. In this way restriction enzymes help to destroy foreign DNA and maintain the integrity of the host DNA.

In molecular biological applications restriction enzymes are used to cut double stranded DNA at specific sites not to destroy, but to create DNA fragments that may represent analytic information or be used in further processes.
The most common are type II restriction enzymes that recognise short 4-8 bp palindromic sequences. Additionally many of them make staggered cuts, thus creating short, complementary single stranded ends that are called sticky ends. These sticky ends are well suited to later ligate DNA fragments that have been cut with the same restriction enzyme.
Linearised vectors may additionally be treated with a 5' Phosphatase to prevent religation of the vector without an insert.

All restriction enzymes and according buffers were purchased from New England Biolabs.
The concentration of the used plasmid vector solutions was 0.89 µg/µl for pGBKT7 and 3.65 µg/µl for pGADT7.

A typical mixture had a volume of 50 µl, containing 5 µg plasmid DNA or 39 µl insert DNA PCR product, 2 µl of each restriction enzyme, 5 µl of the appropriate 10x Buffer, 0.1µg/µl BSA (100x = 10 mg/ml) if required and the necessary amount of RO-H_2O.
Complete mixing was ensured by carefully inverting the reaction tube a few times and briefly spinning down the solution eventually.
Incubation lasted three hours at 37 °C for all assays. To the cut vectors 2.5 µl Calf Intestinal Alkaline Phosphatase (CIP, c = 10,000 units/ml) were added to reach 0.5 units/µg DNA and incubation was continued for one hour at 37 °C. Finally the processed DNA was purified as described in 4.1.3, eluted in 50 µl nuclease free H_2O and stored at -20 °C.

Table 4.2 shows which NE Buffer was used for which combination of restriction enzymes and the amount of BSA added. Table 4.2 shows the sequences of the used restriction sites. It also lists to which coding sequence each site was added during PCR and which Yeast Two-Hybrid vector was cut accordingly.

Table 4.2 Buffers and amount of BSA used for specific double restriction digests

Restriction enzymes	NE Buffer	BSA added
BamH I, Pst I	NEB BamH I	5 µg
BamH I, Xho I	NEB BamH I	5 µg
Nde I, Cla I	NEB 4	5 µg

Table 4.3 Sequences of restriction sites and application to genetic constructs

Restriction Site	Sequence	Concentration	Added to cds	Used in vector
BamH I	5'...G▼GATCC...3' 3'...CCTAGG▲...5'	20,000 units/ml	hnRNP M, PLRG1	pGBKT7, pGADT7
Pst I	5'...CTGCAG▼...3' 3'...G▲ACGTC...5'	20,000 units/ml	hnRNP M	pGBKT7
Xho I	5'...C▼TCGAG...3' 3'...GTGCAC▲...5'	20,000 units/ml	PLRG1	pGADT7
Nde I	5'...CA▼TATG...3' 3'...GTAT▲AG...5'	20,000 units/ml	CDC5L	pGADT7
Cla I	5'...AT▼CGAT...3' 3'...TAGC▲TA...5'	5,000 units/ml	CDC5L	pGADT7

For all used restriction sites name, sequence and aliquot concentration are listed, as well as in which coding sequences and vectors they were applied. Triangular arrows indicate the exact point of cleavage.

pGBKT7 was also cut sequentially with BamH I and Pst I using the following mixture and conditions:

4 µl pGBKT7 (c = 0, 89 µg/µl)
6 µl BamH I Buffer
6 µl 10x BSA
4 µl BamH I
40 µl H_2O

The mixture was incubated for 3 hours at 37 °C.

The cut vector was purified as described in 4.1.3 and was subsequently cut with Pst I using the following mixture and conditions:

44 µl eluated DNA
6 µl NE Buffer 3
6 µl 10x BSA
4 µl Pst I

Incubation lasted for 3 hours at 37 °C.

Then 2.5 µl Calf Intestinal Alkaline Phosphatase (CIP)(10,000 units/ml) were added, to prevent religation of the vector without an insert. The mixture was incubated for 1 hour at 37 °C. Finally the cut vector was purified again, eluted in 50 µl nuclease free H_2O and stored at –20 °C.

4.1.5 DNA Ligation

Complementary sticky ends of linear otherwise double stranded DNA fragments easily base-pair under physiological conditions. A DNA Ligase is used to catalyse the formation of the phosphodiester bonds between neighbouring 5' phosphate and 3' hydroxyl termini. This step is required to covalently join the two ends. When finally the ends of an insert are joined with the two complementary ends of the vector the circular and thus functional form of the plasmid is restored and it can be used for transformation.
Certainly, blunt ends are ligated too, but the original juxtaposition of two DNA fragments is less stable and the orientation of the insert cannot be defined if only blunt ends are used.

T4 DNA Ligase (400,000 units/ml) and the according reaction buffer were purchased from New England Biolabs.

A typical ligation assay had a total volume of 15 µl containing approximately 0.1 µg vector DNA and insert DNA in 5 to 10 fold molar excess over vector DNA, 1 µl T4 DNA Ligase, 1.5 µl 10x T4 DNA Ligase reaction buffer and the necessary amount of RO-H_2O.
The actual DNA solution volumes to be used were estimated from the result of analytic gel electrophoresis (see 4.1.2).
Ligation was performed over night at 16 °C. The DNA Ligase was heat inactivated for 20 minutes at 65 °C afterwards.

4.2 Transformation of Escherichia coli

4.2.1 Generation of electro-competent cells

To be susceptible to transformation by electroporation E. coli has to be made electro-competent.

20 ml LB-medium were inoculated with E. coli TOP10 and incubated on the shaker at 37 °C over night.
10 ml over night culture were then used to inoculate 1 l LB-medium. Incubation was continued under the same conditions until an OD 600 of 0.5 was reached, taking approximately 3 hours. The suspension was put on ice for 10 min then centrifugated at 3500-4000 rpm at 4 °C for 15 min. The supernatant was removed and the cell pellet was resuspended in 200 ml ice-cold 1 mM Hepes and put on ice for 10 min. Centrifugation, resuspension and incubation on ice were repeated. Then centrifugation was done as before and the supernatant was discarded again. 100 ml ice-cold 10 % glycerol were added while gently swirling the cell suspension. Centrifugation was done again as before, the supernatant was discarded. The cell pellet was resuspended in 2 ml ice-cold 10 % Glycerol and transferred to a pre-chilled 2 ml micro-centrifuge tube. 100 µl aliquots were distributed into pre-chilled micro-centrifuge tubes. The aliquots were subsequently equilibrated in liquid nitrogen, then stored at –80 °C.

4.2.2 Electroporation

Electroporation is an effective method to introduce foreign DNA into host cells. The mixture of DNA and host cells is exposed to a short lasting electric field that causes the plasma membranes to become porous at a level that allows macromolecules like DNA to cross it. Migration of DNA is enhanced during the pulse due to its negative charge. Cells have to be treated though before electroporation to become electrocompetent.

The resistance of the cell suspension is crucial to the success of electroporation. It is indicated by the time constant. This is the period of time between ending the pulse and reduction of the voltage to 37 (100 / e) % of its original value. Values above 20 ms indicate a resistance too high for electroporation to occur, while too low resistance will prompt an explosion like release of the applied energy, which results in a high rate of lethal cell damage.
Electric voltage, capacity and resistance have to be set according to the host organism in use.

100 µl aliquots of competent E. coli were thawed on ice, 5 µl over night ligation solution were added.
Electroporation was performed in standard plastic cuvettes. Their flanks were covered with metallic foil inside and outside to serve as electrodes. The gap between the covered sides was 2 mm.

The parameter settings for E. coli are 2.5 kV, 25 µF and 1000 Ohm.
Time constants usually ranged from 16.0 to 19.3.
Immediately after electroporation the cell suspension was transferred into 900 µl SOC medium and incubated at 37 °C for 30 minutes. The complete suspension was then plated on a 9 mm LB-Agar plate containing an antibiotic to select for successfully transformed cells. Kanamycin was used to select for pGBKT7 and ampicillin to select for pGADT7 carrying cells. Incubation was performed over night at 37 °C. The plates were sealed and stored up to one week at 4 °C.

4.3 –80 °C stocks of transformed E. coli

Stocks of E. coli containing plasmids of interest were made from liquid cultures that had been incubated overnight at 28 °C. 200 µl over night culture and 100 µl sterile 87 % glycerol were mixed in a micro-centrifuge tube and immediately frozen at –80 °C.

4.4 Plasmid preparation from E. coli

4.4.1 Cultivation of E. coli

For small scale plasmid preparation 5 ml selective LB medium were inoculated with half a colony and grown over night on the shaker at 28 °C. For medium and large scale plasmid preparation 100 ml selective LB medium were inoculated with a single colony or with cell mass scraped from a –80 °C stock. Incubation was done at 28 °C on the shaker.
The comparatively low incubation temperature of 28 °C was chosen to increase plasmid stability during propagation.

4.4.2 Plasmid preparation

Bacterial cells are removed from their media by centrifugation and resuspended in a buffered solution. The resuspension solution contains RNAse to remove RNA that could not be easily separated from the plasmid DNA of interest otherwise.
Cell lysis is achieved by adding a lysis solution containing NaOH and SDS. SDS dissolves the plasma membrane and denaturates proteins, NaOH denaturates both plasmid and chromosomal DNA. In the subsequent neutralisation step plasmid DNA is renaturated while chromosomal DNA is precipitated, SDS is precipitated as potassium dodecylsulphate, proteins and membrane debris is co-precipitated.

The solution is cleared from the precipitate either be centrifugation or by filtering. The plasmid DNA is then purified by a membrane based anion exchange process as described in 4.1.3 and by isopropanol and ethanol precipitation. The DNA may finally be resolved in RO-H_2O or TE buffer.

Small scale plasmid preparation was performed using the Promega Wizard® *Plus* SV Minipreps DNA Purification System according to the manufacturer's instructions, following the centrifugation protocol.
Each sample of purified plasmid DNA was eluted in 50 µl nuclease-free RO-H_2O.

Medium and large scale plasmid preparation was performed using the Qiagen Plasmid Midi/Maxi kit according to the manufacturer's instructions.
As only exception from this protocol, finally isopropanol precipitated DNA was centrifuged at 2300 x g for 30 minutes, the following centrifugation of ethanol precipitated DNA was performed at 2300 x g for 20 minutes. The supernatant was carefully removed by decanting, the tube was then kept horizontally to ensure effective air drying of the DNA precipitate with a maximum duration of 10 minutes. Each sample of purified plasmid DNA was resuspended in 100 µl TE buffer with pH 8.0.

The approximate concentration of plasmid solutions was determined photometrically at a wavelength of 260 nm, assuming that the concentration of a DNA solution in µg/ml is 50 times the OD 260 value of the sample within an OD range of 0.1 to 1.

4.5 Identification of E. coli clones positive for vector and insert of interest

4.5.1 PCR screening

A PCR mixture is set up to amplify the inserted region of the vector, using vector primers - primers that bind to the vector closely up- and downstream of the multi-cloning site, thus enclosing the inserted region. The DNA included in the cells that are added to the mixture serves as template. If it does not contain the vector at all, the result of PCR is similar to a negative control, if it contains a vector the size of the amplicon indicates whether the vector contains an insert and if it is of correct size.

For each sample a mixture of a total volume of 30 µl was set up.

PCR screening mixture

1 µl T7 primer (c = 10 pmol/µl)
1 µl 3'AD primer for pGADT7, 3'DNA-BD primer for pGBKT7 (c = 10 pmol/µl)
0.5 µl dNTPs (c = 10 nmol/µl for each dNTP)
6 µl 5x Green Go Taq Reaction Buffer (Promega)
0.3 µl Go Taq DNA Polymerase (Promega) (5 units/µl)
21.2 µl H_2O

According PCR program

1x	95°C 3 min
30x	94°C 30 s
	55°C 30 s
	72°C 1 min/bp
1x	72°C 5 min

The initial time for denaturing is increased compare to the standard PCR program (see 4.1.1) because this step also includes cell lysis here.
When screening for CDC5L the time for denaturing was increased to 45 s to account for the larger size of this coding sequence.

The result of PCR was made visible by gel electrophoresis as described in 4.1.2.

4.5.2 Control digest

In a control digest a plasmid is enzymatically cut at least at two sites. The lengths of the resulting restriction fragments indicate if the plasmid is composed in the way originally intended.

Control digests of purified plasmids were performed using the same restriction enzymes as originally used in the vector and insert DNA digest preceding ligation (compare 4.1.4). A typical mixture had a total volume of 10 µl, containing approximately 1 µg DNA, 0.5 µl of each restriction enzyme, 0.1µg/µl BSA and the necessary amount of H_2O. Incubation at 37 °C lasted for 2 hours.
The size of the resulting DNA fragment was determined by gel electrophoresis.

4.6 Determination of DNA sequences

4.6.1 DNA Dye Terminator Sequencing

One strand of the DNA to be sequenced is amplified by PCR, using just one primer. The PCR mixture contains differently fluorescent labelled 2'-3'-di-deoxyribonucleotides (ddNTPs) in addition to the usual dNTPs. Whenever a ddNTP is added during elongation chain termination occurs, because without the 3' hydroxyl group no further elongation of this chain is possible.
A resulting chain has two properties that carry exact information about position and type of its 3' terminus nucleotide in respect to the sample sequence.
The first information a resulting chain bears is the position of the added ddNTP. All resulting chains have a common starting point and therefore their length measured in nucleotides is the position of the last nucleotide relative to the common starting position.
The second information it bears is the type of the ddNTP. Due to the different labels, the four types of nucleotides can be distinguished.
To retrieve these information, the DNA fragments are denaturated and separated by capillary electrophoresis. Similar to gel electrophoresis described in 4.1.2 the fragments are separated by size, but here a resolution of just one nucleotide is achieved. Thus the fragments are ordered by size, and this means their 3' termini resemble the original DNA sequence.
The DNA finally travels through a detection cell. There, Laser light illuminates the passing DNA causing the fluorescent labels to emit light at their specific wavelengths. The emitted light is separated by wavelength and detected by a charge-coupled device (CCD).
The result of this process is an electropherogram, a sequence of emission peaks of four different colours, each representing one type of DNA nucleotide. The electropherogram is usually interpreted into DNA code by specific software. Interpretation mistakes that may for example occur due to improperly separated signals can be manually corrected.

Today, approximately 1000 bp of DNA can be sequenced in one run.
Limiting factors to the length of a DNA sequence to be determined at once are PCR, the size of the capillary and a statistical effect: At which position a ddNTP is added to a growing chain is a stochastic process. Certainly, the chance for a certain chain length to occur decreases with length, thus the resulting signals get continually weaker for nucleotide positions more remote from the start.

The sequencing machine used for this work was an Applied Biosystems/Hitachi ABI Prism® 3100-Avant Genetic Analyzer. Usually it was possible to obtain electropherograms of slightly

more than 500 unambiguously readable nucleotides. Therefore sequencing primers were designed to divide all inserts above this size into stretches of maximally 500 bp. The fragment sequences were then joined using the software SeqMan (DNASTAR, Inc., USA).

4.6.2 Sample DNA amplification and purification

Sequencing is negatively influenced by impurities and by low amounts of template DNA. Therefore the coding sequences were amplified by PCR. Vector primers were used like for PCR screening, because the primer binding site itself is not sequenced, since no ddNTPs can be included there. Also sequencing often does not work properly for the first few nucleotides. In addition, just one set of vector primers is needed for each vector, regardless of the insert sequence.

For each sample a mixture of a total volume of 50 µl was set up.

PCR mixture to amplify DNA before sequencing

1 µl template (low scale plasmid preparation)
1 µl T7 primer (c = 10 pmol/µl)
1 µl 3'AD primer for pGADT7, 3'DNA-BD primer for pGBKT7 (c = 10 pmol/µl)
1 µl dNTPs (c = 10 nmol/µl for each dNTP)
10 µl 5x Colourless Go Taq Reaction Buffer (Promega)
1 µl Go Taq DNA Polymerase (Promega) (5 units/µl)
35 µl H_2O

According PCR program

1x	95°C 2 min
30x	94°C 30 s
	55°C 30 s
	72°C 1 min/bp
1x	72°C 5 min

Resulting amplicons were purified as described in 4.1.3, eluted in 50 µl nuclease free H_2O. 2 µl of the eluate were applied to analytical gel electrophoresis to verify the sample purity.

4.6.3 Sequencing PCR

A BigDye™ Terminator v3.0 Ready Reaction Mix (Applied Biosystems, USA) was used. It contains DNA polymerase, dNTPs and labelled ddNTPs.
Sequencing PCR mixtures were set up to a total volume of 10 µl.

Sequencing PCR mixture

4 µl purified DNA
4 µl BigDye™ Terminator v3.0 Ready Reaction Mix
1.2 µl 10x BigDye™ v3.0 Dilution Buffer
0.8 µl primer (c = 10 pmol/µl)

Sequencing PCR program

30x	95°C 30 s
	50°C 15 s
	60°C 4 min
1x	20°C 1 min

4.6.4 Ethanol Precipitation

The labelled amplicons were then purified by ethanol precipitation: Each sample was transferred to a microcentrifuge tube. 8 µl H_2O and 32 µl 95 % ethanol were added and the mixture was vortexed. The sample was kept for 15 minutes at room temperature and was then centrifuged for 20 minutes at room temperature. The supernatant was removed without disturbing the precipitate. In a washing step 125µl 70 % ethanol were added, the sample was vortexed and centrifuged for 10 minutes at room temperature. The supernatant was removed again. The sample was vacuum dried in a heated vacuum drier for 15 minutes. Storage was possible up to one week at 4 °C.
During incubation at room temperature and for eventual storage the samples were covered to avoid light induced degradation of the fluorescent labels.

4.6.5 Sample preparation for sequencing

20 µl formamide were added to each sample. The samples were kept at 95 °C on a heat block for 2 minutes to denature the DNA. 20 µl of each sample were then loaded, using a 96 well sequencing plate. Air bubbles must be avoided.

NB: Four samples are processed simultaneously, therefore the number of wells used must be a multiple of four. Wells that will be required by the machine but do not contain samples must be filled with formamide only.

4.6.6 Sequencer settings

The sequencing software has to be set according to the machine itself, the applied dye terminator and to the capillary.
The ABI Prism® 3100-Avant Genetic Analyzer required the Dye Set Z and the mobility file "DT3100POP4(BDV3)V1.mob". For the long capillary used in this work the according run module was "long seq 80 POP4 Default Module", the analysis module was "BC-3100 POP4 80cm seq".

4.7 Transformation of the Yeast Two-Hybrid reporter strain AH109

4.7.1 Yeast cultivation

A BD Yeastmaker™ Kit was purchased from BD Biosciences Clontech, including a stock of the Saccharomyces cerevisiae strain AH109.
The yeast strain AH109 was streaked on YPDA Agar and incubated at 30 °C for 3 days, then sealed and stored at 4 °C. 3 ml YPDA were inoculated with a single colony and incubated at 28 °C on the shaker over night. The 3 ml suspension was transferred into 50 ml YPDA. Incubation was continued until an OD 600 of 0.4 – 0.5 was reached. This incubation step usually took 3 hours, OD 600 was measured regularly.

4.7.2 Generation of competent yeast for PEG/Lithium Acetate mediated transformation

The 50 ml suspension was split into two 50 ml centrifuge tubes and centrifuged for 5 minutes at 700 g. The supernatants were poured off and each pellet was resuspended in 15 ml sterile RO-H_2O. Centrifugation was done as before and the supernatants were poured off again. Each pellet was resuspended in 0.75 ml 1.1x TE/LiAc solution. Each resulting suspension was transferred into a microcentrifuge tube and centrifuged at high speed, 13400 rpm, in an Eppendorf MiniSpin centrifuge with a setting of 15 seconds. Note that the maximum speed was kept for a few seconds only while the longer part was required for acceleration. The

supernatants were immediately carefully removed without disturbing the pellets. Each pellet was resuspended in 0.3 ml 1.1x TE/LiAc solution. Storage up to a few hours at room temperature was possible before transformation.

4.7.3 Transformation

Herring testes DNA (c = 10 mg/ml) as carrier DNA was denaturated by heating at 100 °C for 5 minutes, then chilling on ice immediately and repeating this procedure once. For co-transformation of the reporter strain with two plasmids 2.5 µg of each plasmid and 5 µl denaturated carrier DNA were mixed in a microcentrifuge tube. 50 µl competent yeast suspension were added and gently mixed by briefly vortexing twice. 0.5 ml PEG/LiAc solution were added and thoroughly mixed by vortexing three times. Incubation at 30 °C for 30 minutes followed immediately. Additional mixing was performed after 15 minutes of the incubation period. The tube was kept at exactly 42 °C in a waterbath for 15 minutes. After five and ten minutes the suspension was quickly vortexed. The transformed yeast was subsequently centrifuged at high speed for 15 seconds as described above. The supernatant was carefully removed and the pellet was resuspended in 400 µl 0.9 % NaCl solution.

4.7.4 Plating and incubation

100 µl of each suspension were spread on a Synthetic 2xDO (-Leu, -Trp) plates using a sterile Trigalski spatula. In order not to harm the yeast cells, a minimum number of strokes necessary to achieve even distribution was performed.

After the fluid had been absorbed by the medium the plates were turned bottom up. To save the plates from drying a sterile filter paper was put in each cover and wetted with 2 ml H_2O. During incubation water was regularly added to keep the filter papers, and thus the plates, moist.

The plates were incubated at 30 °C until colonies arose. Usually 3 to 6 days of incubation were required.

4.8 Liquid Yeast-Two-Hybrid

Yeast growth is the parameter that indicates protein-protein interactions in the Yeast Two-Hybrid System. In Liquid Yeast Two-Hybrid the co-transformed reporter strain is transferred to liquid 4xDO (-Leu, -Trp, -His, -Ade) medium. Growth of the reporter strain is then easily monitored by consecutive OD measurment.

4.8.1 Preparation of the 96-well microtiter plates and inoculation

A 96-well microtiter plate was filled with 220 µl 4xDO medium per well. A second plate was filled with 180 µl 4xDO medium per well. Each well was inoculated with a comparatively small part of a separate colony grown on 2xDO agar. Twelve individual colonies of one co-transformation plate were chosen to complete one row in the microtiter plate. The cells were suspended carefully.

20 µl of each well were then transferred to the corresponding wells on the second plate, using a multi-pipette, to create a 1:10 dilution. This step was included to ensure that the observed growth behaviour was not depending on the inoculum concentration. The second plate was then treated like the original one.

Note that the initial OD 600 values should be above 0.1 to allow for correct relation to later values.

4.8.2 OD measurement

OD 600 was measured every 24 hours approximately. The initial measurement was performed immediately after inoculation.

A Sunrise Tecan photometer suitable for 96-well microtiter plates was used. Measurement was performed after 30 seconds of high speed shaking the plate, at 600 nm at normal accuracy. The plate cover was removed for OD measurement. The results were saved to Microsoft Excel sheets.

4.8.3 Incubation

The plates were incubated at 28 °C on a Heidolph Instruments Titramax 101 microtiter plate shaker at 500 rpm. Incubation lasted for three to five days.

4.8.4 Evaluation

Biomass increase within three days was used to assess an interaction. Increase between 100 % and 200 % accounted for weak interactions and increase above 200 % was defined to indicate strong interactions.

The biomass increase of a twelve well sample was determined using a linear homogeneous regression model. According to this model the OD 600 values after three days of incubation (Y) are related to the initial OD 600 values (X) as $Y = b * X$. The factor b indicates the mean relative OD 600 increase. The standard error of b was calculated to account for stochastic deviations.

Required formulas

$$Y = b * X + E \quad (1)$$

$$\hat{b} = \frac{\sum_{i=1}^{n} x_i y_i}{\sum_{i=1}^{n} x_i^2} \quad (2)$$

$$SE_{\hat{b}} = \sqrt{\frac{\sum_{i=1}^{n} (y_i - \hat{b}x_i)^2}{(n-1)\sum_{i=1}^{n} x_i^2}} \quad (3)$$

As Liquid Yeast Two-Hybrid was developed during this work its evaluation is described in detail in the Results chapter 5.2.

Growth curves were drawn by depicting the mean OD 600 values of a sample versus the measurement time.

4.8.5 Confirmation on 4xDO plates

To confirm observed growth on solid medium, cell suspensions were transferred to 4xDO plates with a single streak. The plates were incubated as described in 4.7.4. The formation or absence of colonies was taken as conformation for positive or negative results respectively.

4.9 Yeast cell size measurement

4.9.1 Measurement

20 µl cell suspension from a Liquid Yeast Two-Hybrid well were applied to microscopy after the Liquid Yeast Two-Hybrid experiment had been finished. Pictures were taken at 100 x magnification using an Olympus BX 51 microscope with an attached camera. The size of single cells was measured according to the following rules: The longest possible way across the cell as it was depicted was measured. Buds were not included in the examination. If cells had not separated properly after division all cells that were of approximately the same size as their mother cell were included though. The minimum number of single cells to be examined per sample was 25.
Microscopy samples were taken from wells that showed yeast growth values most closely to the average of their twelve-well samples.

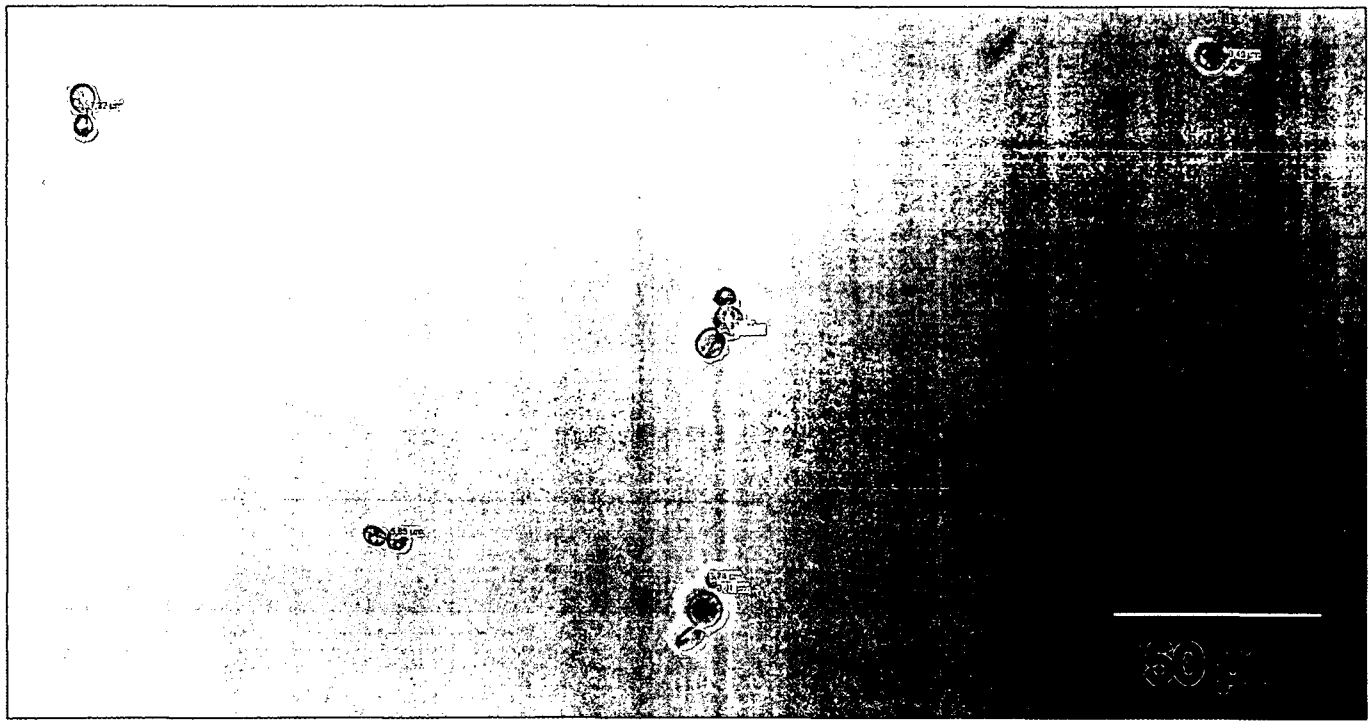


Figure 4.1 Yeast cell size measurement. *This picture shows yeast AH109 cells at 100 x magnification. These cells have been transformed with CDC5L-pGADT7 and hnRNP M-pGBKT7 and have then been incubated for six days in 4xDO medium. The size of each cell was measured as the longest possible line across the cell as marked by coloured lines.*

4.9.2 Evaluation

Cell size distributions were depicted by box-and-whisker plots. The mean size of each co-transformation type was compared to the mean size of each other type by Student's T-Test.

4.10 Yeast-Two-Hybrid library scale transformation

4.10.1 Library scale creation of competent yeast AH109

Three glass tubes containing 3 ml YPDA each were inoculated with AH109 and incubated overnight at 28 °C on the shaker at 180 rpm.

150 ml YPDA were inoculated using two 3 ml cultures (one remaining as reserve). The suspension was incubated as before for approximately 24 hours.

15 ml of the 150 ml culture were used to inoculate 400 ml YPDA to an OD 600 value of 0.2 to 0.3. To reach an OD 600 of 0.5 +/- 0.1 the culture was incubated as before for four hours. OD 600 was measured every hour.

The culture was split into two centrifuge tubes and centrifuged at room temperature for 5 minutes at 700 g. The supernatants were decanted and each pellet was resuspended in 100 ml RO-H_2O. Centrifugation was repeated and the supernatants were decanted again. Each pellet was resuspended in 2 ml 1.1x TE/LiAc

4.10.2 Library scale transformation of AH109

Transformation Mixture

500 µl hAorta Library in pACT2 (c=2.15 µg/µl)
100 µl NOPSI in pGBKT7 (c=1.57 µg/µl)
1 ml denaturated Herring testes DNA (c=10 mg/ml)
4 ml competent AH109
27.4 ml PEG/LiAc

The mixture was set up in a 250 ml flask, vortexed and incubated for 45 minutes at 28 °C on the shaker.
Then 7 ml DMSO were added and mixed by swirling.
Heat shock was performed subsequently in a water bath at 42 °C for 20 minutes with mixing after 7 and 14 minutes. Centrifugation at 700g for 5 minutes at room temperature followed. The supernatant was removed and the pellet was resuspended in 35 ml YPDA Plus (BD Biosciences Clontech, USA) and incubated for 90 minutes at 28 °C on the shaker. Centrifugation was performed as before, the supernatant was removed and the pellet was finally resuspended in 4 ml 0.9 % NaCl.

4.10.3 Library scale plating and incubation

Four quadratic Bio-Assay dishes (Nunc, Denmark) sized 243 x 243 x 18 mm were filled with 500 ml 4xDO agar each. Approximately 1 ml of the final yeast suspension was spread on each plate. The plates were incubated at 28 °C for 14 days.

5 Results

5.1 Genetic constructs

In order to use hnRNP M, hnRNP M deletion mutants, PLRG1, CDC5L and NOPSI in Yeast Two-Hybrid assays, the coding sequences of these proteins were cloned into the Yeast Two-Hybrid vectors that would express them as Yeast Two-Hybrid fusion proteins.
hnRNP M and its deletion mutants and NOPSI were cloned into pGBKT7, PLRG1 and CDC5L were cloned into pGADT7 (see Table 5.1).

Table 5.1 Genetic constructs

Coding sequence insert	Size or range of wild type (bp)	Vector
hnRNP M	2076	pGBKT7
hnRNP M mutant 1	1 - 744	pGBKT7
hnRNP M mutant 2	1 - 1749	pGBKT7
hnRNP M mutant 3	744 - 2076	pGBKT7
hnRNP M mutant 4	1035 - 1749	pGBKT7
hnRNP M RRM 1	1 - 465	pGBKT7
hnRNP M RRM 2	466 - 1035	pGBKT7
hnRNP M RRM 3	1844 - 2076	pGBKT7
PLRG1	1548	pGADT7
CDC5L	2409	pGADT7
NOPSI	1401	pGBKT7

The sequences for hnRNP M and PLRG1 were kindly provided by Marco Denegri from the University of Dundee, Scotland. The sequences for CDC5L and NOPSI were available at the laboratory of our working group.

5.1.1 Generation of genetic constructs

PCR was used to amplify the coding sequences of all proteins of interest. Deletion mutant sequences of hnRNP M were obtained by using primers that bound to the 5' and 3' ends of the desired region. Restriction sites were added to all sequences by appropriate PCR primers that were designed according to the rules described in 4.1.1. The chosen restriction sites were not present within the sequence but were available within the MCS of the target vector.
Sequences of all PCR primers used in this work are listed in Appendix D. PCR mixtures without template served as negative control to ensure that amplification arose from the intended template.

The amplicons were applied to Agarose Gel Electrophoresis (refer to 4.1.2). Bands matching the size of the desired sequence were cut out of the gel. The obtained gel slices were dissolved and the contained DNA was retained using a Promega Clean-Up System (refer to 4.1.3). All amplicons and were digested with restriction enzymes matching their added restriction site and purified as described in 4.1.4 and 4.1.3. Aliquots of the Yeast Two-Hybrid vectors pGBKT7 and pGADT7 were cut at the restriction sites corresponding to those of their intended inserts. The vectors linearised in this way were treated with 0.5 units Calf Intestinal Phosphatase per µg DNA to prevent self-ligation. Figure 5.1 shows CDC5L amplicons, Figure 5.2 shows a selection of digested and purified insert and vector DNA prepared for ligation, made visible by gel electrophoresis. (See Appendix E for detailed depictions of all used DNA standards.)

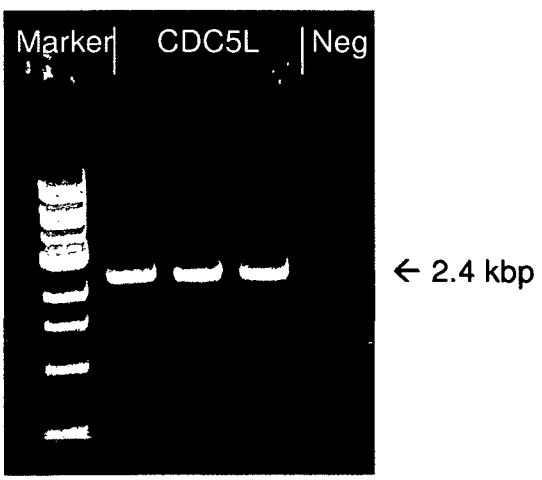


Figure 5.1 CDC5L amplicons. *Marker: NEB 1kb DNA Ladder. CDC5L: The 2.4 kbp CDC5L coding sequence was amplified using PCR and made visible on an 1%. agarose gel by ethidium bromide staining. Restriction sites for later digest and ligation were added using appropriately designed PCR primers. Neg: Negative control. A PCR mixture without template resulted in no amplification.*

Ligation of inserts and vectors was performed over night at 16 °C as in described in 4.1.5. Figure 5.3 shows the sample composition after ligation of CDC5L and pGADT7 using Nde I and Cla I restriction sites. Approximately 50 % of applied vector molecules were successfully ligated with an insert in this representative example, but this amount was sufficient for following transformation of the E. coli host strain TOP10.

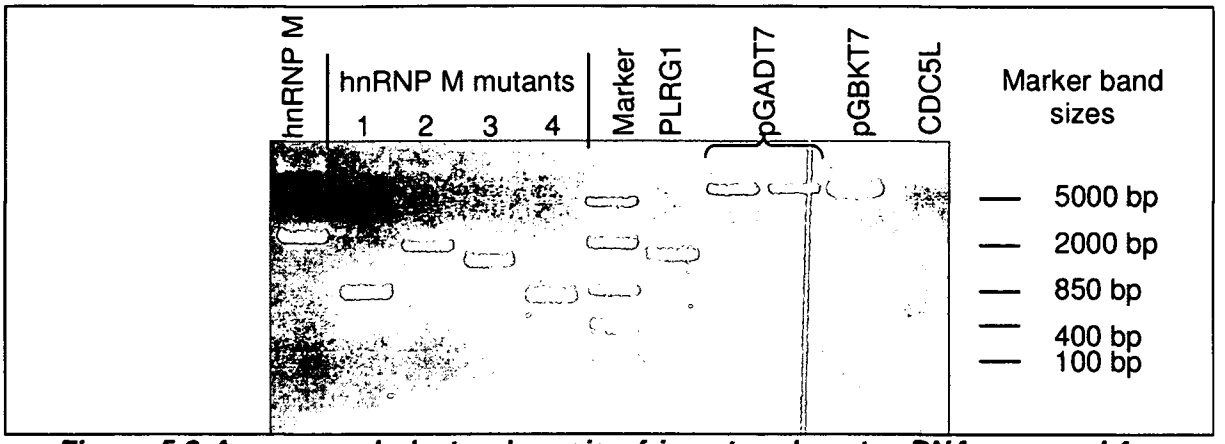


Figure 5.2 Agarose gel electrophoresis of insert and vector DNA prepared for ligation. *Insert sequences of hnRNP M, hnRNP M deletion mutants 1 – 4, PLRG1 and CDC5L were made visible by agarose gel electrophoresis (accordingly labelled lanes). Lanes pGADT7 and pGBKT7 show the according Yeast Two-Hybrid vectors linearised for ligation with insert sequences. Marker: Fermentas FastRuler™ DNA ladder, Middle range.*

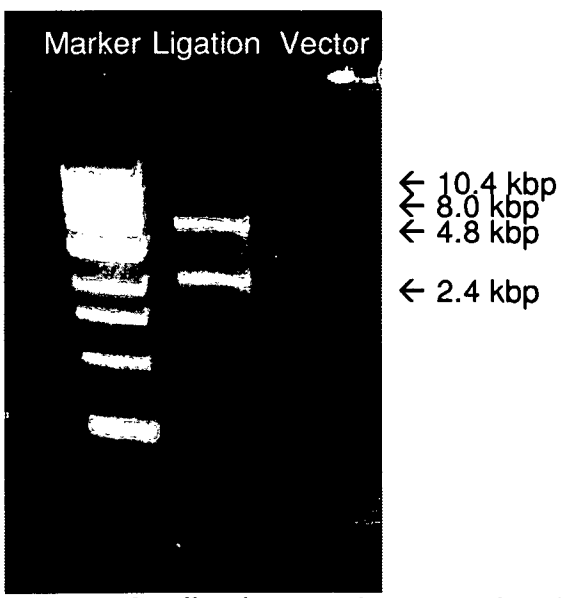


Figure 5.3 Representative result of a ligation on the example of CDC5L and pGADT7. *Marker: NEB 1kb DNA Ladder. Ligation: Four bands are visible, the uppermost band of approximately 10.4 kbp represents the succesfully ligated vector with insert. The 8.0 kbp band resembles unligated vector. The strong band at 4.8 kbp containes CDC5L dimers and the lowest band at 2.4 kbp is unligated CDC5L. Vector: Linearised pGADT7 with a size of 8.0 kbp ensures the identity of the two upper bands in the ligation sample separated in lane Ligation.*

5.1.2 E. coli transformation and transformant selection

The E. coli strains TOP 10 or DH10B were used to propagate the newly created vector-insert constructs to provide sufficient amounts for analysis and for later application in Yeast Two-Hybrid assays.

All vector-insert constructs were used to transform competent E. coli TOP10 by electroporation as described in 4.2. After transformation E. coli was cultivated on LB medium containing Kanamycin to select for pGBKT7 constructs or Ampicillin to select for pGADT7 constructs. Untransformed E. coli were separately plated as negative control. Not all resulting clones contained the desired construct. Some clones contained plasmids without inserts and some did not contain the desired plasmid at all. By PCR screening and subsequent control digest of purified plasmids it was possible to identify clones positive for the desired construct. Figure 5.4 shows all possible outcomes of a PCR screening. On the example of pGBKT7-hnRNP M Figure 5.5 shows the result of a PCR screening of 6 clones. 5 turned out to be positive and this result was verified by control digest.

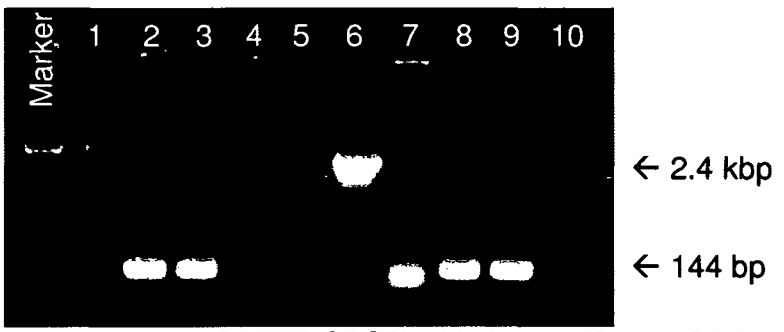


Figure 5.4 PCR screening for pGADT7-CDC5L. *Marker: NEB 1kb DNA ladder. Lanes 2, 3, 8 and 9: Clear bands at very low size indicate the short stretch that is amplified from an empty vector by the primers used for PCR screening. Lanes 1, 4, 5, 7 and 10: Clones without plasmid do not produce clear bands. Lane 6: The approximately 2.4 kbp band indicates CDC5L amplified from CDC5L-pGADT7.*

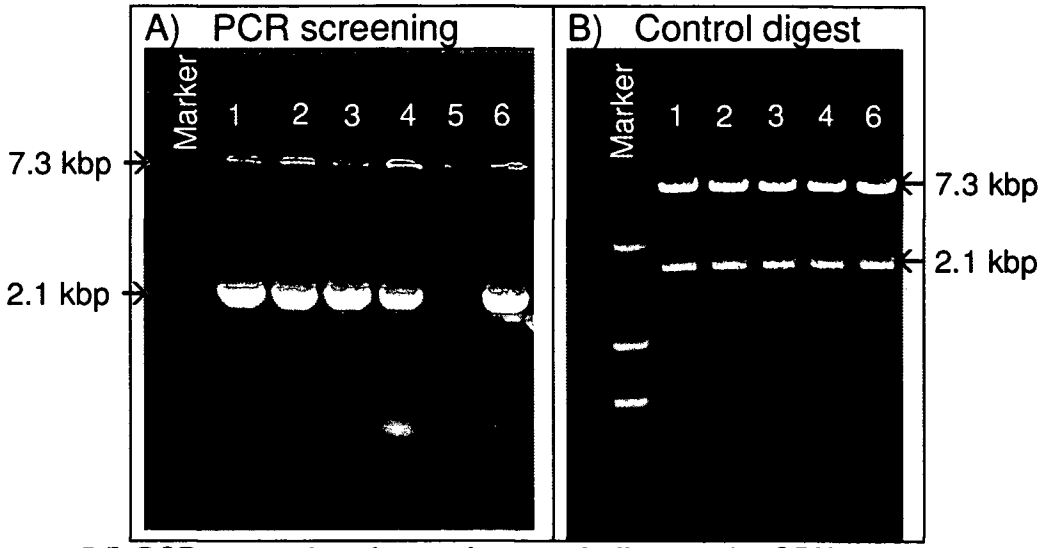


Figure 5.5 PCR screening for and control digest of pGBKT7-hnRNP M. A) *Marker: NEB 1kb DNA ladder. Lanes 1 – 6: PCR screening was used to identify clones positive for pGBKT7-hnRNP M. The strong bands at 2.1 kbp indicate hnRNP M in all but one clone. A negative control was provided but is not shown.* ***B)*** *Marker: NEB TriDye 2-log. Lanes 1 – 4 and 6: Plasmids were prepared from the clones found to be positive by PCR screening and were subjected to control digest. All five clones show a vector band and a 2.1 kbp band for hnRNP M.*

–80 °C stocks of all positive clones were established. Plasmids of positive clones were subject to verification of the desired insert sequence.

5.1.3 Verification of insert DNA sequences

When DNA is multiplied by PCR or within the bacterial host its sequence can be mutated. Therefore it was necessary to determine the insert DNA sequence of the plasmids that would be used in Yeast Two-Hybrid assays, to ensure the resulting proteins were unaltered.

Two clones of each insert-vector construct were chosen. The DNA sequences of these inserts were determined by DNA Dye Terminator Sequencing as described in 4.6. To obtain DNA of sufficient amount and purity the sequences were first amplified by 30 PCR cycles and then purified using the Promega Clean-Up System. In a second PCR procedure the purified sequences were used as templates to create single stranded DNA chains that were terminated by a fluorescent-labelled ddNTP. Sequences of slightly more than 500 bp could be determined at a time and therefore sequencing primers were designed to cover all sequences of interest in blocks of a maximum length of 500 bp. The resulting labelled chains were purified by ethanol precipitation.

An Applied Biosystems/Hitachi ABI Prism® 3100-Avant Genetic Analyzer was used to separate these chains by capillary electrophoresis and detect the fluorescence labels that indicate the terminal nucleotide of each chain. The resulting electropherograms were translated into DNA code and examined using DNAStar software.
The sequences of hnRNP M, and hnRNP M mutants 1, 2 and 4 were determined at an early stage by IBL, Austria.

The determined sequences for wild type hnRNP M were compared to the hnRNP M var 2 nucleotide sequence provided by NCBI, accession number NM_031203, and the hnRNP M amino acid sequence provided by Swiss-Prot, accession number P52272. These two databank sequences deviate in four amino acid positions. The clones of hnRNP M were found to resemble the Swiss-Prot sequence in three of these four differing positions and match the NCBI sequence in one differing position. This characteristic was found in all clones of hnRNP M. The template received from the University of Dundee can therefore be assumed to encode a natural variation of hnRNP M. An hnRNP M clone that contained only these deviations, and one silent mutation when compared to the NCBI nucleotide sequence, was obtained. See Table 5.2 for the differences between cloned hnRNP M and databank sequences.

Table 5.2 Differences between cloned hnRNP M and databank sequences

Amino acid position	hnRNP M clone	hnRNP M P52272	hnRNP M NM_031203
aa 34 (nt 100 - 102)	Pro (CCG)	Pro	Ser (TCG)
aa 143 (nt 427 - 429)	Lys (AAA)	Lys	Lys (AAG)
aa 152 (nt 454 - 456)	His (CAT)	His	Cys (TGT)
aa 159 (nt 475 - 477)	Lys (AAG)	Gln	Lys (AAG)
aa 488 (nt 1462 - 1464)	Leu (CTG)	Leu	Pro (CCG)

The sequences of cloned hnRNP M deletion mutants were compared to the established wild type sequence. Deletion mutants 1 and 2 and the RRM domains 2 and 3 were obtained without further deviations. hnRNP M deletion mutant 3 was obtained containing one silent mutation and one resulting in an amino acid change from Leucine to Proline roughly at the middle of its sequence. In the deletion mutant 4 the very C-terminal amino acid is shifted from Alanine to Threonine. This mistake can only have occurred due to a mistake at primer manufacturing, since the according triplet lies within the primer binding site. In the deletion mutant containing RRM 1, the deviation found for amino acid 152 when comparing hnRNP M clones to NCBI databank sequence appears to be reverted. Again, the according triplet lies within the primer binding site. Because the primer sequence was designed based on the

NCBI data, the hnRNP M clone sequence was changed at this position during PCR. See Table 5.3.

Table 5.3 Differences between cloned hnRNP M deletion mutants and cloned wild type hnRNP M

Amino acid position	hnRNP M mutant 3	hnRNP M wild type
aa 340 (nt 1018 - 1020)	Ala (GCG)	Ala (GCA)
aa 341 (nt 1021 - 1023)	Pro (CCG)	Leu (CTG)
	hnRNP M mutant 4	
aa 583 (nt 1747 - 1749)	Thr (ACC)	Ala (GCC)
	hnRNP M RRM 1	
aa 152 (nt 454 - 456)	Cys (TGT)	His (CAT)

The determined sequence of CDC5L was compared to the CDC5L nucleotide sequence provided by NCBI, accession NM_001253. A clone containing two silent mutations was obtained (Table 5.4).

Table 5.4 Differences between cloned CDC5L and databank sequence

Amino acid position	CDC5L clone	CDC5L NM_001253
aa 530 (nt 1588 - 1590)	Arg (CGC)	Arg (CGT)
aa 615 (nt 1843 - 1845)	Tyr (TAC)	Tyr (TAT)

The determined sequence of PLRG1 was compared to the PLRG1 nucleotide sequence provided by NCBI, accession NM_002669 (Table 5.5). Three clones were examined. They showed six identical deviations from the NCBI sequence. Five of these deviations were also found in the template sequence, the last is a silent mutation and it was not examined in the template sequence. Proteins derived from this template have been successfully used by Marco Denegri at the University of Dundee and it can be assumed to be a natural variation. A clone containing only these deviations was obtained.

Table 5.5 Comparison of cloned PLRG1, PLRG1 databank sequence and the PLRG1 template

Amino acid position	PLRG1 clone	PLRG1 NM_002669	PLRG1 template
aa 239 (nt 715 - 717)	Lys (AAG)	Lys (AAA)	Lys (AAG)
aa 261 (nt 781 - 783)	Pro (CCG)	Leu (CTG)	Pro (CCG)
aa 294 (nt 880 - 882)	Leu (CTG)	Leu (TTG)	Leu (CTG)
aa 311 (nt 931 - 933)	Thr (ACG)	Thr (ACT)	Thr (ACT)
aa 344 (nt 1030 - 1032)	Gln (CAG)	Gln (CAA)	Gln (CAG)
aa 406 (nt 1216 - 1218)	Ile (ATA)	Ile (ATT)	n.d.

5.1.4 Preparation of all plasmids in sufficient amounts

E. coli stocks containing the plasmids of intended sequence were cultivated as described in 4.4.1 for medium and large scale plasmid preparation. Plasmids were prepared using the Qiagen Plasmid Midi/Maxi kit. As final control all plasmid preparations were subjected to a control digest as described in 4.5.2. Figure 5.6 shows the positive outcome of this final assay.

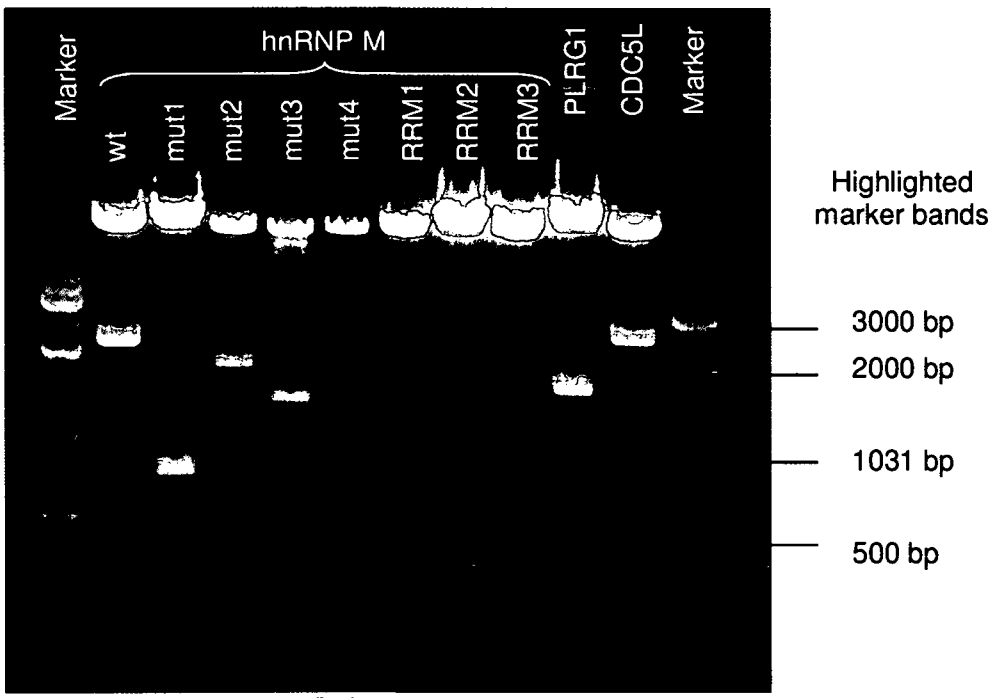


Figure 5.6 Final control digests. *Before applying the generated constructs to a Yeast Two-Hybrid assay their correct composition was checked by a final control digest. hnRNP M wild type (wt), hnRNP M deletion mutants 1 to 4 (mut 1 – 4) and sequences for hnRNP M RNA recognition motifs 1 to 3 (RRM 1 – 3) were successfully cut from pGBKT7. PLRG1 and CDC5L were successfully cut from pGADT7. Marker: Fermentas O'Gene Ruler DNA Ladder Mix (see Appendix E for detailed labelling of all marker bands).*

5.2 Liquid Yeast Two-Hybrid

5.2.1 Establishment of Liquid Yeast Two-Hybrid

In order to identify the interaction domain of hnRNP M towards PLRG1 and CDC5L a GAL4 based Yeast Two-Hybrid system was used to examine which hnRNP M deletion mutants interact with PLRG1 and CDC5L.

The yeast reporter strain AH109 was co-transformed with all possible combinations of hnRNP M deletion mutants as bait proteins and PLRG1 and CDC5L as prey proteins. Successfully co-transformed clones were selected on 2xDO medium lacking tryptophan and leucine, using the nutritious markers on each plasmid (refer to 2.5).

In this system further growth of the yeast reporter strain on selective 4xDO medium is possible only if the proteins encoded on the Yeast Two-Hybrid plasmids interact and thus activate the reporter genes HIS3 and ADE2, as described in 2.5. This re-enables the synthesis of histidine and adenine that are both absent in 4xDO medium. Resulting yeast growth is the parameter indicating an interaction.

In usual Yeast Two-Hybrid assays the selective 4xDO medium is agar based. A protein-protein interaction is then indicated by the growth of colonies.

In this work at least hnRNP M was found to have a growth inhibiting effect on the yeast reporter strain, as is described in detail in 5.4. Most probably therefore, yeast growth on 4xDO medium was too weak to be observed. The incapability to indicate very weak growth caused by detrimental effects of the examined proteins is a disadvantage of usual Yeast Two-Hybrid approaches.

To overcome this disadvantage co-transformed clones were transferred not to 4xDO agar but to small volumes of liquid 4xDO medium, contained in a standard 96-well microtiter plate. For each co-transformation type twelve clones were chosen to build one sample. Yeast growth was then monitored by consecutive OD-measurement (see 4.8.2). In this way it was possible to observe even small changes in yeast cell mass. Figure 5.7 shows this capability of Liquid Yeast Two-Hybrid.

As an additional advantage the resulting data is not depending on eye judgment but is generated by photometric measurement. It is numeric and can thus be instantly applied to statistic procedures.

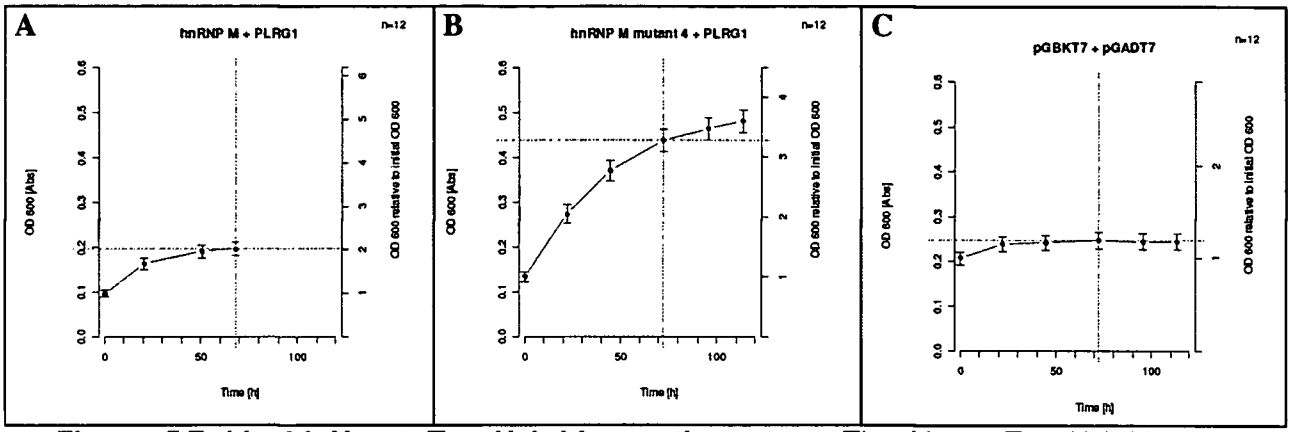


Figure 5.7 Liquid Yeast Two-Hybrid growth curves. *The Yeast Two-Hybrid reporter strain AH109 was co-transformed with pGADT7-PLRG1 plus pGBKT7-hnRNP M (A) and pGBKT7-hnRNP M aa 345-583 (B) and with empty pGADT7 plus pGBKT7 (C). The co-transformants were cultivated in liquid 4xDO medium for at least three days. In Liquid Yeast Two-Hybrid, growth of the reporter strain is determined by consecutive OD 600 measurement. The mean OD 600 values were plotted versus time, error bars indicate the standard error of the mean. Growth correlates with the protein-protein interaction and hybrid protein effects on the host organism:* ***A)*** *hnRNP M and PLRG1 interact but the cell growth is inhibited by detrimental effects of hnRNP M. Still, growth is clearly detectable.* ***B)*** *The hnRNP M region between aa 345 and aa 583 is interaction positive with PLRG1 and has less adverse effects on the host.* ***C)*** *Interaction negative clones are unable to grow. For all samples n = 12.*

5.2.2 Liquid Yeast Two-Hybrid data

Sample data resulting from a Liquid Yeast Two-Hybrid assay consists of an OD 600 value for each clone at each point of measurement.

To characterise a protein-protein interaction the mean increase in cell mass for a sample had to be determined. Initial OD 600 values (X) were compared to the OD 600 values after three days of incubation (Y). Each sample consists of twelve clones starting from twelve different initial OD values and reaching twelve different OD values within three days. But each single clone should fit into the model

$$Y = b * X + E \qquad (1)$$

which is a simple homogenous linear regression model (Bartlett, 1933). Figure 5.8 shows that this regression model is well suited to describe Liquid Yeast Two-Hybrid data. In this model the factor b describes the relative increase in cell mass while E refers to a random

error caused by the clone itself or unknown factors. An assessed value for b can be calculated easily as

$$\hat{b} = \frac{\sum_{i=1}^{n} x_i y_i}{\sum_{i=1}^{n} x_i^2} \qquad (2)$$

Since b^ is an assessed value gained from a limited number of observations it is necessary to identify its probable range. The standard error for b^ indicates this range. It is calculated as

$$SE_{\hat{b}} = \sqrt{\frac{\sum_{i=1}^{n} (y_i - \hat{b}x_i)^2}{(n-1)\sum_{i=1}^{n} x_i^2}} \qquad (3)$$

$\hat{b}$ is used to characterise a protein-protein interaction. In this work growth between 100 % and 200 %, indicated by an increase factor $\hat{b}$ between 1 and 2, was viewed to account for weak interactions. Growth above 200 %, or $\hat{b}$ greater than 2, was defined to indicate strong interactions.

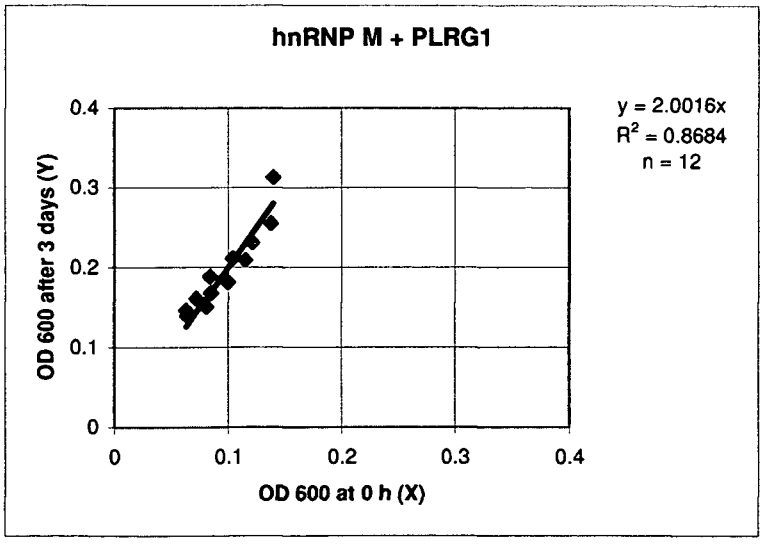


Figure 5.8 Well correlated data can be obtained. *This scatter-plot shows initial OD 600 values (X) vs. OD 600 values after three days of incubation (Y) for twelve wells of an hnRNP M plus PLRG1 co-transformation. The data are well correlated and increase in biomass can be described by the linear regression model Y = b * X + E. The factor b describes the relative increase in cell mass, which is the parameter directly linked to protein-protein interactions.*

5.3 Interaction domain mapping

Using the method of Liquid Yeast Two-Hybrid each hnRNP M deletion mutant was tested for interaction with PLRG1 and CDC5L. As the main result of this work, the interaction domain of hnRNP M with both PLRG1 and CDC5L was found to be located between amino acid 345 and 583 of hnRNP M transcript variant 2.

5.3.1 Interaction of hnRNP M with PLRG1

The yeast reporter strain AH109 was co-transformed with hnRNP M and its deletion mutants encoded on pGBKT7 and PLRG1 encoded on pGADT7. Co-transformed clones were applied to Liquid Yeast Two-Hybrid. Co-transformants containing untreated pGBKT7 and pGADT7 served as negative control. Figure 5.9 shows representative results for this examination. A complete list of results is included in Appendix A.

hnRNP M and hnRNP M deletion mutants	**Interaction with PLRG1**	
	Interaction	b (+/- SE)
Wild type 1 — 691	+	**2.002** (+/- 0.053)
Mutant 1 1 — 248	-	**1.126** (+/- 0.012)
Mutant 2 1 — 583	-	**1.328** (+/- 0.018)
Mutant 3 248 — 691	+	**2.492** (+/- 0.103)
Mutant 4 345 — 583	++	**3.186** (+/- 0.127)
	Negative control	
pGBKT7 + pGADT7	-	**1.196** (+/- 0.018)

Figure 5.9 Mapping the interaction domain of hnRNP M and PLRG1. *The hnRNP M region between aa 345 and aa 583 was found to be responsible for interaction with PLRG1. This region contains a methionine and arginine rich repeat motif (blue box). The deletion mutant ranging from aa 1 to aa 583 includes the interaction domain but is interaction negative in this assay probably due to adverse effects of hnRNP M on the reporter strain.*

5.3.2 Interaction of hnRNP M with CDC5L

The interaction of hnRNP M with CDC5L has been examined in exactly the same way as that between hnRNP M and PLRG1. AH109 clones co-transformed with hnRNP M deletion mutants and CDC5L were selected and applied to Liquid Yeast Two-Hybrid. Representative results are shown in Figure 5.10, while all results are listed in Appendix A. The same region was found to be responsible for interaction between hnRNP M and CDC5L as for hnRNP M and PLRG1.

hnRNP M and hnRNP M deletion mutants		Interaction with CDC5L	
		Interaction	b (+/- SE)
Wild type	1 — 691	-	**1.491** (+/- 0.110)
Mutant 1	1 — 248	-	**1.510** (+/- 0.014)
Mutant 2	1 — 583	-	**1.499** (+/- 0.020)
Mutant 3	248 — 691	+	**2.600** (+/- 0.094)
Mutant 4	345 — 583	+	**2.982** (+/- 0.199)
		Negative control	
pGBKT7 + pGADT7		-	**1.280** (+/- 0.046)

Figure 5.10 Mapping the interaction domain of hnRNP M and CDC5L. *Interaction of hnRNP M with CDC5L appears to be similar to its interaction with PLRG1. The interaction domain is located between amino acids 345 and 583 and adverse effects of hnRNP M and its deletion mutant ranging from amino acids 1 to 583 probably inhibit growth of the reporter strain.*

5.3.3 Growth on solid medium

Growth resulting from positive interactions was also made visible directly. After incubation in liquid 4xDO medium clones containing hnRNP M mutant 4 plus PLRG1 or CDC5L were streaked out on 4xDO plates (refer to 4.8.5). They were well able to form colonies within a few days. Clones co-transformed with untreated pGBKT7 and pGADT7 were treated in the same way to serve as negative control. No growth was observed for control clones. See Figure 5.11.

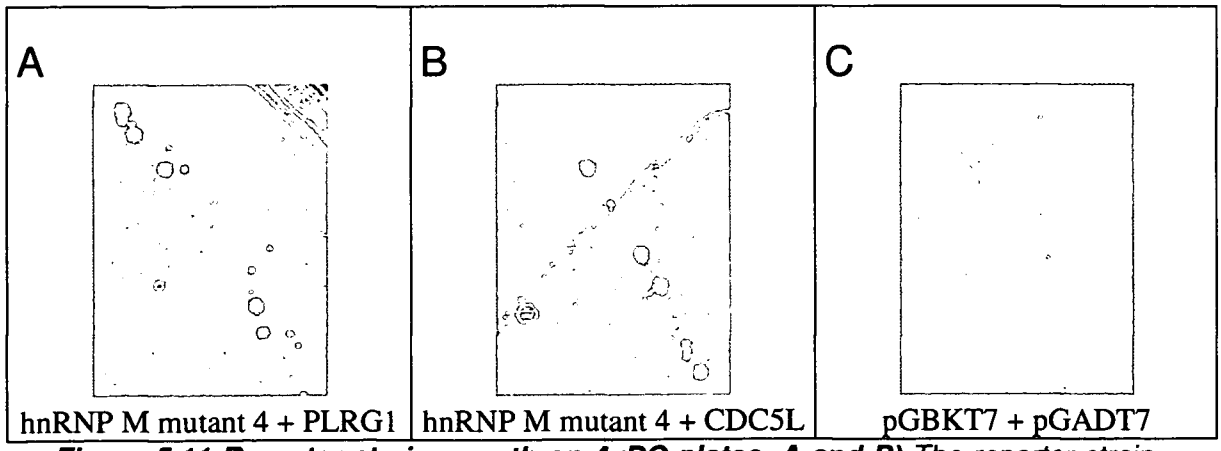


Figure 5.11 Reporter strain growth on 4xDO plates. A and B) *The reporter strain co-transformed with the interaction domain containing hnRNP M mutant 4 and PLRG1 or CDC5L was streaked out directly from Liquid Yeast Two-Hybrid wells after incubation there. Clearly visible colonies were obtained.* ***C)*** *Negative control streaks produced no colonies.*

5.3.4 Interaction domain has no auto-activation ability but some other effect

In a Yeast Two-Hybrid study it is crucial to verify that positive results did not arise from the ability of one protein to auto-activate the reporter genes (see 2.5.6).
To do so, the reporter strain was co-transformed with the hnRNP M mutant found to include the interaction domain and untreated pGADT7 and applied to the Liquid Yeast Two-Hybrid assay. Again untreated pGBKT7 plus pGADT7 served as negative control.
As described in 4.8.1 Liquid Yeast Two-Hybrid assays were performed undiluted and roughly 1:10 diluted. Usually the 1:10 dilution did not produce valid data because the initial OD 600 values were almost equally low as values for pure medium (OD 600 = 0.045). In the second run examining the hnRNP M interaction domain plus pGADT7 the mean inoculum density for twelve wells was 0.466 and it was not diluted exactly 1:10 but to a mean value of 0.080. This fact allows for the detection of some effect that might have been unidentified otherwise. Table 5.6 shows the resulting increase factors.

Table 5.6 Interaction domain auto-activation examination

	b (+/- SE)	
	Assay undiluted	Assay diluted
hnRNP M aa 345-583 + pGADT7	**1.656** (+/- 0.111)	**4.382** (+/- 1.184)
negative control	**1.280** (+/- 0.046)	**1.088** (+/- 0.021)

On the first view, a positive result found for the diluted assay is positive while it should be negative. On a second look the very high standard error of the increase factor b indicates abnormally high deviations from the model described in 5.2. An even closer look on the OD 600 values of all sample clones, depicted in Figure 5.12, reveals the presence of two subpopulations. The sample does no longer fit into the linear regression model and it is unsuitable for directly producing an interaction negative or positive result.

Both subpopulations can be viewed separately (Figure 5.13). They then fit well into the regression model. Subpopulation 1, consisting of five clones, turns out to be unusually fast growing while subpopulation 2 which contains seven clones shows almost no growth, as was originally expected. A similar outcome was found when the same examination was done before. No exact result was obtained in this case, because initial OD 600 values were below 0.05. Still a division into two subpopulations, one comparatively fast growing and one almost without growth is visible (Figure 5.14 A).

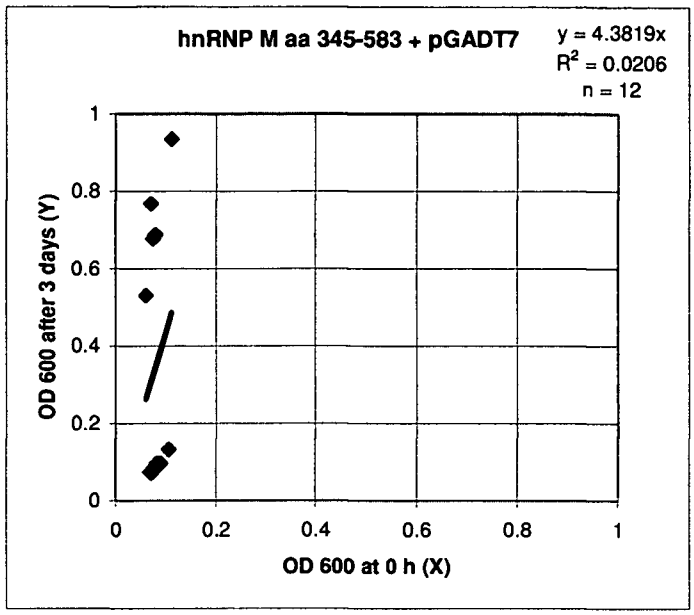


Figure 5.12 Scatter plot of hnRNP M interaction domain plus pGADT7 Liquid Yeast Two-Hybrid data. *In this depiction it becomes obvious that the sample of twelve hnRNP M aa 345-583 plus pGADT7 co-transformed clones consists of two clearly separate subpopulations. The regression model cannot describe the growth behaviour of the overall sample.*

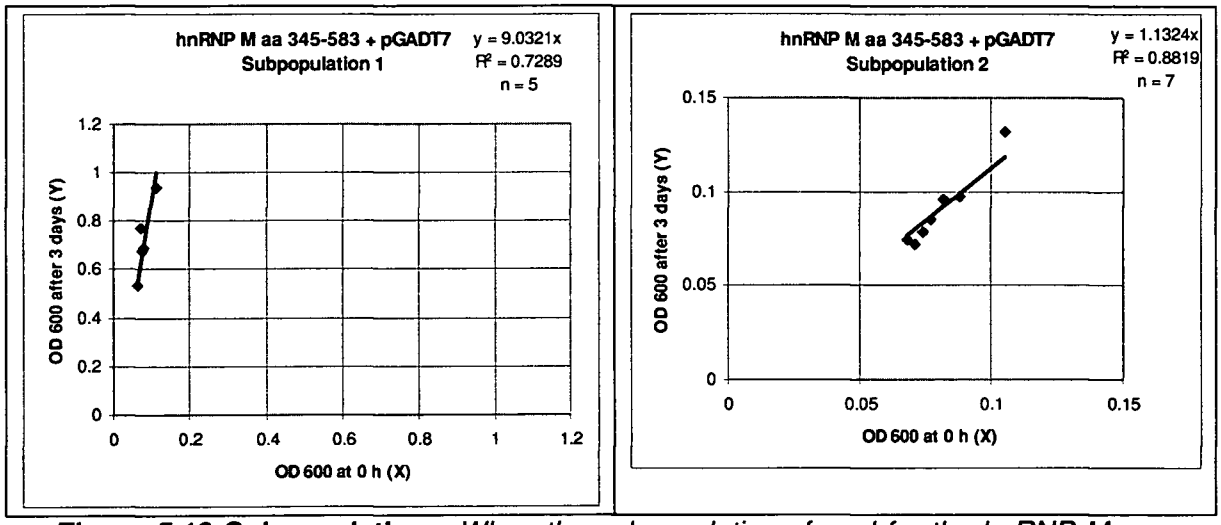


Figure 5.13 Subpopulations. *When the subpopulations found for the hnRNP M aa 345-583 plus pGADT7 co-transformation are separated their growth can be described by the linear homogenous regression model. Subpopulation 1 is fast growing, but the larger subpopulation 2 has an increase factor of only b = 1.13 (+/- 0.04). The growth behaviour of subpopulation 2 indicates that hnRNP M aa 345-583 has no auto-activation activity. The strong biomass increase observed in subpopulation 1 probably accounts for some mutagen or random effect of this hnRNP M region.*

In a final assay a similar result was obtained using a standard Yeast Two-Hybrid approach. The reporter strain AH109 was co-transformed with hnRNP M aa 345-583 plus CDC5L, hnRNP M aa 345-583 plus pGADT7 and pGBKT7 plus pGADT7 as negative control. Co-transformed clones were selected on 2xDO medium. Three clones of each type were resuspended in 0.9 % NaCl and then spread on 4xDO plates with a single streak and incubated at 30 °C. All hnRNP M aa 345-583 plus CDC5L clones were growth positive and all pGBKT7 plus pGADT7 clones were negative. Of the hnRNP M aa 345-583 plus pGADT7 clones two were negative and one was positive (Figure 5.14 B).

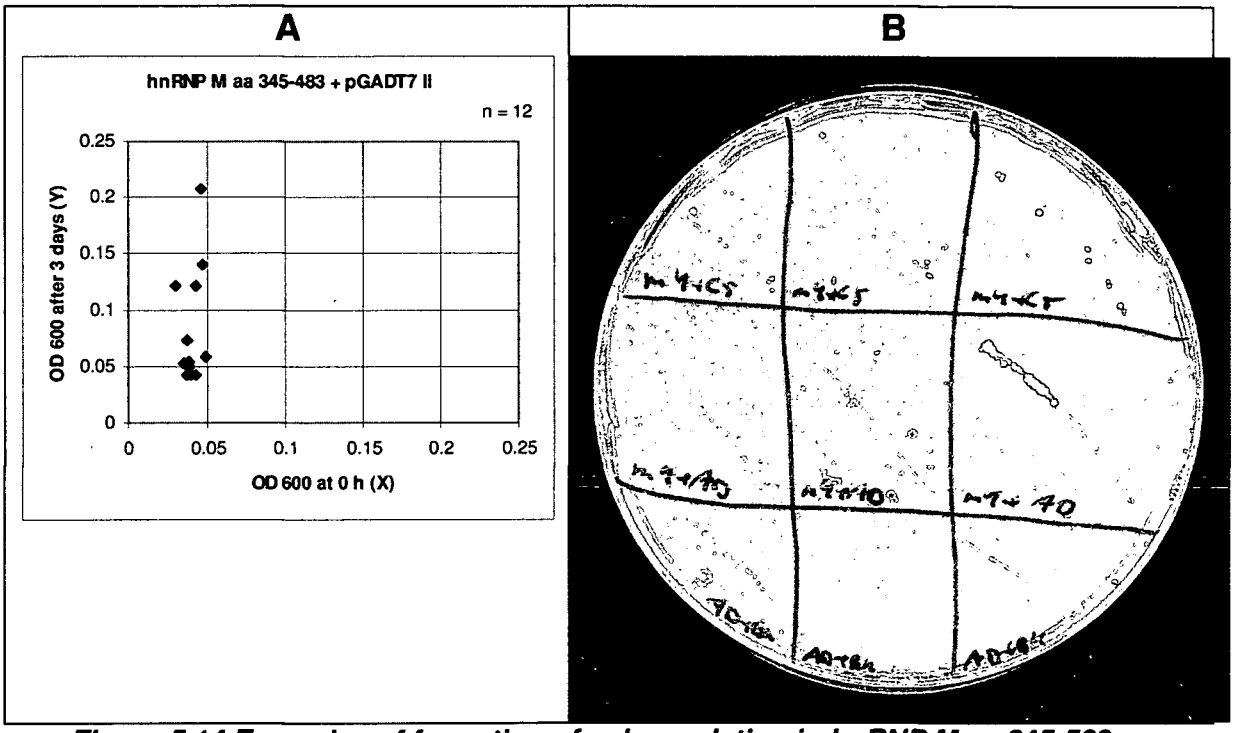


Figure 5.14 Examples of formation of subpopulation in hnRNP M aa 345-583 plus pGADT7 co-transformation samples. A) *This Liquid Yeast Two-Hybrid of hnRNP M aa 345-583 plus pGADT7 resulted in the formation of two subpopulations as observed before.* ***B)*** *Row 1: All three clones of hnRNP M plus CDC5L examined in a standard Yeast Two-Hybrid procedure are positive. Row 2: Of three examined hnRNP M aa 345-583 plus pGADT7 clones two are negative and one is positive. Row 3: The three pGBKT7 plus pGADT7 clones that serve as negative control are all negative.*

This leads to the assumption that the hnRNP M interaction region has no auto-activation ability, because this would result in growth of all or at least almost all according clones. Also the undiluted sample showed an increase factor reliably below 2.0. Still this hnRNP M region has an effect that occurs with a certain probability and allows the affected clone to grow under otherwise selective conditions.

5.3.5 PLRG1 has no auto-activation ability

A possible auto-activation ability of PLRG1 was examined in a similar way. The reporter strain was co-transformed with untreated pGBKT7 and PLRG1 encoded in pGADT7 and applied to Liquid Yeast Two-Hybrid. PLRG1 was found to have no auto-activation activity, nor any other growth enhancing effect (Table 5.7).

Table 5.7 PLRG1 auto-activation examination

	b (+/- SE)
PLRG1 + pGBKT7	**1.185** (+/- 0.032)

5.4 Yeast cell size measurment

The phenotype of transformed yeast was routinely investigated. In doing so, cells co-transformed with hnRNP M and PLRG1 were found to be visibly smaller than cells that were co-transformed with PLRG1 and an empty pGBKT7 (Figure 5.15). Aberrations in shape were not observed.

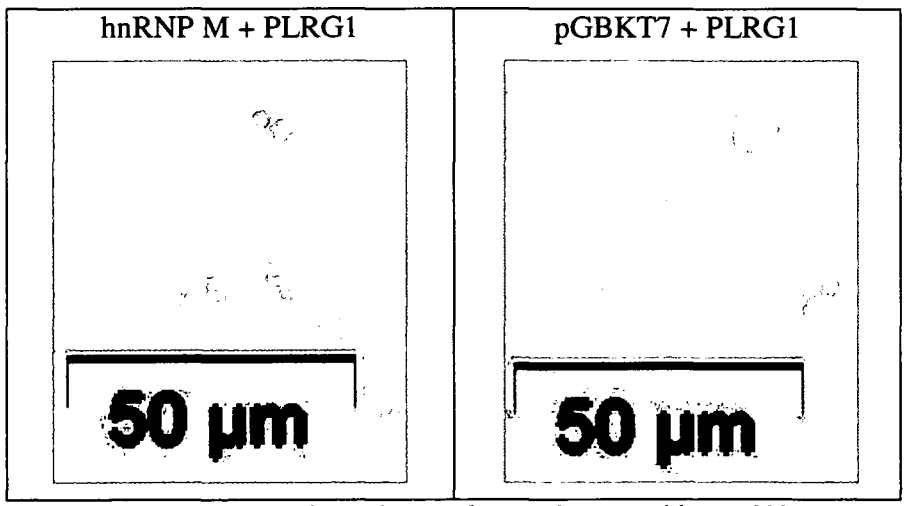


Figure 5.15 Microscopy samples of transformed yeast. *Yeast AH109 transformed with hnRNP M plus PLRG1 were observed to be smaller than yeast transformed with an empty vector pGBKT7 and PLRG1. Both co-transformants were simultaneously grown for 71 hours in liquid 4xDO. The pictures were taken at 100x magnification.*

To thoroughly examine if transformation with certain genetic constructs affects cell size, at least 25 cells of each co-transformation type were measured after incubation in liquid 4xDO medium. The cell size was determined as the longest possible diameter as described in 4.9.1. Two groups of samples were examined: All co-transformation types containing PLRG1 and all co-transformation types containing CDC5L.

The clones containing PLRG1 were grown simultaneously in one Liquid Yeast Two-Hybrid assay. The CDC5L containing clones were grown in two following assays. In each assay yeast co-transformed with empty pGBKT7 and pGADT7 (BK+AD) was included as control. In a first step the cell size distributions of all three control samples were compared using one-way ANOVA. No significant differences in cell size were found for these control clones ($p = 0.593$). This ensures that cell size is constant in repeated assays. Differences in cell size should thus be depending on the co-transformation type only.

The found cell-size distributions are depicted in Figure 5.16 as box and whisker plots. For the CDC5L containing group the two control samples were unified to one sample for this depiction and further calculation.

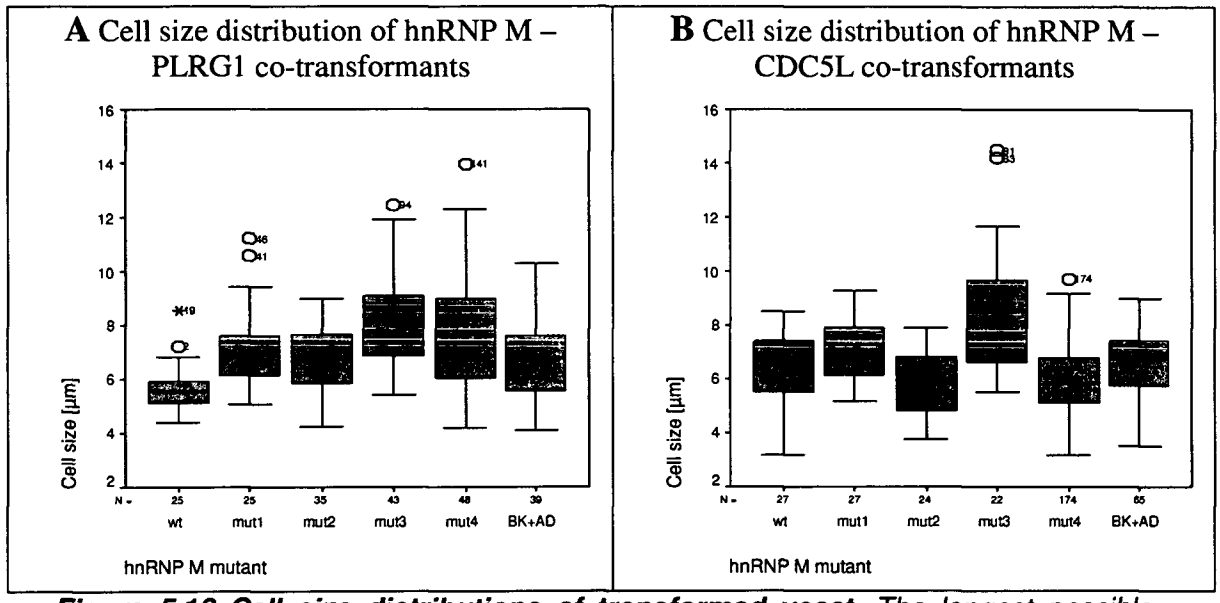


Figure 5.16 Cell size distributions of transformed yeast. *The longest possible diameter was measured of at least 25 cells of each co-transformation type. Cell size distribution of all clones containing wild type hnRNP M (wt) or an hnRNP M mutant (mut1-4) plus PLRG1 (A) or CDC5L (B) are depicted as box-and-whisker plots. Cell size distribution of control clones containing pGBKT7 and pGADT7 (BK+AD) are included as well.*

To find out which samples of one group were significantly different in cell size, they were compared pair-wise using Student's T-test. The results are summarised in Figure 5.17.

A **Mean cell sizes of yeast co-transformed with hnRNP M and PLRG1**

hnRNP M mutant	mut3	mut4	mut1	BK+AD	mut2	wt
Mean cell size	8.17 µm	7.47 µm	7.14 µm	6.77 µm	6.72 µm	5.66 µm
Significance						

B **Mean cell sizes of yeast co-transformed with hnRNP M and CDC5L**

hnRNP M mutant	mut3	mut1	BK+AD	wt	mut4	mut2
Mean cell size	8.70 µm	7.11 µm	6.52 µm	6.22 µm	5.93 µm	5.91 µm
Significance						

Figure 5.17 Significant differences in mean cell size. *T-Test was applied to pair-wise compare the mean cell sizes of co-transformed clones. Underlined values do not show significant difference at $\alpha = 0.05$. A and B refer to the same groups as in Figure 5.16.*

Yeast transformed with wild type hnRNP M plus PLRG1 turned out to be significantly smaller than all other samples of its group ($p < 0.001$) . The original observation was thus confirmed. However, yeast co-transformed with wild type hnRNP M and CDC5L was equally sized as the control cells and most other CDC5L containing co-transformants. In both groups the hnRNP M mutant 3 containing cells were the largest ones, but overall the order concerning hnRNP M mutants was not the same and is therefore not determined by the type of hnRNP M mutant alone.

This implies that hnRNP M may inhibit increase in yeast cell size, but this effect also depends on other factors. At very least it is manifest in clones that are also transformed with PLRG1.

5.5 Library scale co-transformation with NOPSI and hAorta cDNA Library

To find novel interaction partners of NOPSI (WBRSC20A) a Yeast-Two-Hybrid screening of this protein versus the Matchmaker Aorta cDNA library (Clontech Laboratories, Palo Alto, USA) was performed.

5.5.1 Determination of transformation efficiency

First the transformation efficiency was determined in a small scale assay. The reporter strain AH109 was co-transformed with NOPSI in pGBKT7 and the hAorta Library in pACT2 three times independently. After transformation, suspension was spread on -Leu, -Trp and 2xDO plates. DNA amounts and suspension volumes were as described in 4.7. The proceeding of this assay can be followed in Table 5.8. The amount of applied DNA, the number of resulting colonies and the according transformation efficiency are listed there.

Table 5.8 Determination of transformation efficiency

Drop out condition	-Trp	-Leu	2xDO
Selecting for	Nopsi	Library	Nopsi+Library
Amount of selecting DNA	0.3125 µg	0.3125 µg	1.25 µg
Run I	2984 cfu	353 cfu	170 cfu
Run II	2784 cfu	155 cfu	178 cfu
Run III	1366 cfu	140 cfu	105 cfu
Mean	2378 cfu	216 cfu	151 cfu
Transformation efficiency	7609.6 cfu/µg	691.2 cfu/µg	120.8 cfu/µg

Amount of selecting DNA is the amount in the plated volume of that construct that is selected for. In case of the 2xDO plates it is the combined amount of both constructs. Transformation efficiency is the mean number of colonies per amount of selecting DNA.

The construct NOPSI-pGBKT7 turned out to have more than ten times the transformation efficiency of the cDNA Library, which was therefore the limiting factor. The efficiency for co-transformation was 120.8 cfu/µg total DNA. When calculated for the limiting factor the co-transformation efficiency was 241.6 cfu/µg Library DNA. To obtain 10^6 cfu approximately 16600 µg Library DNA should be used in the interaction screening.

5.5.2 Library scale transformation

Only 1570 µg hAorta cDNA Library were available. This amount was still expected to provide a sufficient probability to identify novel interaction partners for NOPSI. Library scale co-transformation was performed as described in 4.10. The amount of applied NOPSI-pGBKT7 was 157 µg, the final resuspension volume was 4 ml. Approximately 1 ml was spread on each of four 530 cm^2 4xDO dishes. To monitor the transformation 50 µl were plated on -Trp and -Leu DO plates and 100 µl on a 2xDO plate. On the -Trp plate 2906 colonies were grown, indicating successful transformation with NOPSI-pGBKT7. On the -Leu plate the number of colonies could not be determined, because it had not been dried properly and the yeast was therefore smeared across the plate. The 2xDO plate, indicating successful co-

transformation, contained 3078 colonies. The transformation efficiency calculated from this value was 100 cfu/µg DNA.

In total 1416 colonies were obtained on 4xDO medium. Of these 656 were of white colour, 760 were red. Red colonies may indicate back mutations of the reporter strain, therefore white colonies were chosen for further examination. Table 5.9 shows the colony numbers for all four plates.

Table 5.9 Result of Library scale transformation

Plate	White colonies	Red colonies	Total colonies
1	69	57	126
2	111	134	245
3	306	271	577
4	170	298	468
Sum	656	760	1416

The counting of colonies and all further steps including plasmid preparation from presumably interaction positive clones and insert DNA sequencing were performed by Christine Marizzi and are described in her Diploma Thesis (Marizzi, 2007).

6 Discussion

6.1 Genetic constructs

Main aim of this work was to identify the interaction domain of hnRNP M with the CDC5L complex members CDC5L and PLRG1 using Yeast Two-Hybrid. Therefore deletion mutant coding sequences of hnRNP M were created by PCR. These sequences were cloned into the Yeast Two-Hybrid vector pGBKT7. The coding sequences of CDC5L and PLRG1 were cloned into the vector pGADT7 and all plasmids were propagated in E. coli.

E. coli clones containing a plasmid with the correct insert were identified by PCR screening and control digest of the purified plasmid.

During PCR and propagation in E. coli mutations of the insert sequence can occur. A crucial step at this stage was therefore to identify plasmids that correctly encoded the examined proteins. The insert sequences were determined by DNA dye termination sequencing and compared to the according sequences from the NCBI (National Center for Biotechnology Information) databank. The determined sequence of hnRNP M was also compared to the according sequence provided by the Swiss-Prot Protein knowledgebase.

The coding sequences of hnRNP M and PLRG1 used in this work were kindly provided by Marco Denegri from the University of Dundee. Interestingly, it turned out that these sequences slightly deviated from the databank sequences. Concerning hnRNP M the sequences provided by NCBI and Swiss-Prot deviate at four amino acid positions. The cloned coding sequence of hnRNP M was found to match the Swiss-Prot sequence in three of these cases and the NCBI sequence in one case. It seems therefore that the hnRNP M sequence used in this work resembles an additional natural variation. This variation is distinguished from the databank sequences only at positions already known to be variable. It might therefore be interesting to find out if additional variations of hnRNP M exist that represent other combinations of these four variable amino acids.

For PLRG1 six deviations of single nucleotides were found, of which one results in an amino acid change. This PLRG1 sequence is also assumed to be a natural variation, since proteins derived from this sequence have been successfully used at the University of Dundee. Especially the interaction of hnRNP M with PLRG1 was first observed using the same variants as in this work.

The coding sequence for CDC5L was available at our laboratory. A clone was obtained that contained two silent random mutations only.

The sequences of the hnRNP M deletion mutants were compared to the wild type sequence. Of the finally used sequences mutants 1 and 2 were perfectly correct.

Deletion mutant 3 contained one random silent mutation and one mutation in which a leucine residue was exchanged for a proline at amino acid position 340. This mutation was regarded not as a risk to the experiment but as a small chance to reveal further details about the interactions of hnRNP M because this amino acid position was included in one other deletion mutant and certainly in the wild type. If the mutated amino acid had turned out to affect an interaction the interaction domain could have been located to a region close around this position. Finally (compare 4.2 and 6.2) this position was found not to lie within the interaction domain of hnRNP M.
In deletion mutant 4 the very last amino acid was mutated. This definitely occurred due to a mistake at primer manufacturing because the mutated position was included in the primer binding sequence. Since the proteins contained fusion moieties anyway and this was a terminal amino acid the mutation can be regarded to have no detectable influence. In addition the position was included in other constructs in the correct form as well.
The power of a primer to change the amplified sequence was also observed in a clone containing just the first RNA recognition motif (RRM 1). At its 3' end this sequence contained one of the sites naturally deviating from the databank sequence. In detail the triplet for amino acid 152 was TGT, coding for cysteine while the databank sequence listed CAT, coding for histidine. So, two consecutive nucleotides were different. The according PCR primer was designed to suit to the databank sequence and thus the 3' ends of the PCR amplicons completely resembled the databank sequence and not that of the template.

Summarised, it was possible to obtain most required sequences in the appropriate Yeast Two-Hybrid vectors without mutations. The mutations contained in sequences that were finally used were not able to limit the success of the further interaction domain mapping but rather had the potential to give additional insight.

6.2 The Liquid Yeast Two-Hybrid

Yeast Two-Hybrid is an effective and simple to use technique to identify protein-protein interactions (compare 2.5). In this work it should be used to find out which hnRNP M deletion mutants interacted with CDC5L and PLRG1.
In a first assay the yeast reporter strain was co-transformed with hnRNP M and PLRG1. It was able to form colonies on medium selecting just for the presence of the plasmids. These colonies were transferred to 4x Drop out (DO) plates that selected for transcription of the reporter genes HIS3 and ADE2 due to protein-protein interaction. On this medium the colonies were unable to grow further although the interaction of hnRNP M with PLRG1 was strongly assumed and was in fact the basis for this work.

Therefore the fusion proteins had most probably an inhibiting effect on the reporter strain. Still growth might have occurred but was too weak to be observed.

In this situation it was necessary to improve the Yeast Two-Hybrid approach. The need to accurately detect small changes in reporter strain biomass led to the development of Liquid Yeast Two-Hybrid (compare 5.2). In this method successfully co-transformed clones are not transferred to another plate to select for protein-protein interactions but to small volumes of similar but liquid 4xDO medium. Growth in this liquid medium too indicates a positive interaction. In Liquid Yeast Two-Hybrid the reporter strain growth is easily monitored by consecutive measurements of the samples' optical density. As containment for the medium standard 96-well microtiter plates were chosen. All samples contained in one plate were suitably measured at once, using an appropriate photometer.

The first major advantage of this method was its sensitivity. A mere doubling of biomass was sufficient to distinguish interaction positive clones from negative controls that usually showed a biomass increase of 20 % to 40 % only. This method might therefore strongly increase the range of proteins suitable for Yeast Two-Hybrid among those proteins that could so far not be examined due to detrimental effects on the yeas host.

Secondly Liquid Yeast Two-Hybrid is an easily applied and fast method. Only standard laboratory equipment is required and several samples can be simultaneously examined using just one microtiter plate. Usually interaction indicating growth was detectable within 24 hours. Final results were obtained within three to five days.

As its third advantage Liquid Yeast Two-Hybrid makes quantification of the yeast growth independent from human judgment. This is especially intriguing since the quantity of reporter strain growth was found to roughly correlate with the strength of fusion proteins' binding strength (Estojak et al., 1995). Especially in understanding regulatory or dynamic functions of an interaction it is certainly not only of interest to know possible interacting factors but also the quantity of their affinity. If Yeast Two-Hybrid is used to estimate this quantity in correlation to reporter strain growth a most objective and unbiased method is required. Liquid Yeast Two-Hybrid meets this demand and so might increase the capacity of Yeast Two-Hybrid to semi-quantify interactions. At first, the rough quantification as weak or strong that was established in this work will have to be verified by independent methods, e.g. Biacore surface plasmon resonance assays. In this work this was not possible due to time limits.

6.3 Identification of the hnRNP M interaction domain with CDC5L and PLRG1

The major aim of this work was to identify the interaction domain of hnRNP M with CDC5L and PLRG1. All hnRNP M deletion mutants were tested for their interaction with CDC5L and PLRG1 using the newly developed method of Liquid Yeast Two-Hybrid.

The interaction domain to both CDC5L and PLRG1 was clearly found to be contained in hnRNP M deletion mutant 4 ranging from amino acids 345 to 583 of the wild type sequence. This mutant contained only the methionine and arginine rich repeat motif but no RNA recognition domains. hnRNP proteins typically contain RNA binding domains and distinct domains mediating protein-protein interactions (reviewed by Dreyfuss et al., 2002). This characteristic was herewith confirmed for hnRNP M.
Not all hnRNP M sequences that included the interaction domain yielded interaction positive results though, due to the detrimental effects initially observed.

Figure 6.1 shows how hnRNP M may be connected to the CDC5L complex based on the findings of this work.
However an interaction of hnRNP M might occur only with one of the two examined partners, since CDC5L and PLRG1 interact with each other and all proteins used in this study are well conserved across species. In a study about SNEV which too is a CDC5L complex member a chimera protein of SNEV and its yeast orthologue Prp19 was able to restore the phenotype of a Prp19 mutant yeast strain. The yeast part of the chimera protein was not sufficient to do so (Grillari et al., 2005). So it is possible that human hnRNP M binds only to CDC5L or PLRG1 and that interaction to the other protein is mediated over the yeast orthologue of the direct interactor in the Yeast Two-Hybrid assay. To find out which interactions are of direct nature independent methods, especially GST pull-down, will have to be applied. A part of the required experiments has already been performed by Marco Denegri, one of our partners at the University of Dundee, though the results are yet unpublished.

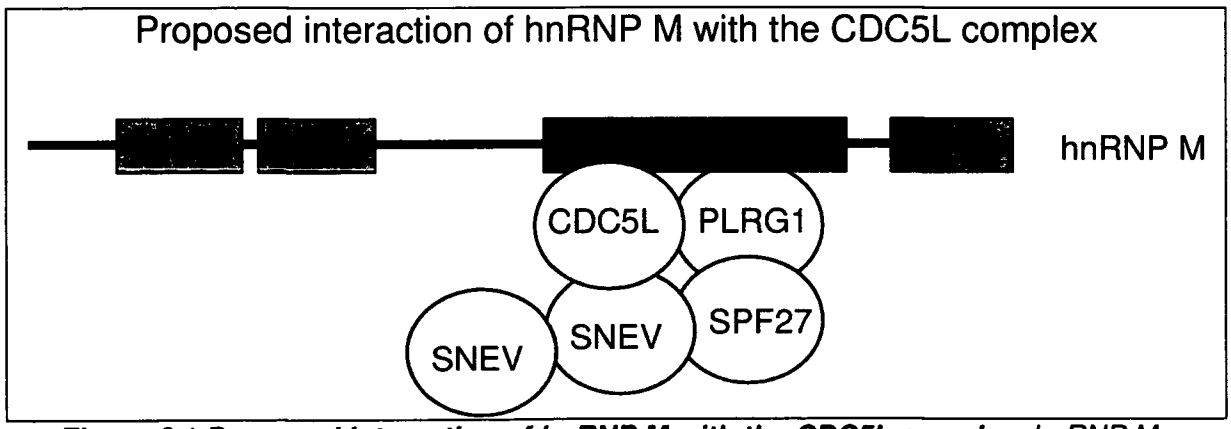


Figure 6.1 Proposed interaction of hnRNP M with the CDC5L complex. *hnRNP M was found to interact with the CDC5L complex proteins PLRG1 and CDC5L over a region between amino acids 345 and 583, containing just its methionine and arginine rich repeat motif (blue box). Due to these interactions a functional link between hnRNP M and the complete CDC5L complex seems possible. The depiction of the CDC5L complex includes its four components, CDC5L, PLRG1, SNEV and SPF27, and also accounts for the known homooligomerisation of SNEV (also see 2.2).*

Concerning the function of the interaction of hnRNP M with CDC5L and PLRG1 it is well possible that hnRNP M directs the CDC5L complex to the site of splicing.

hnRNP M has also been suggested to be a stress dependent trigger for splicing. It leaves the overall hnRNP population and binds to the nuclear matrix upon stress inducement. This redistribution seems to cause splicing to halt at an early stage of spliceosome assembly (Gattoni et al., 1996).
Three works from our partner group at the University of Dundee have examined the involvement of CDC5L, PLRG1and the CDC5L complex at this early stage of splicing. In one study immunodepleting nuclear extracts from CDC5L was observed to inhibit splicing just at the second catalytic step (Ajuh et al., 2000). But when the direct interaction of CDC5L and PLRG1 was inhibited by truncated CDC5L or PLRG1 peptides, spliceosome assembly was found to be inhibited in nuclear extracts (Ajuh et al., 2001). In the third work smaller peptides of just 15 or 24 amino acids that were derived from the interaction domains of CDC5L and PLRG1 with each other were used. These inhibited pre-mRNA splicing, but did not block the interaction of CDC5L and PLRG1 and they did not prevent assembly of the spliceosome (Ajuh and Lamond, 2003).
It seems therefore that CDC5L and the complete CDC5L complex are required only after spliceosome assembly but PLRG1 alone may well be required at pre-spliceosomal stages.
So as alternative to thinking that hnRNP M directs the complete CDC5L complex to the pre-mRNA we may imagine that hnRNP M usually directs PLRG1 to the pre-mRNA but fails to do so when it is redistributed due to cellular stress, which eventually inhibits splicing at the stage of spliceosome assembly.

Further experiments will have to examine the effect of blocking the interaction found between hnRNP M and CDC5L and PLRG1. They will reveal which if any of these models resembles actual nuclear processes.

6.4 Effects of hnRNP M on the yeast host

The proteins examined in this work are nuclear proteins and they have orthologes in yeast. In the Yeast Two-Hybrid assay they are expressed constitutively. So it is possible that they interfere with the host metabolism and cause the inability of the reporter strain to grow despite an interaction of the fusion proteins.
The detrimental effect of hnRNP M on the yeast reporter strain turned out to be visible in the phenotype of transformed yeast cells (see 5.4). The phenotype of transformed yeast was examined after incubation in 4xDO medium, selecting for protein-protein interactions. Here it

was initially observed that clones that were co-transformed with hnRNP M-pGBKT7 plus PLRG1-pGADT7 were of smaller cell size than clones that just contained PLRG1 and untreated pGBKT7.
In a subsequent step the cell size distributions of all co-transformants used in this work were determined. Yeast transformed with wild type hnRNP M and PLRG1 were found to be significantly smaller in cell size than control cells and all other clones containing PLRG1 and an hnRNP M mutant. For CDC5L containing clones a similar effect could not be observed.
These findings suggest that hnRNP M in some way inhibits increase in yeast cell size, but this effect or its manifestation seems to be dependent on additional factors.

Another interesting influence was found for hnRNP M mutant 4, the deletion mutant containing the interaction domain to PLRG1 and CDC5L, when it was tested for auto-activation ability in the Yeast Two-Hybrid System. In three independent assays clones co-transformed with hnRNP M mutant 4-pGBKT7 and untreated pGADT7 were found to divide into two sub-populations when incubated in 4xDO medium. The larger sub-population always showed no growth, but the smaller one was able to grow although in spite of selective conditions that should require reporter gene transcription.
Since only a fraction of clones was able to grow, the effect probably depends on a mutation, while an inherent auto-activation ability of hnRNP M mutant 4 can be excluded.
Most likely, either hnRNP M mutant 4 has a mutagen influence on the reporter strain that allows it to grow under selective conditions, or mutant 4 has an auto-activation ability that is only effective in combination with some mutational event in the host.
Astonishingly in all three assays almost exactly one third of the examined clones belonged to the fast growing sub-population while two thirds were not growing (compare 5.3.4). The effect seems thus to occur to a single clone with a probability of one third.
As hnRNP M and its functions are studied in more detail the actual nature of this effect will hopefully be revealed too.

6.5 Library screening for NOPSI interactors

Nopsi (WBSCR20A) is one of several genes deleted on chromosome 20 in Williams-Beuren syndrome (compare 2.4). It is linked to the spliceosome by its interaction with the splicing factor SKIP (Schörgenhumer, 2005).
In order to identify novel interactors of NOPSI a Yeast Two-Hybrid library screening with NOPSI as bait and a human aorta cDNA library as prey was planned. In this work the yeast reporter strain was co-transformed with NOPSI as bait and a human aorta cDNA library as prey. The co-transformed yeast was plated on 4xDO plates. All subsequent steps including

plasmid preparation and sequence determination were done by Christine Marizzi and are described in her Diploma Thesis (Marizzi, 2007).

The efficiency of co-transformation with NOPSI-pGBKT7 and the library in pACT2 was examined in a small scale experiment. The amount of library DNA available for the library screening was limited. Based upon the found transformation efficiency the available amount of library DNA would yield approximately 10^5 *cfu* instead of the aimed for 10^6 *cfu*. This was nonetheless viewed to be sufficient to identify novel interaction partners of NOPSI.

To scale up the transformation to the amount of library DNA available a modified transformation protocol was worked out (see 4.10) and applied. The transformation efficiency then determined for the large scale transformation was not significantly different from the efficiency observed in the small scale experiment. So scale up using the modified protocol was successful.

A surprisingly high number of colonies was eventually found on the 4xDO plates, counted by Christine Marizzi. Slightly more than half of these were of red colour instead of the usual white. Red colour is an indication of *ADE2* mutants, because a deficiency in ADE2 leads to accumulation of its substrate, the purine precursor 5-aminoimidazole-ribonucleotide, which is moved to the vacuole where it forms red pigments (Jones and Fink, 1982; Zonneveld and Van der Zanden, 1995). Red colonies in Yeast Two-Hybrid therefore transcribe the reporter gene *ADE2* only weakly and probably not to a fusion protein interaction but due to incomplete control of the responsible promoter.

The high number of colonies and the abundance of the red phenotype imply that the selective conditions were not stringent enough in this assay. Nonetheless several interesting NOPSI interactors were finally identified, among them the human ribosomal protein 15 (Marizzi, 2007), from the clones obtained in this work.

References

Ajuh P, Lamond AI. (2003) Identification of peptide inhibitors of pre-mRNA splicing derived from the essential interaction domains of CDC5L and PLRG1. Nucleic Acids Res. 31(21):6104-16.

Ajuh P, Sleeman J, Chusainow J, Lamond AI. (2001) A direct interaction between the carboxyl-terminal region of CDC5L and the WD40 domain of PLRG1 is essential for pre-mRNA splicing. J Biol Chem. 276(45):42370-81.

Ajuh P, Kuster B, Panov K, Zomerdijk JC, Mann M, Lamond AI. (2000) Functional analysis of the human CDC5L complex and identification of its components by mass spectrometry. EMBO J. 19(23):6569-81.

Alberts B, Johnson A, Lewis J, Raff M, Roberts K, Walter P. (2002) Molecular biology of the cell, 4th ed. New York: Garland Science.

Barr RK, Hopkins RM, Watt PM, Bogoyevitch MA. (2004) Reverse two-hybrid screening identifies residues of JNK required for interaction with the kinase interaction motif of JNK-interacting protein-1. J Biol Chem. 279(41):43178-89.

Bartlett M.S. (1993) On the theory of statistical regression. Proc. Royal Soc. Edinburgh, 53, 260.

Bellugi U, Lichtenberger L, Mills D, Galaburda A, Korenberg JR. (1999) Bridging cognition, the brain and molecular genetics: evidence from Williams syndrome. Trends Neurosci. 22(5):197-207. Review.

Bernstein HS, Coughlin SR. (1998) A mammalian homolog of fission yeast Cdc5 regulates G2 progression and mitotic entry. J Biol Chem. 273(8):4666-71.

Bernstein HS, Coughlin SR. (1997) Pombe Cdc5-related protein. A putative human transcription factor implicated in mitogen-activated signaling. J Biol Chem. 272(9):5833-7.

Burd CG, Dreyfuss G. (1994) Conserved structures and diversity of functions of RNA-binding proteins. Science. 265(5172):615-21. Review.

Chiodi I, Biggiogera M, Denegri M, Corioni M, Weighardt F, Cobianchi F, Riva S, Biamonti G. (2000) Structure and dynamics of hnRNP-labelled nuclear bodies induced by stress treatments. J Cell Sci. 113 (Pt 22):4043-53.

Datar KV, Dreyfuss G, Swanson MS. (1993) The human hnRNP M proteins: identification of a methionine/arginine-rich repeat motif in ribonucleoproteins. Nucleic Acids Res. 21(3):439-46.

De Andrade HH, Marques EK, Schenberg AC, Henriques JA. (1989) The PSO4 gene is responsible for an error-prone recombinational DNA repair pathway in Saccharomyces cerevisiae. Mol Gen Genet. 217(2-3):419-26.

Deckert J, Hartmuth K, Boehringer D, Behzadnia N, Will CL, Kastner B, Stark H, Urlaub H, Luhrmann R. (2006) Protein composition and electron microscopy structure of affinity-purified human spliceosomal B complexes isolated under physiological conditions. Mol Cell Biol. 26(14):5528-43.

Dempsey LA, Sun H, Hanakahi LA, Maizels N. (2001) G4 DNA binding by LR1 and its subunits, nucleolin and hnRNP D, A role for G-G pairing in immunoglobulin switch recombination. J Biol Chem. 274(2):1066-71.

Depienne C, Heron D, Betancur C, Benyahia B, Trouillard O, Bouteiller D, Verloes A, Leguern E, Leboyer M, Brice A. (2007) Autism, language delay and mental retardation in a patient with 7q11 duplication. J Med Genet. 2007 Mar 30; [Epub ahead of print]

Diep CQ, Peng G, Bewley M, Pilauri V, Ropson I, Hopper JE. (2006) Intragenic suppression of Gal3C interaction with Gal80 in the Saccharomyces cerevisiae GAL gene switch. Genetics. 172(1):77-87.

Doll A, Grzeschik KH. (2001) Characterization of two novel genes, WBSCR20 and WBSCR22, deleted in Williams-Beuren syndrome. Cytogenet Cell Genet. 95(1-2):20-7.

Dreyfuss G, Kim VN, Kataoka N. (2002) Messenger-RNA-binding proteins and the messages they carry. Nat Rev Mol Cell Biol. 3(3):195-205. Review.

Estojak J, Brent R, Golemis EA. (1995) Correlation of two-hybrid affinity data with in vitro measurements. Mol Cell Biol. 15(10):5820-9.

Fields S, Song O. (1989) A novel genetic system to detect protein-protein interactions. Nature 340(6230):245-6.

Fiset S, Chabot B. (2001) hnRNP A1 may interact simultaneously with telomeric DNA and the human telomerase RNA in vitro. Nucleic Acids Res. 29(11):2268-75.

Gattoni R, Mahe D, Mahl P, Fischer N, Mattei MG, Stevenin J, Fuchs JP. (1996) The human hnRNP-M proteins: structure and relation with early heat shock-induced splicing arrest and chromosome mapping. Nucleic Acids Res. 24(13):2535-42.

Grillari J, Ajuh P, Stadler G, Loscher M, Voglauer R, Ernst W, Chusainow J, Eisenhaber F, Pokar M, Fortschegger K, Grey M, Lamond AI, Katinger H. (2005) SNEV is an evolutionarily conserved splicing factor whose oligomerization is necessary for spliceosome assembly. Nucleic Acids Res. 33(21):6868-83.

Grillari J, Hohenwarter O, Grabherr RM, Katinger H. (2000) Subtractive hybridization of mRNA from early passage and senescent endothelial cells. Exp Gerontol. 35(2):187-97.

Groenen PM, Vanderlinden G, Devriendt K, Fryns JP, Van de Ven WJ. (1998) Rearrangement of the human CDC5L gene by a t(6;19)(p21;q13.1) in a patient with multicystic renal dysplasia. Genomics 49(2):218-29.

Habelhah H, Shah K, Huang L, Ostareck-Lederer A, Burlingame AL, Shokat KM, Hentze MW, Ronai Z. (2001) ERK phosphorylation drives cytoplasmic accumulation of hnRNP-K and inhibition of mRNA translation. Nat Cell Biol. 3(3):325-30.

Hatakeyama S, Yada M, Matsumoto M, Ishida N, Nakayama KI. (2001) U box proteins as a new family of ubiquitin-protein ligases. J Biol Chem. 276(35):33111-20.

Jones EW, Fink GR. (1982) Regulation of amino acid and nucleotide biosynthesis in yeast. Pp.181-299 in The Molecular Biology of the Yeast Saccharomyces: Metabolism and Gene Expression, edited by Strathern JN, Jones EW and Broach JR. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press.

Jurica MS, Licklider LJ, Gygi SR, Grigorieff N, Moore MJ. (2002) Purification and characterization of native spliceosomes suitable for three-dimensional structural analysis. RNA 8(4):426-39.

Kafasla P, Patrinou-Georgoula M, Lewis JD, Guialis A. (2002) Association of the 72/74-kDa proteins, members of the heterogeneous nuclear ribonucleoprotein M group, with the pre-mRNA at early stages of spliceosome assembly. Biochem J. 363(Pt 3):793-9.

Kafasla P, Patrinou-Georgoula M, Guialis A. (2000) The 72/74-kDa polypeptides of the 70-110 S large heterogeneous nuclear ribonucleoprotein complex (LH-nRNP) represent a discrete subset of the hnRNP M protein family. Biochem J. 350 Pt 2:495-503.

Lei XH, Shen X, Xu XQ, Bernstein HS. (2000) Human Cdc5, a regulator of mitotic entry, can act as a site-specific DNA binding protein. J Cell Sci. 113 Pt 24:4523-31.

Lewin B. (2000) Genes, 7th ed. Oxford: Oxford University Press.

Löscher M, Fortschegger K, Ritter G, Wostry M, Voglauer R, Schmid JA, Watters S, Rivett AJ, Ajuh P, Lamond AI, Katinger H, Grillari J. (2005) Interaction of U-box E3 ligase SNEV with PSMB4, the beta7 subunit of the 20 S proteasome. Biochem J. 388(2):593-603.

Lu X, Legerski RJ. (2007) The Prp19/Pso4 core complex undergoes ubiquitylation and structural alterations in response to DNA damage. Biochem Biophys Res Commun. 354(4):968-74.

Marizzi Ch. (2007) Interactions within the CDC5L Complex via Yeast Two Hybrid System and Screening for Novel Interacting Partners of NOPSI. Wien: Diplomarbeit Universität Wien.

Mayeda A, Krainer AR. (1992) Regulation of alternative pre-mRNA splicing by hnRNP A1 and splicing factor SF2. Cell. 68(2):365-75.

McDonald WH, Ohi R, Smelkova N, Frendewey D, Gould KL. (1999) Myb-related fission yeast cdc5p is a component of a 40S snRNP-containing complex and is essential for pre-mRNA splicing. Mol Cell Biol. 19(8):5352-62.

Merla G, Ucla C, Guipponi M, Reymond A. (2002) Identification of additional transcripts in the Williams-Beuren syndrome critical region. Hum Genet. 110(5):429-38.

Michelotti EF, Michelotti GA, Aronsohn AI, Levens D. (1996) Heterogeneous nuclear ribonucleoprotein K is a transcription factor. Mol Cell Biol. 16(5):2350-60.

Morris CA, Demsey SA, Leonard CO, Dilts C, Blackburn BL. (1988) Natural history of Williams syndrome: physical characteristics. J Pediatr. 113(2):318-26.

Nemeth K, Salchert K, Putnoky P, Bhalerao R, Koncz-Kalman Z, Stankovic-Stangeland B, Bako L, Mathur J, Okresz L, Stabel S, Geigenberger P, Stitt M, Redei GP, Schell J, Koncz C. (1998) Pleiotropic control of glucose and hormone responses by PRL1, a nuclear WD protein, in Arabidopsis. Genes Dev. 12(19):3059-73.

Neubauer G, King A, Rappsilber J, Calvio C, Watson M, Ajuh P, Sleeman J, Lamond A, Mann M. (1998) Mass spectrometry and EST-database searching allows characterization of the multi-protein spliceosome complex. Nat Genet. 20(1):46-50.

Ohi R, McCollum D, Hirani B, Den Haese GJ, Zhang X, Burke JD, Turner K, Gould KL. (1994) The Schizosaccharomyces pombe cdc5+ gene encodes an essential protein with homology to c-Myb. EMBO J. 13(2):471-83.

Piganeau N, Schroeder R. (2006) Identification and detection of RNA-RNA interactions using the yeast RNA hybrid system. Nat Protoc. 1(2):689-94.

Potashkin J, Kim D, Fons M, Humphrey T, Frendewey D. (1998) Cell-division-cycle defects associated with fission yeast pre-mRNA splicing mutants. Curr Genet. 34(3):153-63.

Putz U, Skehel P, Kuhl D. (1996) A tri-hybrid system for the analysis and detection of RNA--protein interactions. Nucleic Acids Res. 24(23):4838-40.

Schörgenhumer F. (2005) Screening for protein-protein interacting partners of SKIP, a component of the human spliceosome. Wien: Diplomarbeit Universität für Bodenkultur Wien.

Smith TF, Gaitatzes C, Saxena K, Neer EJ. (1999) The WD repeat: a common architecture for diverse functions. Trends Biochem Sci. 24(5):181-5. Review.

Sobhanifar S. (2003) Yeast Two Hybrid Assay: A Fishing Tale. BioTeach Journal 1:81-87

Tarn WY, Hsu CH, Huang KT, Chen HR, Kao HY, Lee KR, Cheng SC. (1994) Functional association of essential splicing factor(s) with PRP19 in a protein complex. EMBO J. 13(10):2421-31.

Van Criekinge W, Beyaert R. (1999) Yeast Two-Hybrid: State of the Art. Biol Proced Online. 2:1-38.

Vertegaal AC, Ogg SC, Jaffray E, Rodriguez MS, Hay RT, Andersen JS, Mann M, Lamond AI. (2004) A proteomic study of SUMO-2 target proteins. J Biol Chem. 279(32):33791-8.

Voglauer R, Chang MW, Dampier B, Wieser M, Baumann K, Sterovsky T, Schreiber M, Katinger H, Grillari J. (2006) SNEV overexpression extends the life span of human endothelial cells. Exp Cell Res. 312(6):746-59.

Zhang N, Kaur R, Lu X, Shen X, Li L, Legerski RJ. (2005) The Pso4 mRNA splicing and DNA repair complex interacts with WRN for processing of DNA interstrand cross-links. J Biol Chem. 280(49):40559-67.

Zhang J, Lautar S. (1996) A yeast three-hybrid method to clone ternary protein complex components. Anal Biochem. 242(1):68-72.

Zonneveld BJ, van der Zanden AL. (1995) The red ade mutants of Kluyveromyces lactis and their classification by complementation with cloned ADE1 or ADE2 genes from Saccharomyces cerevisiae. Yeast 11(9):823-7.

Abbrevations

bp – Base pair(s)

cfu – Colony forming units

CTD – Carboxy-terminal domain of RNA polymerase II

ddNTP – 2'-3'-di-deoxyribonucleotide

DMSO – Dimethyl sulfoxide

dNTP – Deoxynucleotidetriphosphate

DO – Drop out

EDTA – Ethylenediamine tetraacetic acid

Ethidium bromide – 3,8-diamino-5-ethyl-6-phenylphenanthridinium bromide

Hepes – 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid

kbp – Kilo base pairs

NB – Nota bene

LiAc – Lithium acetate

mut - Mutant

nt – Nucleotide(s)

OD – Optical density

OD 600 – Optical density determined at 600 nm

PEG – Polyethylene glycol

RO – Reverse osmosis

rpm – Rotations per minute

RRM – RNA recognition motif

TAE – Tris-Acetate-EDTA

TE – Tris-EDTA

Tris – 2-amino-2-hydroxymethyl-1,3-propanediol (trishydroxymethylaminomethane)

UAS – Upstream activation sequence

wt – Wild type

YNB – Yeast nitrogen base

Appendix A – All Yeast Two-Hybrid interaction domain mapping results

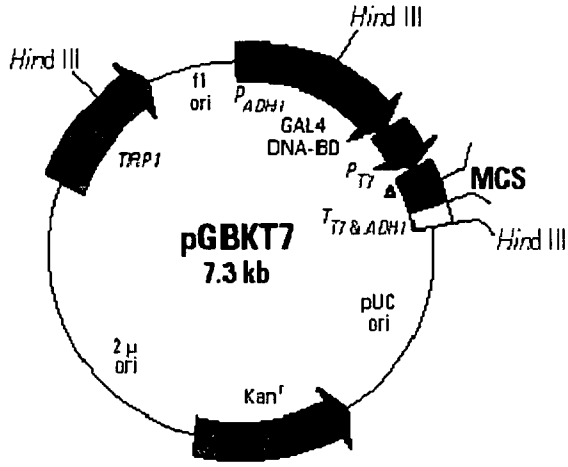
DNA binding domain fusion protein		Activation domain fusion protein	Run 1 b	Run 1 SE	Run 1 Interaction	Run 2 b	Run 2 SE	Run 2 Interaction	Run 3 b	Run 3 SE	Run 3 Interaction	Run 4 b	Run 4 SE	Run 4 Interaction
hnRNP M wild type	+	PLRG1	2,002	+/- 0,053	+	1,303	+/- 0,014	-	1,228	+/- 0,041	-			
hnRNP M mutant 1	+	PLRG1				1,315	+/- 0,022	-	1,126	+/- 0,012	-	1,277	+/- 0,043	-
hnRNP M mutant 2	+	PLRG1							1,328	+/- 0,018	-	1,225	+/- 0,025	-
hnRNP M mutant 3	+	PLRG1				3,633	+/- 0,492	++	2,492	+/- 0,103	+	1,220	+/- 0,117	-
hnRNP M mutant 4	+	PLRG1							3,186	+/- 0,127	++	2,098	+/- 0,143	+
hnRNP M mutant 4	+	empty vector												
empty vector	+	PLRG1	1,185	+/- 0,032	-									
empty vector	+	empty vector				1,360	+/- 0,017	-	1,196	+/- 0,018	-	1,278	+/- 0,019	-

DNA binding domain fusion protein		Activation domain fusion protein	Run 5 b	Run 5 SE	Run 5 Interaction	Run 6 b	Run 6 SE	Run 6 Interaction	Run 7 b	Run 7 SE	Run 7 Interaction	Run 8 b	Run 8 SE	Run 8 Interaction
hnRNP M mutant 3	+	PLRG1	3,884	+/- 0,157	++									
hnRNP M mutant 4	+	PLRG1	5,223	+/- 0,313	++									
hnRNP M mutant 4	+	empty vector				2,071	+/- 0,355	+	1,656	+/- 0,111	-	4,382	+/- 1,184	++
hnRNP M wild type	+	CDC5L	1,491	+/- 0,110	-									
hnRNP M mutant 1	+	CDC5L				1,510	+/- 0,014	-						
hnRNP M mutant 2	+	CDC5L	1,499	+/- 0,020	-									
hnRNP M mutant 3	+	CDC5L				1,609	+/- 0,055	-	2,600	+/- 0,094	+			
hnRNP M mutant 4	+	CDC5L	11,209	+/- 1,621	++	7,232	+/- 0,266	++	2,982	+/- 0,199	+			
empty vector	+	empty vector	1,538	+/- 0,028	-	1,361	+/- 0,025	-	1,280	+/- 0,046	-	1,088	+/- 0,021	-

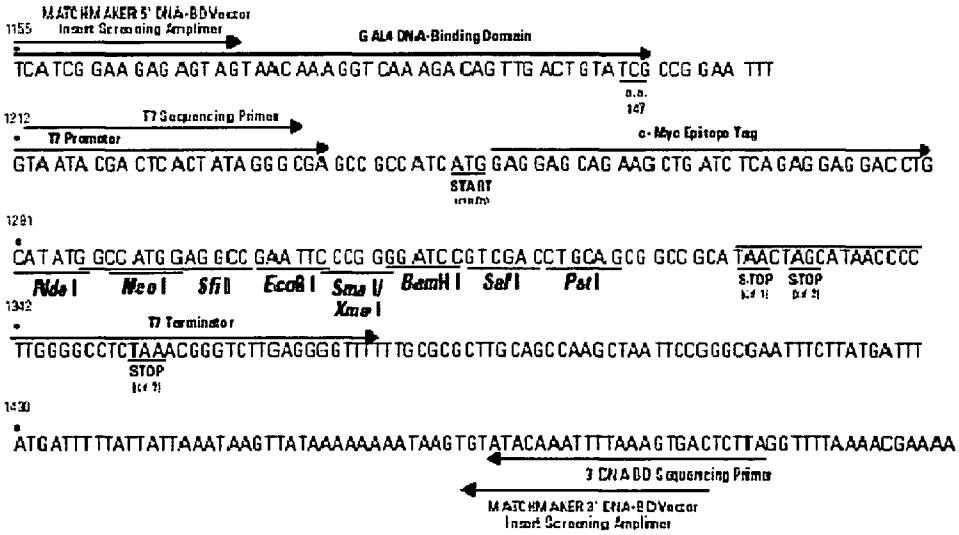
N = 12 for all samples with the exception of hnRNP M mutant 4 + PLRG1 in Run 2, for which n = 4.

Appendix B – Vector maps

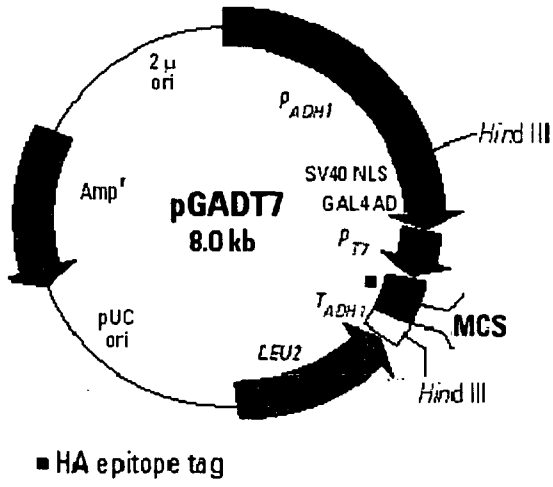
pGBKT7

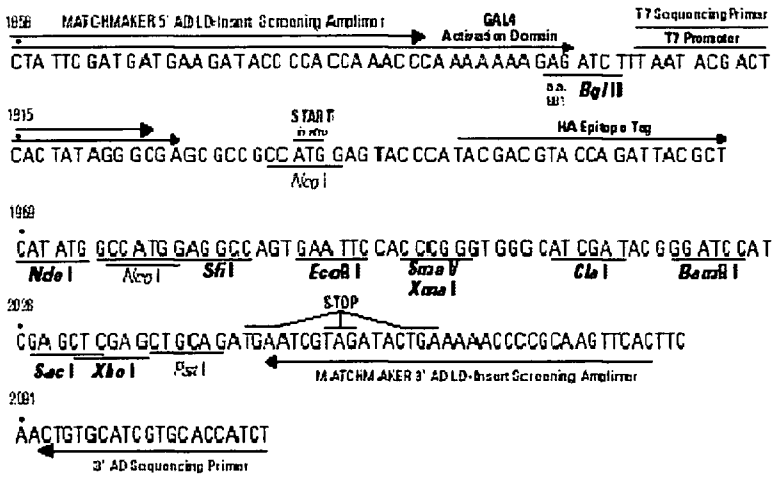


▲ c-Myc epitope tag



pGADT7


2 μ ori
P_{ADH1}
Hind III
SV40 NLS
GAL4 AD
Amp^r
pGADT7
8.0 kb
P_{T7}
T_{ADH1}
MCS
pUC ori
LEU2
Hind III
■ HA epitope tag


1858
MATCHMAKER 5' AD LD-Insert Screening Amplimer
GAL4 Activation Domain
T7 Sequencing Primer
T7 Promoter
CTA TTC GAT GAT GAA GAT ACC CCA CCA AAC CCA AAA AAA GAG ATC TTT AAT ACG ACT
Bgl II
1915
START
HA Epitope Tag
CAC TAT AGG GCG AGC GCC GCC ATG GAG TAC CCA TAC GAC GTA CCA GAT TAC GCT
Nco I
1969
CAT ATG GCC ATG GAG GCC AGT GAA TTC CAC CCG GGT GGG CAT CGA TAC GGG ATC CAT
Nde I
Nco I
Sfi I
EcoR I
Sma I
Xma I
Cla I
BamH I
2026
STOP
CGA GCT CGA GCT GCA GATGAATCGTAGATACTGAAAAACCCCGCAAGTTCACTTC
Sac I
Xho I
Pst I
MATCHMAKER 3' AD LD-Insert Screening Amplimer
2091
AACTGTGCATCGTGCACCATCT
3' AD Sequencing Primer

Appendix C – Determined coding sequences

hnRNP M splicing variant 2 (2076 nucleotides)

ATGGCGGCAGGGGTCGAAGCGGCGGCGGAGGTGGCGGCGACGGAGATCAAAATGGA
GGAAGAGAGCGGCGCGCCCGGCGTGCCGAGCGGCAACGGGGCTCCGGGCCCTAAGG
GTGAAGGAGAACGACCTGCTCAGAATGAGAAGAGGAAGGAGAAAAACATAAAAAGAGG
AGGCAATCGCTTTGAGCCATATGCCAATCCAACTAAAAGATACAGAGCCTTCATTACAAA
CATACCTTTTGATGTGAAATGGCAGTCACTTAAAGACCTGGTTAAAGAAAAAGTTGGTGA
GGTAACATACGTGGAGCTCTTAATGGACGCTGAAGGAAAGTCAAGGGGATGTGCTGTT
GTTGAATTCAAGATGGAAGAGAGCATGAAAAAAGCTGCGGAAGTCCTAAACAAGCATAG
TCTGAGCGGAAGACCACTGAAAGTCAAAGAAGATCCTGATGGTGAACATGCCAGGAGA
GCAATGCAAAAGGCTGGAAGACTTGGAAGCACAGTATTTGTAGCAAATCTGGATTATAA
AGTTGGCTGGAAGAAACTGAAGGAAGTATTTAGTATGGCTGGTGTGGTGGTCCGAGCA
GACATTCTTGAAGATAAAGATGGAAAAAGTCGTGGAATAGGCACTGTTACTTTTGAACAG
TCCATTGAAGCTGTGCAAGCTATATCTATGTTCAATGGCCAGCTGCTATTTGATAGACCA
ATGCACGTCAAGATGGATGAGAGGGCCTTACCAAAAGGAGATTTCTTCCCTCCTGAGCG
TCCACAACAACTTCCCCATGGCCTTGGTGGTATTGGCATGGGGTTAGGACCAGGAGGG
CAACCCATTGATGCCAATCACCTGAATAAAGGCATCGGAATGGGAAACATAGGTCCCGC
AGGAATGGGAATGGAAGGCATAGGATTTGGAATAAATAAAATGGGAGGAATGGAGGGG
CCCTTTGGTGGTGGTATGGAAAACATGGGTCGATTTGGATCTGGGATGAACATGGGCA
GGATAAATGAAATCCTAAGTAATGCACTGAAGAGAGGAGAGATCATTGCAAAGCAGGGA
GGAGGTGGAGGTGGAGGAAGCGTCCCTGGGATCGAGAGGATGGGTCCTGGCATTGAC
CGCCTCGGGGGTGCCGGCATGGAGCGCATGGGCGCGGGCCTGGGCCACGGCATGGA
TCGCGTGGGCTCCGAGATCGAGCGCATGGGCCTGGTCATGGACCGCATGGGCTCCGT
GGAGCGCATGGGCTCCGGCATTGAGCGCATGGGCCCGCTGGGCCTCGACCACATGGC
CTCCAGCATTGAGCGCATGGGCCAGACCATGGAGCGCATTGGCTCTGGCGTGGAGCG
CATGGGTGCCGGCATGGGCTTCGGCCTTGAGCGCATGGCCGCTCCCATCGACCGTGT
GGGCCAGACCATTGAGCGCATGGGCTCTGGCGTGGAGCGCATGGGCCCTGCCATCGA
GCGCATGGGCCTGAGCATGGAGCGCATGGTGCCCGCAGGTATGGGAGCTGGCCTGGA
GCGCATGGGCCCCGTGATGGATCGCATGGCCACCGGCCTGGAGCGCATGGGCGCCAA
CAATCTGGAGCGGATGGGCCTGGAGCGCATGGGCGCCAACAGCCTCGAGCGCATGGG
CCTGGAGCGCATGGGTGCCAACAGCCTCGAGCGCATGGGCCCCGCCATGGGCCCGGC
CCTGGGCGCTGGCATTGAGCGCATGGGCCTGGCCATGGGTGGCGGTGGCGGTGCCA
GCTTTGACCGTGCCATCGAGATGGAGCGTGGCAACTTCGGAGGAAGCTTCGCAGGTTC
CTTTGGTGGAGCTGGAGGCCATGCTCCTGGGGTGGCCAGGAAGGCCTGCCAGATATTT

GTGAGAAATCTGCCATTCGATTTCACATGGAAGATGCTAAAGGACAAATTCAACGAGTG
CGGCCACGTGCTGTACGCCGACATCAAGATGGAGAATGGGAAGTCCAAGGGGTGTGGT
GTGGTTAAGTTCGAGTCGCCAGAGGTGGCCGAGAGAGCCTGCCGGATGATGAATGGCA
TGAAGCTGAGTGGCCGAGAGATTGACGTTCGAATTGATAGAAACGCTTAA

CDC5L (2409 nucleotides)

ATGCCTCGAATTATGATCAAGGGGGGCGTATGGAGGAATACCGAGGATGAAATTCTGAA
AGCAGCGGTAATGAAATATGGGAAAAATCAGTGGTCTAGGATTGCCTCATTGCTGCATA
GAAAATCAGCAAAGCAGTGCAAAGCCAGATGGTATGAATGGCTGGATCCAAGCATTAAG
AAGACAGAATGGTCCAGAGAAGAAGAGGAAAAACTCTTGCACTTGGCCAAGTTGATGCC
AACTCAGTGGAGGACCATTGCTCCAATCATTGGAAGAACAGCGGCCCAGTGCTTAGAAC
ACTATGAATTTCTTCTGGATAAAGCTGCCCAAAGAGACAATGAAGAGGAAACAACAGAT
GATCCACGAAAACTTAAACCTGGAGAAATAGATCCAAATCCAGAAACAAAACCAGCGCG
GCCTGATCCAATTGATATGGATGAGGATGAACTTGAGATGCTTTCTGAAGCCAGAGCCC
GCTTGGCTAATACTCAGGGAAAGAAGGCCAAGAGGAAAGCAAGAGAGAAACAATTGGA
AGAAGCAAGACGTCTTGCTGCCCTCCAAAAAAGAAGAGAACTTCGAGCAGCTGGCATA
GAAATTCAGAAGAAAAGAAAAAGGAAGAGAGGAGTTGATTATAATGCCGAAATCCCATT
TGAAAAAAAGCCTGCCCTTGGTTTTTATGATACTTCTGAGGAAAACTACCAAGCTCTTGA
CGCAGATTTCAGGAAATTAAGACAACAGGATCTTGATGGGGAGCTAAGATCTGAAAAAG
AAGGAAGAGATAGAAAAAAAGACAAACAGCATTTGAAAAGGAAAAAAGAATCTGATTTAC
CATCAGCTATTCTTCAAACTAGTGGTGTTTCTGAATTTACTAAAAAGAGAAGCAAACTAG
TACTTCCTGCCCCTCAGATTTCAGATGCAGAACTCCAGGAAGTTGTAAAAGTAGGCCAA
GCGAGTGAAATTGCACGTCAAACTGCCGAGGAATCTGGCATAACAAATTCTGCTTCCAG
TACACTTTTGTCTGAGTACAATGTCACCAACAACAGCGTTGCTCTTAGAACACCACGAAC
ACCAGCTTCCCAGGACAGAATTCTGCAGGAAGCCCAGAACCTCATGGCCCTCACCAAT
GTGGACACCCCATTGAAAGGTGGACTTAATACCCCATTGCATGAGAGTGACTTCTCAGG
TGTAACTCCACAGCGACAAGTTGTACAGACTCCAAACACAGTTCTCTCTACTCCATTCAG
GACTCCTTCTAATGGAGCTGAAGGGCTGACTCCCCGGAGTGGAACAACTCCCAAACCA
GTTATTAACTCTACTCCGGGTAGAACTCCTCTTCGAGACAAGTTAAACATTAATCCCGAG
GATGGAATGGCAGACTATAGTGATCCCTCTTACGTGAAGCAGATGGAAAGAGAATCCCG
AGAACATCTCCGTTTAGGGTTGTTGGGCCTTCCTGCCCCTAAGAATGATTTTGAAATTGT
TCTACCAGAAAATGCCGAGAAGGAGCTGGAAGAACGTGAAATAGATGATACTTACATTG
AAGATGCTGCTGATGTGGATGCTCGAAAGCAGGCCATACGAGATGCAGAGCGCGTAAA
GGAAATGAAACGAATGCATAAAGCTGTCCAGAAAGATCTGCCAAGACCATCAGAAGTAA
ATGAAACTATTCTAAGACCCTTAAATGTAGAACCGCCTTTAACAGATTTACAGAAAAGTG
AAGAACTAATCAAAAAAGAAATGATCACAATGCTTCATTATGACCTTCTACATCACCCTTA

TGAACCATCTGGAAATAAAAAAGGCAAAACTGTAGGGTTTGGTACCAATAATTCAGAGCA
CATTACCTACCTGGAACATAATCCTTATGAAAAGTTCTCCAAAGAAGAGCTGAAAAAGGC
CCAGGATGTTTTGGTGCAGGAGATGGAAGTGGTTAAACAAGGAATGAGCCATGGAGAG
CTCTCAAGTGAAGCTTATAACCAGGTGTGGGAAGAATGCTACAGTCAAGTTTTATATCTT
CCTGGGCAGAGCCGCTACACACGGGCCAATCTGGCTAGTAAAAAGGACAGAATTGAAT
CACTTGAAAAGAGGCTCGAGATAAACAGGGGTCACATGACGACAGAAGCCAAGAGGGC
TGCAAAGATGGAAAAGAAGATGAAAATTTTGCTTGGGGGTTACCAGTCTCGTGCTATGG
GGCTCATGAAACAGTTGAATGACTTATGGGACCAAATTGAACAGGCTCACTTGGAGTTA
CGCACTTTTGAAGAACTCAAGAAACATGAAGATTCTGCTATTCCCCGGAGGCTAGAGTG
TCTAAAAGAAGACGTTCAGCGACAACAAGAAAGAGAAAAGGAACTTCAACATAGATATG
CTGATTTGCTGCTGGAGAAAGAGACTTTAAAGTCAAAATTCTGA

PLRG1 (1548 nucleotides)

ATGGTCGAGGAGGTACAGAAACATTCTGTACACACCCTTGTGTTCAGGTCGTTGAAGAG
GACCCATGACATGTTTGTAGCTGATAATGGAAAACCTGTGCCTTTAGATGAAGAGAGTC
ACAAACGAAAAATGGCAATCAAGCTTCGTAATGAGTATGGTCCTGTGTTGCATATGCCTA
CTTCAAAAGAAAATCTTAAAGAGAAGGGTCCTCAGAATGCAACGGATTCATATGTTCATA
AACAGTACCCTGCCAATCAAGGACAAGAAGTTGAATACTTTGTGGCAGGTACACATCCA
TACCCACCAGGACCTGGGGTTGCTTTGACAGCAGATACTAAGATCCAGAGAATGCCAAG
TGAATCAGCTGCACAGTCCTTAGCGGTGGCATTACCTTTGCAGACCAAGGCTGATGCAA
ATCGTACTGCCCCTAGTGGAAGTGAATACCGACATCCTGGGGCTTCTGACCGTCCACA
GCCTACAGCGATGAATTCAATTGTCATGGAGACTGGCAATACCAAGAACTCTGCACTGA
TGGCTAAAAAAGCCCCTACAATGCCAAAACCCCAGTGGCACCCACCGTGGAAACTCTAC
AGGGTTATCAGTGGGCATCTTGGCTGGGTTCGATGTATTGCTGTGGAACCTGGAAATCA
GTGGTTTGTTACTGGATCTGCTGACAGAACTATAAAGATCTGGGACTTGGCTAGTGGCA
AATTAAAGCTGTCATTGACTGGGCATATTAGTACTGTGCGGGGCGTGATAGTAAGCACA
AGGAGCCCATATCCGTTCTCTTGTGGAGAAGACAAACAAGTGAAATGCTGGGATCTCGA
ATACAATAAGGTTATACGGCATTATCATGGACATTTAAGTGCAGTGTATGGTCTGGATTT
GCACCCGACAATCGATGTGTTGGTAACCTGTAGTCGAGATTCAACGGCACGGATTTGG
GATGTGAGAACTAAAGCCAGTGTACACACATTATCTGGACATACAAATGCAGTTGCTACA
GTGAGATGTCAGGCTGCAGAACCACAGATTATTACAGGAAGCCATGATACTACAATTCG
ATTATGGGATCTGGTGGCTGGAAAAACAAGAGTGACATTAACAAATCACAAAAAATCAGT
TAGGGCTGTGGTTTTACATCCAAGACATTACACATTTGCATCTGGTTCTCCAGATAACAT
AAAGCAGTGGAAATTCCCTGATGGAAGTTTCATACAAAATCTTTCCGGTCATAATGCTAT
TATTAACACATTGACGGTAAATTCTGATGGAGTGCTTGTATCTGGAGCTGACAATGGCAC
CATGCATCTTTGGGACTGGAGAACTGGCTACAATTTTCAGAGAGTTCACGCAGCTGTGC

AACCTGGGTCTTTGGACAGTGAATCAGGAATATTTGCTTGTGCTTTTGATCAGTCTGAAA
GTCGATTACTAACAGCTGAAGCTGATAAAACCATTAAAGTATACAGAGAGGATGACACA
GCCACAGAAGAAACTCATCCAGTCAGCTGGAAACCAGAAATTATCAAGAGAAAGAGATT
TTAATGA

Appendix D – Primer sequences

PCR primers

Restriction sites are written in capital letters. Start codons and anti-sense stop codons are underlined.

hnRNP M sense start ATG BamH I
agaGGATCCctatggcggcaggggtc

hnRNP M ATG 155 (var2) sense BamH I
agaGGATCCctatgagagcaatgcaaaaggctg

hnRNP M ATG 166 (var2) sense BamH I
agaGGATCCctatgacagtatttgtagcaaatctgg

hnRNP M ATG 287 sense BamH I
agaGGATCCctatggatttcttccctcctgagcgt

hnRNP M ATG 384 sense BamH I
agaGGATCCctatggagatcattgcaaagcag

hnRNP M ATG 583 (var2) sense BamH I
agaGGATCCctatggccatcgagatggagc

hnRNP M ATG 616 (var2) sense BamH I
agaGGATCCctatgtttgtgagaaatctgccattcg

hnRNP M stop Pst I anti-sense
aagaCTGCAGttaagcgtttctatcaattcg

hnRNP M stop 145 (var2) Pst I anti-sense
aagtCTGCAGttacttgactttccttcagcgtccat

hnRNP M stop 154 (var2) Pst I anti-sense
aagtCTGCAGttacctggcacattcaccatc

hnRNP M stop 287 Pst I anti-sense
aagtCTGCAGttaatctccttttggtaaggccctct

hnRNP M stop 622 Pst I anti-sense
aagtCTGCAGttaggcacggtcaaagctggca

hnRNP M stop 651 Pst I anti-sense
aagaCTGCAGttaggcacggtcaaagctggca

CDC5L sense start Nde I
aagaCATATGatgcctcgaattatgatcaag

CDC5L stop Sma I anti-sense
atgtCCCGGGtcagaattttgacttta

PLRG1 sense start BamH I
agaaGGATCCatatggtcgaggaggtacagaaac

PLRG1 stop Xho I anti-sense
agatCTCGAGtcattaaaatctctttctcttg

DNA sequencing primers

hnRNP M start nt 1273
cacatggcctccagcattg

CDC5L start nt 500
ccaagaggaaagcaagagag

CDC5L start nt 1001
ccagtacacttttgtctgag

CDC5L start nt 1479
gaaggagctggaagaacgtg

CDC5L start nt 1999
caagttttatatcttcctggg

PLRG1 start nt 498
catggagactggcaatacc

PLRG1 start nt 1000
gctacagtgagatgtcagg

Vector primers

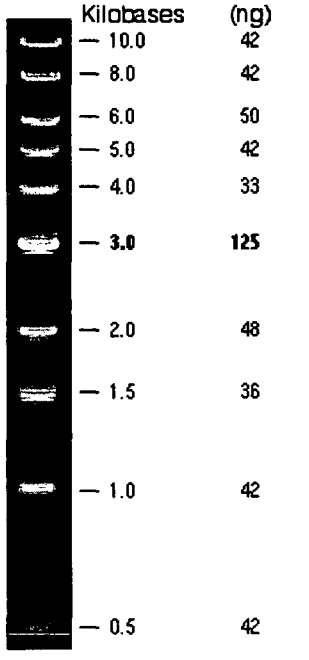
T7 primer
taatacgactcactatagggc

3'AD primer
gacacgtagcacgtggtaga

3'DNA-BD primer
tatgtttaaaatttcactgagaat

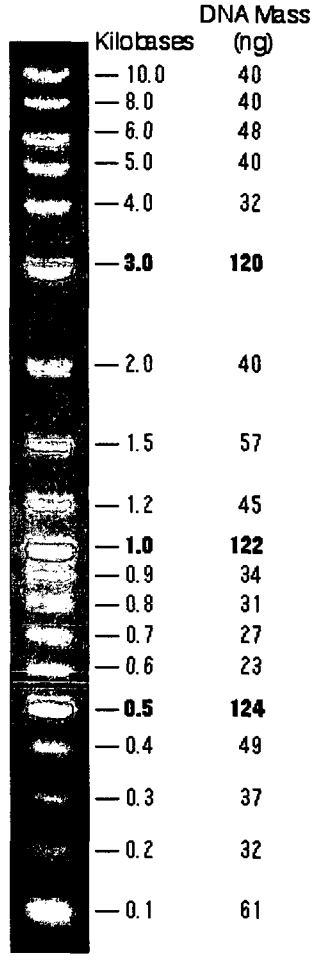
Appendix E – DNA standards

NEB 1 kb DNA Ladder



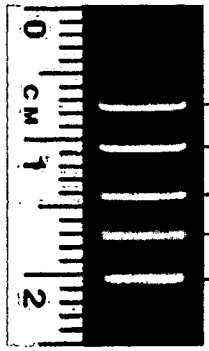
0.5 µg of 1 kb DNA Ladder visualized by ethidium bromide staining on a 0.8% TAE agarose gel.

NEB TriDye 2-Log



1.0 µg of 2-Log DNA Ladder visualized by ethidium bromide staining on a 1.0% TBE agarose gel.

Fermentas FastRuler™ DNA Ladder Middle Range



bp	ng/20µl	ng/15µl	ng/10µl	ng/5µl	ng/3µl
5000	80	60	40	20	12
2000	80	60	40	20	12
850	80	60	40	20	12
400	80	60	40	20	12
100	80	60	40	20	12

20µl/lane, 1X TAE, 7V/cm, 14min
1.0% agarose

Fermentas O'Gene Ruler DNA Ladder Mix

bp	ng/0.5µg	%
10000	10.8	2.2
8000	13.0	2.6
6000	16.2	3.2
5000	16.2	3.2
4000	21.6	4.3
3500	27.5	5.5
3000	**69.0**	13.8
2500	18.0	3.6
2000	**46.0**	9.2
1500	23.9	4.8
1200	16.2	3.2
1031	**29.6**	5.9
900	21.1	4.2
800	18.8	3.8
700	16.5	3.3
600	14.1	2.8
500	**41.7**	8.3
400	19.8	4.0
300	20.0	4.0
200	20.0	4.0
100	20.0	4.0

0.5µg/lane,
8cm length gel,
1X TAE, 7V/cm, 45min

Appendix F – Media formulations

I Bacteria cultivation and transformation

LB Medium

Peptone	10 g/l
Yeast-Extract	5 g/l
NaCl	10 g/l
Agar (for plates only)	20 g/l

Adjust pH to 7 with NaOH.
Autoclave. Cool to 50 °C.

Ampicillin (if required)		100 µg/ml
Kanamycin	(if required)	50 µg/ml

Hanahan's SOC medium

Peptone	2%
Yeast Extract	0.5 %
NaCl	10 mM
KCl	2.5 mM
$MgCl_2$	10 mM
$MgSO_4$	10 mM
Glucose	20 mM

Mix Peptone, Yeast Extract, NaCl and KCl. Adjust pH to 7 with NaOH and autoclave. Autoclave or sterile filtrate $MgCl_2$, $MgSO_4$ and glucose solutions separately and then mix the sterile solutions.

II DNA preparations

TE Buffer

Tris	10 mM
EDTA	1 mM

Adjust pH with HCl.

50x TAE Buffer (1 l)

Tris	242 g
Acetic Acid	57.1 ml
$Na_2EDTA{\cdot}2H_2O$	37.2 g
RO-H2O	to 1 l

1 % Agarose Gel (1 gel slide)

Agarose	1.2 g
50x TAE Buffer	2.4 ml
RO-H_2O	117.6 ml

Melt in microwave oven. Cool to 50 °C.

Ethidium bromide	30 µg (final concentration is 0.25 µg/ml)

TAE Ethidium Bromide Buffer (5 l)

50x TAE Buffer	100 ml
RO-H_2O	4900 ml
Ethidium bromid	1.5 mg

III Yeast cultivation and transformation

YPDA Medium (1 l)

Peptone	20 g
Yeast extract	10 g
Agar (for plates only)	20 g
RO-H_2O	to 950 ml

Autoclave. Cool to 50 °C.

40 % Glucose solution 50 ml (final concentration is 2 %)

0.2 % adenine hemisulfate solution 15 ml (final concentration is 0.003 %)

Glucose and adenine hemisulfate solution are autoclaved separately.

1.1 x TE/LiAc solution (10 ml)

10x TE Buffer pH 7.5	1.1 ml
1 M lithium acetate	1.1 ml
RO-H_2O	to 10 ml

PEG/LiAc solution (10 ml)

50 % PEG 3350	8 ml
10x TE Buffer pH 7.5	1 ml
1 M lithium acetate	1 ml

0.9 % NaCl Solution

NaCl	0.9 g
RO-H_2O	100 ml

IV Yeast Two-Hybrid Drop Out media

10x Glucose solution

Glucose	100 g
RO-H_2O	400 g

Autoclave.

10x Yeast nitrogen base (500 ml)

YNB with ammonium sulfate, without amino acids	33.5 g
RO-H_2O	500 ml

Sterile filtrate. Do not autoclave YNB.

10x Drop Out supplements (100 ml)

-Trp

-Trp DO supplement (Clontech)	740 mg
RO-H_2O	100 ml

-Leu

-Leu DO supplement (Clontech)	690 mg
RO-H_2O	100 ml

2xDO

-Leu, -Trp DO supplement (Clontech)	640 mg
RO-H_2O	100 ml

4xDO

-Leu, -Trp, -His, -Ade DO supplement (Clontech)	600 mg
RO-H_2O	100 ml

Autoclave 10x Drop Out supplements.

Synthetic Drop Out medium (300 ml)

Agar (for plates only)	6 g
RO-H_2O	210 ml

Autoclave.

Temper all solutions to 50 °C before mixing if agar was included.

10x Glucose solution	30 ml
10x Yeast nitrogen base	30 ml
10x Drop Out supplement	30 ml